TECHNISCHE UNIVERSITÄT MÜNCHEN
TUM Campus Straubing für Biotechnologie und Nachhaltigkeit

Perspectives on the Warehouse of the Future

Fabian Johannes Lorson

Vollständiger Abdruck der von der promotionsführenden Einrichtung TUM Campus Straubing für Biotechnologie und Nachhaltigkeit der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Wirtschafts- und Sozialwissenschaften (Dr. rer. pol.)

genehmigten Dissertation.

Vorsitz:	Prof. Dr. Andreas Pondorfer
Prüfer der Dissertation:	1. Prof. Dr. Alexander Hübner
	2. Prof. Dr. Johannes Fottner
	3. Prof. Dr. Sebastian Schiffels

Die Dissertation wurde am 12.08.2022 bei der Technischen Universität München eingereicht und von der promotionsführenden Einrichtung TUM Campus Straubing für Biotechnologie und Nachhaltigkeit am 02.03.2023 angenommen.

Contents

List of Figures

List of Tables

List of Abbreviations

AGV	Automated guided vehicles
AMR	Autonomous mobile robots
AR	Augmented reality
AS/RS	Automated storage and retrieval system
BS	Behavioral science
C	Consultant
DiD	Difference-in-Differences
FAPR	Fully autonomous picking robot
IBR	Intervention-based research
MARR	Minimum attractive rate of return
MIP	Mixed-integer program
OM	Operations management
OPS	Order picking solution

RQ	Research question
SKU	Stock keeping unit
SMR	Shelf-moving robot
SP	System provider
VR	Virtual reality
WO	Warehouse operator

1 Introduction

This doctoral thesis develops new perspectives for the warehouse of the future. It presents a research agenda and theoretical foundation for the optimization of human-machine interactions, introduces a goal-setting intervention for a semi-automated pick-to-light human-machine interaction, and finally develops a mathematical optimization model for the selection of the most suitable order picking solution (OPS).

In this first chapter, warehousing and its role in modern supply chains is explained (Section 1.1). Given the transformational character of automation and resulting human-machine interactions for warehousing, these topics are introduced in 1.2.

The remainder is organized as follows. Chapter 2 gives an overview on the three contributions (articles) that compose the main body of this dissertation. In this way, involved authors and status of publication are provided, while the purpose, methodology and findings of the contributions are summarized. Chapter 3 to Chapter 5 each contain one of the three articles. Finally, Chapter 6 synthesizes the findings and outlines areas of future research.

1.1 The role of warehousing in modern supply chains

Warehousing is the intermediate storage of physical goods between different stages of a supply chain. The basic functions of a warehouse are receiving and inspection, put away, order picking, packing, and shipping (De Koster et al., 2007). Bartholdi III and Hackman (2020) specify two main purposes of warehouses. First, they are crucial to satisfy customer demand by having supply on stock. Second, warehouses are required to consolidate products for cheaper transportation costs and higher customer service levels. In this way, warehouses form a critical part of a firm's logistic setup (De Koster et al., 2007), being responsible for more than 20% of total logistics cost (Rodrigue, 2020). Warehousing thus constitutes a critical research field within operations management (see Azadeh et al. (2019); Boysen et al. (2019, 2021); Fragapane et al. (2021); Jaghbeer et al. (2020); Vanheusden et al. (2022) for recent overviews).

In the last years, warehousing has experienced a substantial transformation from a cost center to a central component in the value proposition of firms. Not only the ever increasing volumes of e-commerce orders has put the efficient orchestration of warehousing operations at the focus of operations managers (Schiffer et al., 2022). Additionally, growing customer demands and delivery expectations are fueling the necessity to push products faster and cheaper through warehouses and supply chains. Hence, they gained a pivotal role to ensure an efficient and effective material flow, especially to create resilient supply chains when facing volatile markets or increasing customer expectations on product range, availability and lead times. To establish efficient warehousing operations, managers recently implemented a large variety of semi- and fully automated warehousing systems, often resulting in novel human-machine interactions (see, for example, Fottner et al. (2021)). We focus on these two topics due to their transformational character for warehousing operations in the following.

1.2 Automation, robotics, and human-machine interactions

Expanding automation and robotization has been the focal point of operations in the recent years (IFR, 2020). Enabled by advances in Internet of Things devices and artificial intelligence, coupled with the advent of new system providers and decreased price points, one surging change in operations evolved to be in the arena of warehousing. In fact, the size of the warehouse automation industry has been increasing by 12% annually between 2014 and 2019 (Statista, 2020; The Logistics iQ, 2020). This market growth goes along with an increasing number of automated and robotized warehousing solutions, especially for order picking. A search on an independent comparison platform delivers more than 200 results for warehousing robots from more than 80 different solution providers (Lots of Bots, 2022). But this may only be the beginning: Huge sums of venture capital investments span over the last years and continue to rise in an unprecedented magnitude and speed (see, for example, Forbes (2021); TechCrunch (2021).) It comes thus at no surprise that the global size of the warehouse automation market is expected to reach USD 41 billion within six years, with an average annual increase of 14% from 2022 to 2027 (The Logistics iQ, 2022). Automated warehousing systems are gaining this large momentum because they enable faster throughput times, reduced cost, higher pick quality, more efficient space utilization, improved ergonomics, and lower dependence on human workers to cope with the ongoing labor shortage (Azadeh et al., 2019; McKinsey & Company, 2021a; Pazour et al., 2014). Innovations in warehouse automation thus play a crucial part in delivering products efficiently and effectively throughout supply chains.

Despite these technological developments, human operators will still be necessary to fulfill operational activities. Humans have distinctive characteristics, skills and capabilities that robots are not able to replicate or perform cost efficiently. For instance, they excel in flexibility when swift reactions are needed to volatility of the picking workload (e.g., during high-peak sales seasons). As automated picking solutions are generally linked to a specific capacity, human operators compensate for these fluctuations and persist to play a decisive role in aligning supply and demand. They are also able to handle a larger product variety along different criteria such as product dimensions, weight, special handling requirements (for fragile products for instance) or packaging types (Gutelius and Theodore, 2019). Automated and robotized systems are typically fixed for certain product specifications, while humans continue to complement or even outperform those in dynamic circumstances with changing specifications (Sgarbossa et al., 2020). Thus, manual workforces and machines will be working alongside each other in the warehouse of the future (Olsen and Tomlin, 2020), leading to the necessity to optimize collaborations among humans and machines.

In this dissertation, these resulting human-machine interactions and the large variety of order picking solutions are addressed (see Figure 1.1). Specifically, Chapter 3 (Article 1) develops a research agenda for human-machine interactions in warehousing including behavioral issues, theoretical foundations and unifying themes. This is the first necessary step to generate a holistic and accurate understanding of this nascent, yet emerging area. The formulation of research questions as well as the development of theoretical foundations and unifying themes is imperative to guide the way for future research. By doing so, incorporating behavioral issues into future optimization approaches for human-machine interactions is required to account for the human factor and to ultimately establish efficient operating policies. One of those issues, that is mental impoverishment and stagnating system performance, is tackled utilizing an intervention-based research (IBR) approach in Chapter 4 (Article 2). Solving this issue is particularly important as maximizing performance for repetitive and monotonous operational activities plays a major role in many organizations' success (Bernstein, 2012; KC, 2020; Staats and Gino, 2012). Especially in emerging human-machine interactions for warehousing tasks, human workers often perform such repetitive assignments (e.g., physical retrieving of items or erecting and folding cartons (Bai et al., 2021; Sun et al., 2021; Wang et al., 2021)). Hence, it is crucial to maximize both human factors and system performance. Finally, a novel mathematical optimization model is introduced and formalized for the strategic OPS selection and assignment problem (Chapter 5, Article 3). This decision support is imperative for warehouse planners, as no suitable model exists that addresses recent development and challenges such as the skyrocketing number of novel OPSs, the ongoing labor shortage, the enlarging product diversity or the increasing importance of space utilization. Thus, Chapter 5 contributes scholars and practitioners alike, particularly by deriving the conceptual background to establish necessary decision variables and constraints, conducting a case study to prove a large cost saving potential, and applying numerical experiments to generate managerial insights.

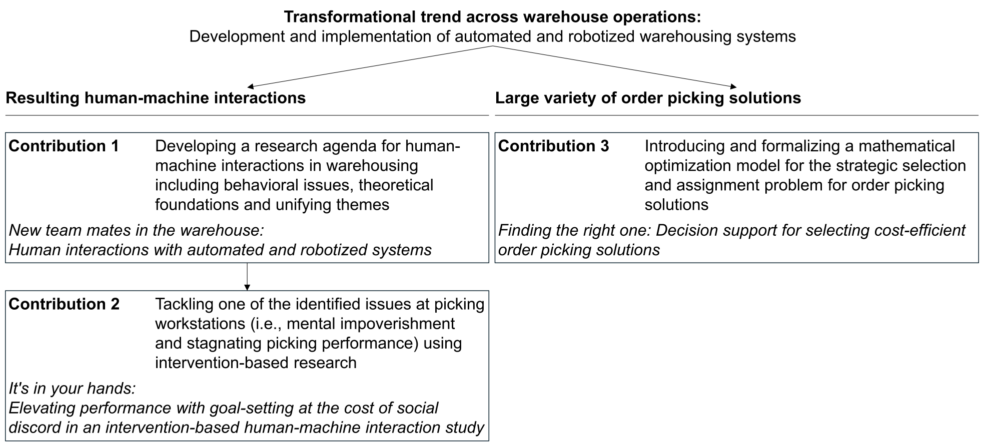


Figure 1.1: Relationship of the three articles

Given the complexity of warehousing and the underlying decision problems, research in this thesis is not limited to one specific research methodology. Instead, a variety of approaches (such as qualitative interviews, literature review, conceptual theory building, mathematical optimization or intervention-based research) are applied to develop a holistic understanding.

2 Contributions

This chapter introduces the three articles (Chapter 3 to Chapter 5) that compose the main body of the doctoral thesis. For each of the articles, it gives an overview on the purpose, methodology and findings. Additionally, Table 2.1 lists the co-authors and states the current status of publication, while Table 2.2 provides the co-author roles along the contributor roles taxonomy provided by Brand et al. (2015).

Article		Co-authors	Status
1	New team mates in the warehouse: Human interactions with automated and robotized systems	Andreas Fügener and Alexander Hübner	Accepted and published online in IISE Transactions (forthcoming)
2	It's in your hands: Elevating performance at the cost of social discord in an intervention-based human-machine interaction study	Andreas Fügener and Alexander Hübner	In the process of submission to Journal of Operations Management as of 09.08.2022
3	Finding the right one: Decision support for selecting cost-efficient order picking solutions	Fabian Schäfer and Alexander Hübner	In the process of submission to IISE Transactions on 09.08.2022

Table 2.1: Status of publication

Article & Author		Contributor roles
1	Lorson, Fabian	Conceptualization, Methodology, Validation, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project administration
	Fügener, Andreas	Conceptualization, Methodology, Validation, Writing - Review & Editing
	Hübner, Alexander	Conceptualization, Methodology, Validation, Writing - Review & Editing, Supervision
2	Lorson, Fabian	Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project administration
	Fügener, Andreas	Conceptualization, Methodology, Validation, Formal analysis, Writing - Review & Editing, Supervision
	Hübner, Alexander	Conceptualization, Methodology, Validation, Writing - Review & Editing, Supervision, Partner acquisition
3	Schäfer, Fabian	Methodology, Software, Validation, Formal analysis, Investigation, Resources, Data Curation, Writing - Original Draft, Writing - Review & Editing
	Lorson, Fabian	Conceptualization, Methodology, Formal analysis, Investigation, Writing - Original Draft, Writing - Review & Editing, Visualization, Project administration
	Hübner, Alexander	Writing - Review & Editing, Visualization, Supervision

Table 2.2: Co-authors roles along the taxonomy of Brand et al. (2015)

Remark The versions of Chapter 3 to Chapter 5 may differ slightly from the versions that were published or submitted to the respective journals. This is due to journal-specific guidelines such as formatting or spelling as well as changes that may be undertaken in the course of the peer review process. Yet, relevance and contributions remain unchanged.

2.1 New team mates in the warehouse: Human interactions with automated and robotized warehousing systems

Purpose Research on human-machine interactions in warehousing, and specifically the role of human behavior in operational activities, is a nascent area with a small, yet growing, body of literature. Hence, the goal of this article is to first establish a systematic framework to analyze and discuss identified behavioral issues in human-machine interactions. To account for the novelty of the topic, a research agenda including theoretical foundations and unifying themes is developed to guide future research.

Methodology To generate a holistic and comprehensive understanding of a novel research field, the triangulation of multiple methods (see Figure 2.1) is imperative. A conceptual foundation is first developed to denote the relationship among important building blocks of human-machine interactions. Expert interviews are conducted to identify the most relevant human-machine interactions and associated behavioral issues. A systematic literature review finally links existing work with the identified issues.

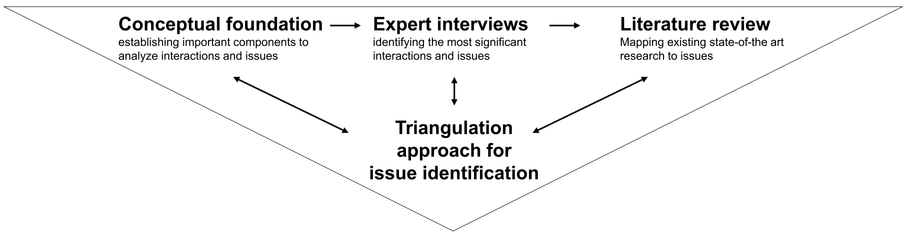


Figure 2.1: Research methodology of Contribution 1

Findings We establish a systematic framework to describe, identify, characterize and derive consequences for human-machine interactions. This framework is used to discuss seven identified behavioral issues with 18 associated research questions across all operational warehousing activities.

In this way, theoretical and managerial insights involving human factors and behavior (e.g., mental workload or satisfaction) are provided for the specific issues. Finally, four unifying themes were derived including theoretical foundations. These themes (such as assigning tasks and developing operating policies among humans and machines, or designing engaging direct interactions) each illustrate a common behavioral aspect across identified issues. The theoretical foundations underpin those themes with prevalent behavioral theories (e.g., goal-setting theory or peer effects) to highlight causalities among the various interconnections (such as human factors and interaction setup component). Figure 2.2 gives an overview.

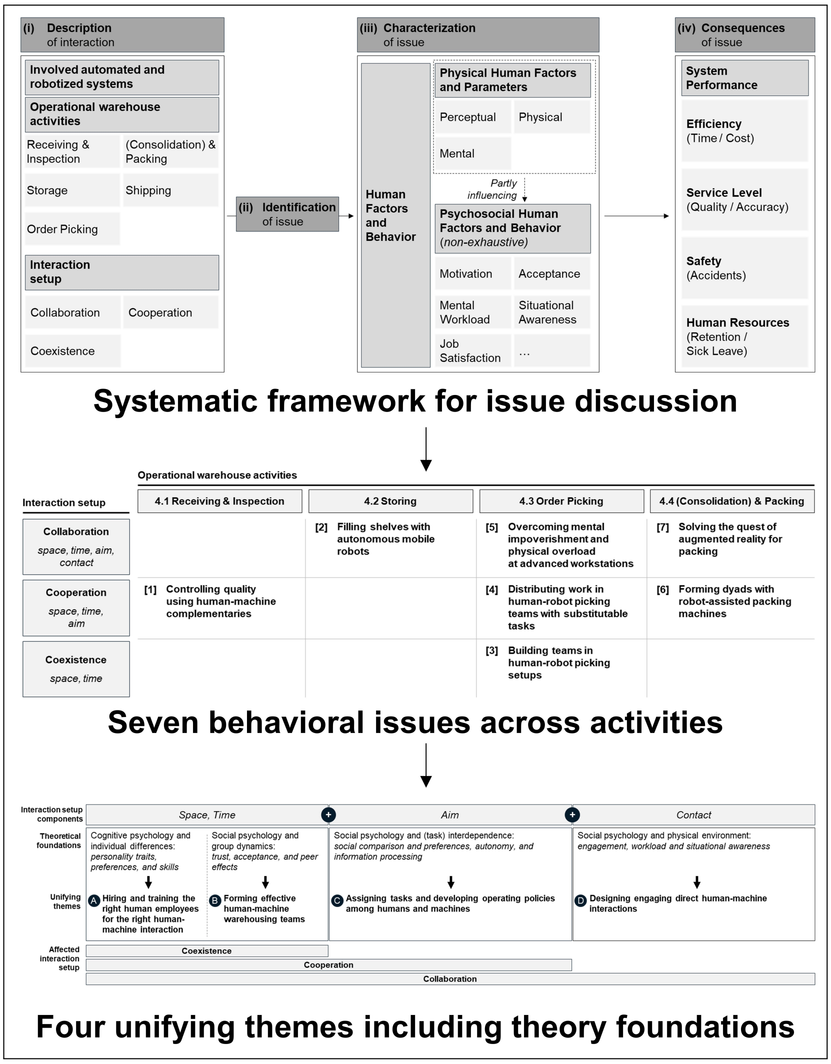


Figure 2.2: Poster summary of findings for Contribution 1

2.2 It's in your hands: Elevating performance with goals at the cost of social discord in an intervention-based human-machine interaction study

Purpose Low satisfaction, self-determination, and perceived fairness (which we call mental impoverishment) paired with stagnating worker performance constitutes a common problem in human-machine interactions for repetitive operational warehouse activities. To tackle this behavioral issue, a goal-setting intervention is introduced which lets the human picker choose out of five different goals (pick amounts) at each workstation. The purpose is to enhance above-mentioned human factors and system performance based on goal-setting mechanisms and theory.

Methodology An IBR approach is utilized to ensure a practice-driven methodology that impacts such a real-world operation policy. By performing a study in the field, a unique opportunity to increase relevance for operations management research is created. To explain findings, IBR approaches often rely on abductive reasoning to iterate between theory and evidence. In this way, plausible explanations about how and why the intervention affected human behavior can be derived.

Findings We find 5.6% performance improvement of worker productivity compared to historical data and a control warehouse (see Figure 2.3). This can be explained by triggered goal-setting mechanisms (such as the higher effort with which workers were engaged in the order picking task) and demonstrates the power of goal-setting theory even in highly physical, operational activities without any kind of monetary incentives. However, scores of worker satisfaction, self-determination, and perceived fairness deteriorated during the intervention. By triangulating surveys, focus interviews, and discussions, we establish the suspension of informal agreements due to the goal-setting intervention as the main reason. Specifically, we find that the goal-setting intervention diminished possibilities for humans to informally organize themselves in their working day, with negative repercussions on the analyzed human factors. Our assessment of both system performance and human factors shows the necessity to account for behavioral aspects when designing human-machine interactions.

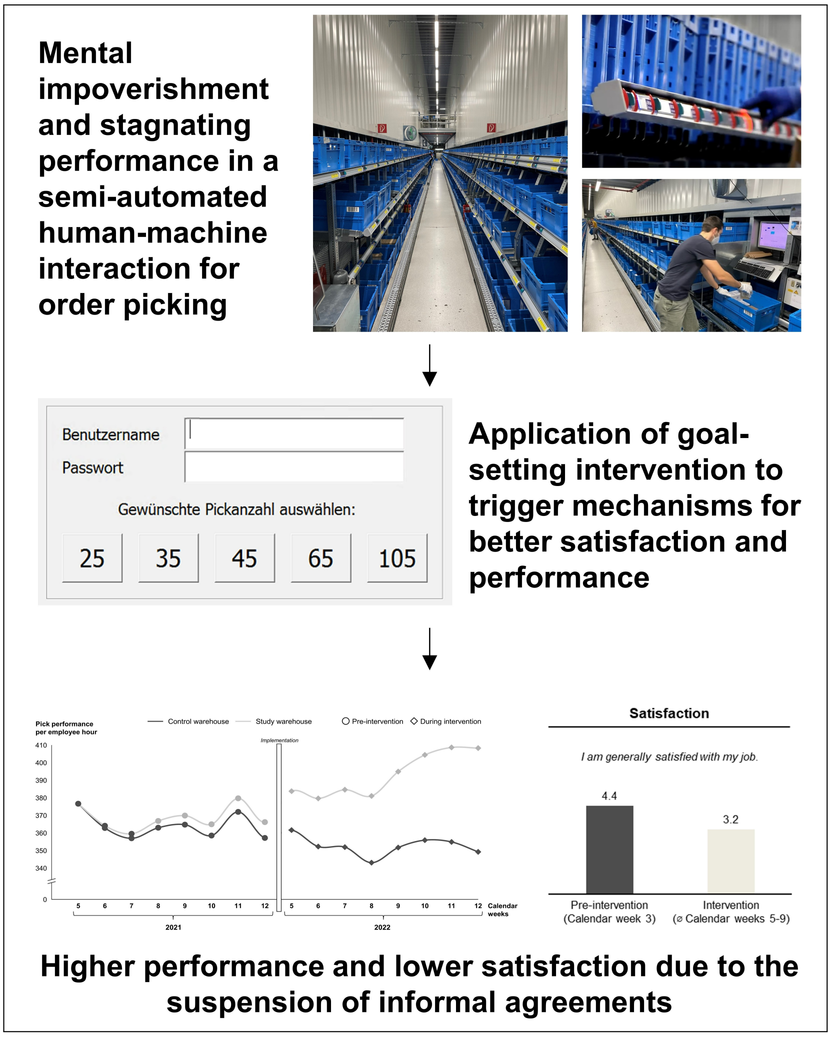


Figure 2.3: Poster summary of findings for Contribution 2

2.3 Finding the right one: Decision support for selecting cost-efficient order picking solutions

Purpose Warehouse managers have to select the most suitable OPSs based out of large variety of potential technologies. To facilitate this decision-making process, a cost-minimizing model that simultaneously selects suitable OPSs and assigns them to available spaces and products is developed. The model is aimed to provide warehouse managers a viable framework for the OPS selection problem, while ensuring all decision-relevant factors and constraints are considered.

Methodology To first account for the novelty of the problem, the conceptual background including decision-relevant costs and constraints is derived by conducting expert interviews and reviewing related literature. The decision problem is then formalized as a mixed-integer program (MIP). By leveraging proprietary data from a business partner, the model selects the most suitable OPSs to minimize total cost while assigning OPS to spaces and products, and adhering to crucial constraints. Numerical experiments are conducted to further produce theoretical and managerial insights.

Findings Decision-relevant cost (e.g., setup, module, labor and error costs) as well as managerial relevant constraints (such as accounting for individual product properties) are first established to conceptualize the innovative OPS selection and assignment problem. The developed mixed-integer cost minimization model solves the problem efficiently, even with varying problem sizes. Utilizing data from a case study retrieved through an industry partner evidences substantial savings potential (up to 57%) when applying the optimal OPS mix of the decision model. In this case, for the underlying set of products and warehouse specifications, the model selects

45 shelve-moving robots (SMRs) and 4 human workers for manual picking as the most suitable OPS. Figure 2.4 provides an overview. Additionally, numerical experiments based on the case study highlight the robustness of the solution and the need to retain human operators until full automation is possible on a large and cost-efficient scale.

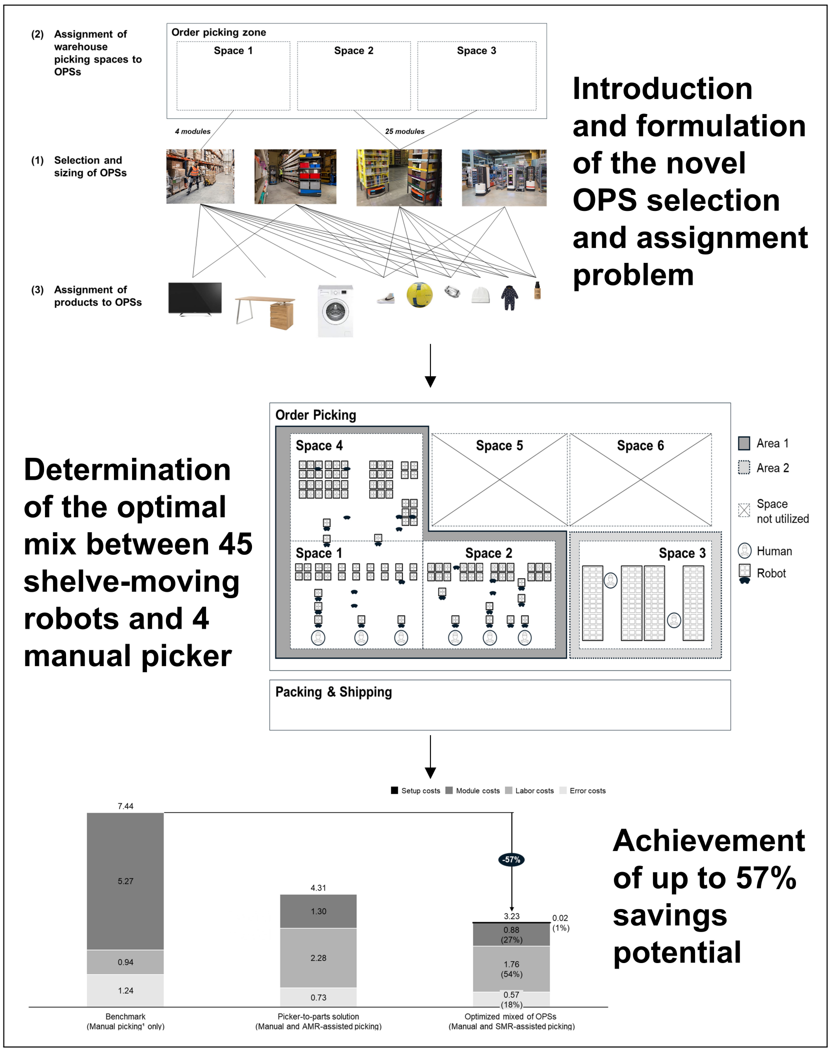


Figure 2.4: Poster summary of findings for Contribution 3

3 New team mates in the warehouse: Human interactions with automated and robotized warehousing systems

Co-authors: Andreas Fügener and Alexander Hübner
Accepted in *IISE Transactions* on April 10, 2022
(forthcoming, DOI: https://doi.org/10.1080/24725854.2022.2072545)

Abstract Despite all the technological progress in the arena of automated and robotized systems, humans will continue to play a significant role in the warehouse of the future due to their distinctive skills and economic advantages for certain tasks. While industry and engineering mainly dealt with the design and functionalities of automated warehouses, the role of human factors and behavior is still underrepresented. Yet, many novel warehousing systems require human-machine interactions, leading to a growing scientific and managerial necessity to consider human factors and behavior, particularly for operational activities. This is the first paper that comprehensively identifies and analyzes relevant behavioral issues of interactions between warehouse operators and machines. To do so, we develop a systematic framework that links human-machine interactions with behavioral issues and implications on system performance across all operational warehouse activities. Insights generated by interviews with warehousing experts are applied to identify the most important issues. We develop a comprehensive research agenda, consisting of a set of potential research questions associated to the identified behavioral issues. The discussion is enriched by providing theoretical and managerial insights from related domains and existing warehousing research. Ultimately, we consolidate our findings by developing overarching theoretical foundations and deriving unifying themes.

3.1 Introduction

Over decades, warehouse operations have traditionally relied on manual processes, due to human operators being more efficient in many aspects such as handling and picking a large variety of products. Enabled by advances in Internet of Things devices and artificial intelligence coupled with the advent of new system providers and more cost-efficient solutions, warehousing has been revolutionized during the last decade: Human operators found themselves next to new robotized and automated teammates (Olsen and Tomlin, 2020). The size of the warehouse automation industry has been growing by 12% annually between 2014 and 2019, and is predicted to double its size from USD 15 billion to USD 30 billion in the next six years (IFR, 2020; Statista, 2020; The Logistics iQ, 2020). The resulting development and utilization of novel automated and robotized systems are boosting the transformation of warehousing from a cost center to a central component in the value proposition of firms. Automated warehousing systems help in this process by enabling faster throughput times, higher service levels, labor cost reductions, efficient space utilization, and improved ergonomics for human workers (see, e.g., Azadeh et al. (2019); Lamballais et al. (2020); Zaerpour et al. (2017)). For instance, Amazon is currently employing more than 200,000 warehouse robots to accelerate its growth in online retail and logistics, driven by faster picking times and lower operating costs (IHCI, 2020). There are many other examples including Hermes, a leading logistics provider, who optimized its return handling processes by installing a new automated inspection and handling system, increasing the capacity by 50% (Logistics Manager, 2020). Innovations in warehouse automation thus play a crucial part in delivering products efficiently and effectively throughout supply chains.

Despite the growing and ubiquitous presence of automated and robotized systems in warehouses, human operators will still be necessary to fulfill operational activities. Tye Brady, the chief technologist of Amazon Robotics, described this with the following words: "*The efficiencies we gain from our* *associates and robotics working together harmoniously – what I like to call a symphony of humans and machines working together – allows us to pass along a lower cost to our customer*" (IEEE, 2020). This statement is just one of many anecdotes evidencing that manual workforces and machines will be working alongside each other in the warehouse of the future (Olsen and Tomlin, 2020). Humans have distinctive characteristics, skills and capabilities that robots are not able to replicate or perform cost efficiently. For instance, they excel in flexibility when swift reactions are needed to volatility of the picking workload (e.g., during high-peak sales seasons). As automated picking solutions are generally linked to a specific capacity, human operators compensate for these fluctuations and persist to play a decisive role in aligning supply and demand. They are also able to handle a larger product variety along different criteria such as product dimensions, weight, special handling requirements (for fragile products for instance) or packaging types (Gutelius and Theodore, 2019). Automated and robotized systems are typically fixed for certain product specifications, while humans continue to complement or even outperform those in dynamic circumstances with changing specifications (Sgarbossa et al., 2020).

Figure 3.1: Simultaneous picking with robots (Source: Magazino)

Figure 3.2: Picking at workstations (Source: Knapp)

As workers collaborate with automated and robotized systems on many tasks across the main operational activities (i.e., receiving, storing, picking and packing - see Fig. 3.1 and Fig. 3.2 for examples), new models, frameworks and concepts are needed to efficiently manage human-machine interactions (De Koster et al., 2020; Olsen and Tomlin, 2020). Such interactions at the operational execution level are part of a socio-technical system with many variables (Monostori et al., 2016; Yang et al., 2019), including human factors and behavior. These systems are usually developed considering the views of engineers or programmers, while the perspectives of the actual blue-collar workers in the loop and corresponding behavioral aspects are often neglected (Moniz and Krings, 2016). However, actions and decisions of the operators may deviate from engineers' expectations and thus impact operations management metrics in both positive and negative directions (Bendoly et al., 2006; Boudreau et al., 2003; Croson et al., 2013; Papadopoulos et al., 2019; Udenio et al., 2017). To establish efficient automated and robotized warehousing systems, it is imperative to understand and account for human factors of workers in operational activities (Donohue et al., 2020), and to consider behavioral methodologies since they provide the opportunity to solve emerging issues in human-machine interactions (Kumar et al., 2018). Combining machine-centric operations management (OM), i.e., the design, plan, control and management of systems and processes, with human-centric behavioral science (BS), i.e., the exploration and integration of human actions, factors and behavior, becomes indispensable to improve decisions and capabilities in automated and robotized warehouses. Fig. 3.3 visualizes the blending of two required perspectives to efficiently manage operational warehouse activities using human-machine interactions.

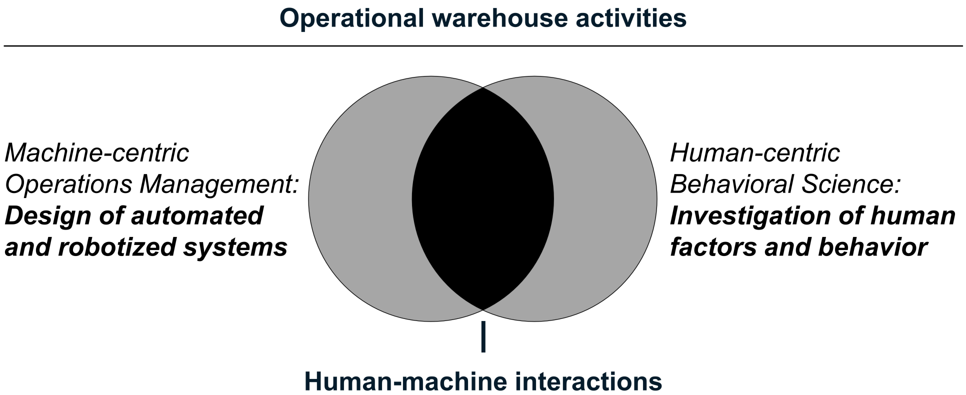


Figure 3.3: Necessary blending of research streams to establish efficient interactions

However, the current literature on such interactions and their behavioral implications on OM in warehousing is rather scarce. Four recent review articles exist from the first literature stream on operational management for automated and robotized warehousing systems (the left of Fig. 3.3):

Azadeh et al. (2019) structure novel systems for storage and picking activities along design and control of technologies, modeling techniques, and research opportunities. Boysen et al. (2019) and Boysen et al. (2021) discuss warehousing systems for their suitability to e-commerce and bricks-and-mortar retailing, respectively. Fragapane et al. (2021) review pertinent work on autonomous mobile intralogistic robots and provide guidance and methods for their planning and control. Additional studies in this stream focus on the development of mathematical operation models and decision support for specific applications in automated storage and order picking systems (see Tappia et al. (2019), Yuan et al. (2019), Lamballais et al. (2020) or Xie et al. (2021) for examples). Azadeh et al. (2019) conclude that further research should be conducted on novel warehousing systems to cope with rapid developments of technologies and increased implementation in practice. Most importantly, none of above-mentioned studies incorporate human factors and behavior into their analysis, nor do these articles focus on the specifics of interactions between operators and depicted systems. Regarding the second stream and the behavioral perspective in warehouses (the right of Fig. 3.3), human factors and behavioral issues for operational activities are discussed by Grosse et al. (2015, 2017) in a content analysis and literature review on human factors in manual order picking. Besides that, only few selected use cases involving behavioral aspects in manual order picking exist (see De Vries et al. (2016a,b), Matusiak et al. (2017), Batt and Gallino (2019) or Glock et al. (2019) for examples). This means that even for conventional warehouses, human factors have not even been adequately addressed up to now. In this sense, it also remains unclear which human-machine interactions and behavioral mechanisms are crucial in automated warehouses, although their analysis and optimization is important to ensure efficient operations.

The controllable and structured environment in warehouses makes many of them incubators for the development and application of automated and robotized systems in supply chains (Azadeh et al., 2019; Fragapane et al., 2021). Specific requirements for operational activities and the necessary collaboration of operators with a growing diversity of machines expose warehousing as a unique research area at the intersection of OM and BS. Despite several calls for research on behavioral implications of operational human-machine interactions in warehousing (see e.g., Azadeh et al. (2019), Boysen et al. (2019), or Jaghbeer et al. (2020)), the amount of existing studies is very limited. Exploring a new research area at an intersection of research domains requires developing a common understanding across literature streams. This needs to be accomplished with a comprehensive identification and ordering of the nascent topics before they can be analyzed in depth and in a structured manner. Naturally, such new areas are insufficiently explored and require the formulation of research questions. Hence, this is the first paper that comprehensively compiles a research agenda for human-machine interactions in the warehouse including theoretical foundations and unifying themes. To ensure a structured approach that connects all relevant dimensions and variables in this domain, we develop a systematic framework. This forms the foundation to identify and analyze the most relevant behavioral issues for these interactions (for the sake of brevity we use the term "issue" in the following), including open research questions. This is amended with theoretical and managerial insights from related domains and existing warehousing research, serving as starting points to improve operational decision-making for human interactions with automated and robotized systems. Ultimately, we consolidate the findings by providing theoretical foundations and unifying themes, guiding the way for future research in human-machine interactions in the warehouse.

Our paper aims at helping OM, and in particular warehousing, researchers to identify potential effects of human behavior. Furthermore, we want to encourage scholars from the field of BS and human factors to consider warehousing as an interesting area of application. When discussing issues in detail, we first identify the issue and associated research questions, before elaborating on the mechanism and consequences on system performance. The latter provides in-depth insights for researchers planning to analyze those or related issues. The overarching theoretical foundations and unifying themes provide causal and salient relationships in warehousing interactions as starting points for further studies.

The remainder begins with Section 3.2, detailing the methodological approach. We build a systematic framework in Section 3.3 to analyze human-machine interactions and behavioral issues in Section 3.4. The findings are summarized by developing theoretical foundations and unifying themes in Section 3.5. Section 3.6 concludes with managerial and theoretical implications, and provides limitations as well as a brief outlook on our study.

3.2 Research methodology

Research on human-machine interactions in warehousing, and specifically the role of human behavior in operational activities, is a nascent area with a small but growing body of literature. As we want to generate a holistic and accurate understanding for this matter, we rely on multi-method approaches which are imperative in such cases (see e.g., Boyer and Swink (2008); Flick et al. (2004)). Further, Lewis-Beck et al. (2004, p. 1142) argue that using methodological triangulation when probing issues "*offers the prospect of enhanced confidence*" in the ensuing findings. Consequently, we triangulate three research methods for issue identification (see Fig. 3.4). We first follow well-established guidelines for emerging topics (Webster and Watson, 2002) and start with the development of a *conceptual foundation*, which is based on central theories in related fields of OM, BS and human-machine interaction. This research step delivers the foundation for the systematic framework in Section 3.3, which denotes the relationships among important building blocks of human-machine interactions in warehousing. Secondly, we conduct *expert interviews* with practitioners to identify the most important human-machine interactions and associated behavioral issues as recommended by Edmondson and Mcmanus (2007). These empirical findings build the main source to derive seven categories (each category represents one issue). Section 3.4 is then structured along the seven issues. A systematic *literature review* is the last, pivotal step to deepen links among managerial issues and existing work (DeHoratius and Rabinovich, 2011). We identify 13 articles that are matched to the identified issues. Only the continuous and comprehensive triangulation of these sources provides the opportunity to structurally identify and analyze relevant issues, create a comprehensive research agenda, and ultimately develop overarching theoretical foundations and unifying themes. For details on the research approach please see the Appendix.

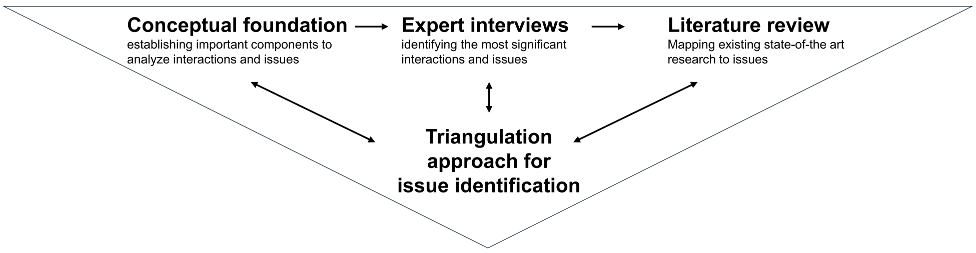


Figure 3.4: Overview of research methodology applied

3.3 Systematic framework to investigate human-machine interactions in warehouses

The systematic framework is designed to structure the investigation of issues by providing a set of important components for human-machine interactions in warehousing and its interconnections. It synthesizes seminal literature and theories from machine-centric OM, human-centric BS, and human-machine interactions outside the warehouse domain. As human-machine interactions in warehousing constitutes a novel area, this exposure to conceptual foundations is essential for comprehensive and structured future research. As such, it constitutes the first contribution of our study and is applied to our analysis at the same time. In this way, we ensure an end-to-end perspective, create a suitable structure for the issue investigation, and uncover open research. We identify important building blocks to discuss behavioral issues which are summarized in Figure 3.5. For the (i) description of interaction, it is necessary to analyze *involved automated and robotized systems*, *operational warehouse activities* and corresponding *interaction setups*. Subsequently to the (ii) identification of issues mainly driven by expert interviews, we follow with the (iii) characterization of those with involved *human factors and behavior*. Finally, the interaction and associated issue have (iv) consequences on *system performance*.

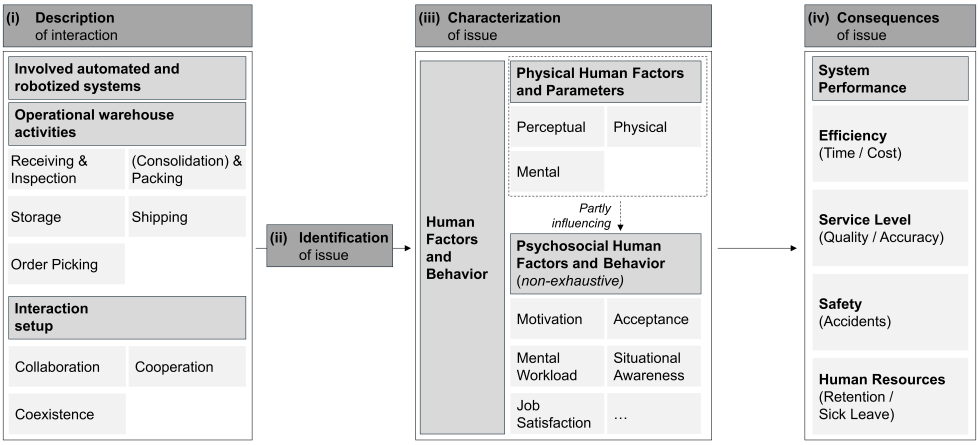


Figure 3.5: Framework to investigate behavioral issues of human-machine interactions

(i) Description of interaction We regard interactions between at least one human worker and one or multiple *involved automated or robotized systems*. Automated systems are defined as machines that carry out a function by themselves that was previously performed by a human (Parasuraman and Riley, 1997). Additionally, physical robots or robotic devices are able to perform tasks with a certain degree of autonomy, and may be able to move within a specific environment (ISO, 2012). Besides these robotized systems, we also regard embedded artificial intelligence in machines as fitting systems for our analysis (Glikson and Woolley, 2020). For an overview we refer to Azadeh et al. (2019), Boysen et al. (2019, 2021) and Fragapane et al. (2021) who cover many potential systems involved. As we deal with human-machine interactions in operational warehouse activities for physical flow, we exclude instances in which humans modify standard workflows without any significant interaction with the machine (e.g., changing picking route).

The main blue-collar, *operational warehouse activities* are receiving and inspection, storing, order picking, packing and shipping (De Koster et al., 2007). Receiving and inspection includes the unloading of products from the delivery vehicle, checking for any quantity or quality inconsistency, and entering master data into the warehousing management system. Additionally, it contains the handling of returns. Subsequently, incoming or returned products are transferred from the unloading to designated put away areas. This process may also include any re-packaging before storing the goods. Once customer orders arrive, the process of order picking consists of retrieving the right products from storage. This may include batching, routing and sequencing. After goods have been retrieved, they are packed (and potentially categorized) for delivery before being shipped to customers. Note that in some cases it is necessary to consolidate orders before packing (e.g., if batch picking is utilized). Packing activities cover boxing or palletizing, packaging (e.g., to protect from transport damage), value-added services (such as labelling, serialization, kitting), or a final quality check. Ultimately, the products are loaded onto the means of transportation to be shipped to the customers or the next step of the supply chain. Similar to Gu et al. (2007b), shipping activities are included in our discussion on receiving and inspection.

The *interaction setup* classifies interactions along proximity and dependency (Schmidtler et al., 2015). In the least intense form, coexistence, the interaction takes place in the same space and time. In cooperation setups, humans and machines also work jointly on the same aim. Collaboration additionally requires physical contact. To translate this for our purposes, we define the space and time as the same warehouse zone and shift. The condition of the same aim is fulfilled if humans and machines work on the same job (e.g., customer order). For physical contact, we consider actions that include either direct physical contact (e.g., wearing a device) or handovers.

(ii) Identification of issue We analyze the described interactions and identify potential behavioral issues based on our expert interviews and theoretical BS foundations. We develop open research questions using the following characterization and consequences of the framework elements.

(iii) Characterization of issue *Human factors and behavior* are the core of the investigation of human-machine interactions. We base our analysis on human factors theory (Sanders and McCormick, 1993; Karwowski, 2005; Salvendy, 2012) and behavioral aspects (e.g., Gino and Pisano (2008)). As a first step, a differentiation needs to be made between *physical human factors and parameters* and *psychosocial human factors and behavior* (Karwowski, 2005). The former are clustered into perceptual, mental and physical. Perceptual parameters include seeing, hearing or perceiving other agents (i.e., humans or machines). Mental or cognitive parameters are processes such as remembering, thinking, judging, decision-making or reasoning. Finally, physical parameters are connected to human movements or activity, such as using body parts, operating, walking or carrying. These physical human factors and parameters are determined by the interaction among humans and machines. Additionally, they may impact and change psychosocial factors and behavior (Karwowski, 2005). These include but are not limited to motivation, acceptance, workload, stress, situational awareness, job satisfaction, trust, reaction to incentives or fairness (see Boudreau et al. (2003), Loch and Wu (2005), Gino and Pisano (2008), Glikson and Woolley (2020), or Parasuraman et al. (2008) for examples). The degree and magnitude of these behavioral aspects may also be decisive for resulting actions and decisions (including biases and heuristics, see Udenio et al. (2017) for instance), and depend on individual human characteristics (such as personality types).

(iv) Consequences of issue The final outcome of the interaction is assessed using *system performance* criteria. For direct measures, a differentiation is made among efficiency and service-level (Staudt et al., 2015). Efficiency measures include the time (such as processing or lead time) or cost of a certain warehouse activity (e.g., the number of order lines picked per time and cost unit). Service level criteria include the quality (e.g., shipped orders without damages) or accuracy (e.g., share of orders delivery without errors). Further, safety metrics (such as number of occupational accidents, see De Koster et al. (2011) and De Vries et al. (2016c)) are another performance criterion as safety issues among blue-collar workers are common, especially in the logistics sector. Finally, we consider criteria for human resource management such as retention or number of sick days. These performance criteria are regarded to acknowledge and account for the impact of human well-being.

The different parts of the systematic framework are utilized in the following to structure the issue analysis in the next section.

3.4 Behavioral issues in human-machine interactions

This section discusses seven issues that have been identified. Fig. 3.6 orders the issues based on the elements *operational warehousing activity* and *interaction setup* of step (i) of our systematic framework. We structure the discussion of concrete behavioral issues along those two dimensions, where we use a subsection for each *operational activity*. Each subsection starts with the (i) description of the interaction with *involved automated and robotized systems*. Along the *interaction setup*, we continue with the (ii) identification of associated issues, including potential open research questions for each issue. Following with the (iii) characterization and (iv) consequences of those issues, we provide theoretical and managerial insights involving *human factors and behavior* and impact on *system performance* based on warehousing and related literature. A number of issues may also be applicable to other activities and the generalization and transfer to those is discussed in Section 3.5. We formulate the future research questions in a generalizable manner to represent challenges and opportunities across multiple activities.

Interaction setup	Operational warehouse activities: 4.1 Receiving & Inspection	4.2 Storing	4.3 Order Picking	4.4 (Consolidation) & Packing
Collaboration *space, time, aim, contact*		[2] Filling shelves with autonomous mobile robots	[5] Overcoming mental impoverishment and physical overload at advanced workstations	[7] Solving the quest of augmented reality for packing
Cooperation *space, time, aim*	[1] Controlling quality using human-machine complementaries		[4] Distributing work in human-robot picking teams with substitutable tasks	[6] Forming dyads with robot-assisted packing machines
Coexistence *space, time*			[3] Building teams in human-robot picking setups	

Figure 3.6: Overview of behavioral issues identified in human-machine interactions

3.4.1 Receiving and inspection: interactions and issues

The inspection process has important human interactions with automated control machines, whereas receiving processes are either manual or fully automated with no relevant human-machine interactions. As such, we focus here on inspections and quality control.

(i) Description of interactions Visual inspection systems that are responsible for checking the quantity or quality as well as measuring product-related data have made great advances in the last years, and many automated systems have been installed in warehouses. However, as these systems are often not able to fully cover a large range of products, human operators complement the process with their input on non-feasible or unclear cases, often by receiving an error message that manual help is required. The products in question are either separated and transported on a conveyor belt to a workstation, or the human worker needs to troubleshoot right at the machine. As no direct physical contact with the machine is required, but the human operator and the machine work on the same incoming products in these instances, the interaction is classified as a cooperation setup.

(ii) Issue identification - [1] **Controlling quality using human-machine complementarities** Warehouse operators and system providers made it clear in their interviews that leveraging the complementary strengths when checking the quality (e.g., identifying defects), quantity or dimensions of incoming products attains higher performance, and is crucial to establishing efficient operating policies. As an example, one warehousing manager [12WO] highlighted the need to combine both human and machine skills: "*We are able to process the basic products with our automated inspection machine. However, we still rely on one additional employee when it comes to SKUs that are hard to distinguish for the machine, for example, if it is a small defect or natural variation, fragile or inconsistent, or unknown such as promotional products.*" Humans thus complement the machine's ability, function as the final decision maker by judging whether the products meet pre-defined criteria, and act as a supervisor or troubleshooter. Such hybrid settings are needed in many cases, as neither unaided humans nor full automation is as effective as combined work. As humans enhance the process through advantages in flexibility and skill ranges, the following research questions evoke:

[1.1] What is the impact of setup choices (such as communication, control criteria and process order) on performance, and what are the underlying mechanisms and psychological factors?
[1.2] What is the degree of perceived transparency and feedback influencing inspection performance, what behavioral aspects may explain individual reactions to those factors, and why?

(iii) Characterization and (iv) Consequences of issue Building on knowledge from cognate applications in OM, decision-making biases (such as anchoring) may be present when troubleshooting is required. Setup choices that influence these include the communication (e.g., should the machine give a recommendation when delegating), distribution of the products (e.g., which products should be delegated to the human based on experience) or the design of the control criteria (e.g., unrealistic or complex accept or reject criteria).

The behavioral influence of these choices need to be analyzed and may depend on individual characteristics of the employees. Some may be prone to anchoring biases, some may be robust to potential false indications of the machines, and some tend to lower their effort by exhibiting a high degree of automation complacency (Parasuraman and Manzey, 2010). It is also important to understand the order of the process to decrease throughput times (e.g., when should the human pro-actively prepare incoming goods for the inspection machine due to bulky items or broken pallets). As in similar setups inspection systems show better performance when combining human and machine skills (see See et al. (2017) for instance), it is necessary to find the optimal incorporation of the above-mentioned setup choices to stimulate human action and minimize inspection errors in warehousing settings, too. Moreover, human motivation, mood and satisfaction can be impacted in this process (Bainbridge, 2002; Lughofer et al., 2009). For example, system performance increases when operators know that their input will be included in the algorithm of the system ex-ante (Lughofer et al., 2009), or when they feel a machine is making intelligible decisions (Kellogg et al., 2020). Motivation may be particularly impacted by the technical architecture (Bendoly et al., 2010). As such, the ability to provide feedback to the machine and to the operator is crucial to increase the interest and willingness for smooth joint work, and emphasizes the machine's ability to learn from the human as an expert (Kadir and Broberg, 2020). Also, it is important to create cognitive and emotional trust to enhance the success of automation integration, particularly by achieving high system transparency and reliability (Glikson and Woolley, 2020). Clearly, analyzing setup choices, process order and motivational, trust and feedback aspects are promising starting points for future research. Note that in some cases this inspection activity may also be classified as collaboration depending on the specific machine that is utilized.

3.4.2 Storing: interactions and issues

Instead of storing goods either manually with a forklift or with automated storage and retrieval systems (AS/RS), semi-automated solutions are utilized that result in significant interactions between humans workers and automated or robotized storing systems.

(i) Description of interactions In such a hybrid setup, humans are supported in filling storage shelves by automated guided vehicles (AGVs) or autonomous mobile robots (AMRs). As main aspects of the following discussion are connected to AGVs and AMRs, the findings may also be applicable for order picking as the reverse application to storing. In general, these machines transport the products when traveling through the aisles next to shelves (Fragapane et al., 2021). In some cases, certain types offer a seating possibility, or even assume the lifting aspect of the storing activity (via a robotic arm or lift). Both the flexibility of the human operator to store a variety of products on shelves and the technological advances of the machines can be leveraged in such systems. This is particularly suitable when individual items need to be handled instead of full pallets. Examples of applications include spare parts warehouses, micro fulfillment centers, and supermarket shelves. Such hybrid approaches require the hand-over of products from machine to human, and we therefore focus in the following on a collaboration setup.

(ii) Issue identification - [2] **Filling shelves with autonomous mobile robots** As the replenishment process frequently accounts for a large share of working time of employees (Boywitz et al., 2019), the efficient orchestration of human and machine leads to several open research questions:

[2.1] What are optimal design choices of collaborative robots for replenishing products, and how and why do psychosocial factors, incentive schemes or personality traits influence such setups?

[2.2] Which incentive schemes or personality traits are beneficial for human employees in a fixed or a floating AMR operational policy, and what mechanisms may explain individual differences?

(iii) Characterization and (iv) Consequences of issue The operator moves with the robot to the storing locations. The physical put away process is conducted by the human, but strenuous bending down is prevented as the machine elevates the product towards a comfortable position for the human. Finally, the operator needs to decide where to place the product, while the robot supports the human with physically demanding tasks (walking, carrying and lifting the products). Such machines (and in particular their arms and lifts) are able to work in different speed settings, often deviating from natural levels of physical human movement. The robot configuration may thus be limited by human abilities that may differ between individual operators.

As a starting point, Roy and Edan (2018) found out that the working pace or default speed of such robots should be the average working speed of the operator to reduce fatigue and stress. The authors base their judgment on human-human experiments and directly derive the implication from their findings. While this may hold true for human-human handover tasks, a further analysis needs to be conducted on the human-machine specifics, and most importantly, on their impact on system performance criteria. Further, when evaluating the behavioral benefits of fixed (AMRs are assigned to a specific worker) or free-floating (AMRs serve multiple pickers) policy (Boysen et al., 2019), it is important to understand which prove to be more efficient depending on individual personality traits (see De Vries et al. (2016a) for a related manual warehouse example). Sauppé and Mutlu (2015) show that employees like to treat robots as a social entity, eventually boosting the perception of their coworker. A fixed strategy thus may increase individual human acceptance as the machine is assigned to the specific employee and satisfies the desire of monopolizing the support (Gombolay et al., 2015). In this light, incentive schemes, that also incorporate potential robot throughput, are interesting research inquiries as they may have a large influence on overall performance. Pasparakis et al. (2021) study another policy option of AMRs: Should the human lead or follow the robot to the picking (or storing) spots. The authors find that for larger efficiency, human leading is superior, while greater accuracy is achieved when humans follow the robots. Further, prevention regulatory focus (as a personality traits) moderates the effect of the different policies on pick speed. However, there is no clear theoretical foundation why these concepts should be interrelated, and hence, it would be interesting to understand which other behavioral mechanisms may play a significant role in this setup. Moreover, the underlying slotting strategy also impacts the interaction. For example, humans may improve performance in dedicated approaches due to learning effects (Weidinger and Boysen, 2018a). Additionally, the benefits of exploiting favorable storing (and consequently picking) locations when deciding on the slotting strategy provides further research potential given ergonomic benefits (Petersen et al., 2005).

3.4.3 Order picking: interactions and issues

Recent warehouse automation efforts have been heavily concentrated on robotized order picking (e.g., Lamballais et al. (2020)), and a large number of companies are offering a variety of systems for this purpose. In particular, focus has been put on minimizing traveling time, as this is the most time consuming task in the picking activity (Tompkins et al., 2010).

(i) Description of interactions We differentiate between picker-to-parts and parts-to-picker solutions, as well as along the degree of automation (Boysen et al., 2019). In picker-to-parts setups, the picker (or in the case of full automation the robot) moves to the storage area to retrieve the products, while in parts-to-picker designs the products are carried to the picker by a transportation system. The degree of automation is the ability and intelligence of the machine to fulfill a single picking task autonomously. Fig.

3.7 provides an overview. Note that regardless of the degree of automation, significant interactions exist.

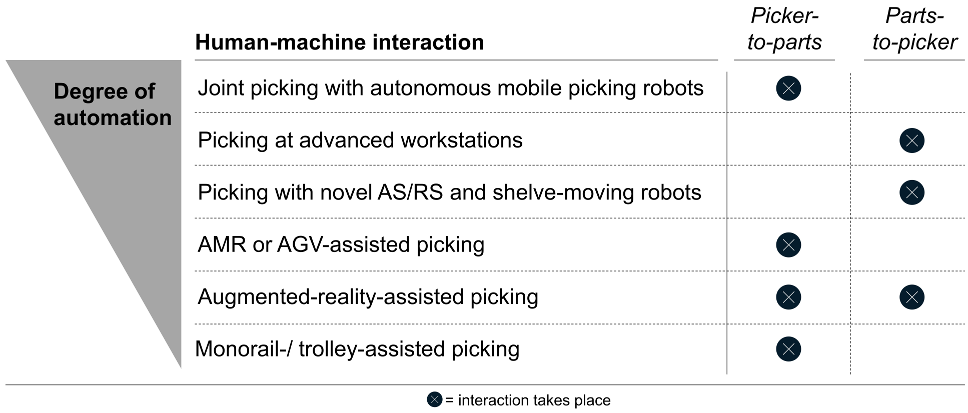


Figure 3.7: Relevant human-machine interactions at order picking

In fully automated picker-to-parts setups, robots are able to fulfill the picking process autonomously without the help of humans (Fottner et al., 2021). Interactions occur when employees are deployed in the same area (see Fig. 3.1), where both an autonomous mobile picking robot and an operator pick in the same aisle. In semi-automated picker-to-parts setups, AMRs, AGVs or trolleys hanging from a monorail help to reduce travel time by allowing pickers to put items on machines that travel to the base. Additionally, virtual and augmented reality applications can support pickers. These systems indicate instructions using perception (via head-mounted displays for instance). In automated parts-to-picker systems with human interaction, pickers are located at advanced workstations and interact using various interfaces (such as buttons or touch screens). Fig. 3.2 shows an example in which a human operator receives input from a display to pick items from arriving totes. Parts are supplied via AS/RS, pouch sorter, shuttle-based technology or shelve-moving mobile robots (Azadeh et al., 2019; Boysen et al., 2021; Yuan et al., 2019). In the latter case, the robots bring back the shelves to a repository or to the next picker after the successful picking process (Weidinger and Boysen, 2018b). Typically, humans fulfill the picking task supported by visualization methods such as pick-to-light. All types of interaction setup are found given the variety of systems and design options available for order picking. We start with an analysis of a coexistence issue, follow with cooperation, and end with the most proximate and dependent setup in a collaboration setup.

(ii) Issue identification - [3] **Building teams in human-robot picking setups** Beginning at human interactions with autonomous picking robots, one key challenge revealed in our expert interviews is determining the team structure, that is, how many humans and how many autonomous picking robots to employ for a given picking zone during the same shift: "*I will have to form new teams, and this will change the human dynamics significantly depending on how many robots I will include* [3WO]." This results in interesting research questions:

[3.1] How does the share of robots impact efficiency and retention of human operators, which behavioral mechanisms govern the differences, and what is the optimal composition and policy in which constellations?
[3.2] Which behavioral traits and skills impact performance when teaming with autonomous robots, what behavioral aspects may explain differences, and why?

(iii) Characterization and (iv) Consequences of issue In such mixed teams, humans see the robots, hear their noises, and maybe even smell their robotic odour (see Fig. 3.1). Humans think about robots as team mates, their role within the team, and how to deal with them. Movements need to be orchestrated to both ensure human safety and robot productivity. Many experts reported different ways employees have of coping with such human-machine coexistence, with one manager [13WO] pointing to unknown consequences: "*We do not know yet what the short- and long-term influence on human social components will be when we employ more and more robots.*"

Insights from BS regarding team composition in general and for human-machine interactions in particular serve as a starting point to analyze this issue. One key aspect of managing teams is to deal with interpersonal processes such as conflict and affect management or collective motivation building to avoid performance problems (Marks et al., 2001). Employees care about human relationships and identify with colleagues (Urda and Loch, 2013), and these social interactions have a large impact on motivation and performance (Cantor and Jin, 2019). In line with that, Stein and Scholz (2019) encourage automation-oriented diversity management when building groups and Gombolay et al. (2015) establish that people value humans more than robots as team members. Hence, psychosocial factors such as motivation, satisfaction or loyalty of employees may vary depending on the human-robot team structure in warehouse operations, too. Additionally, findings about peer effects (Mas and Moretti, 2009; Schultz et al., 2010; Tan and Netessine, 2019) may also exist for such human-machine teaming and impact optimal operating policies. The physical presence of autonomous robots further influences trust and actions, depending on the individual human being (Glikson and Woolley, 2020). Consequently, a thorough understanding of which personalities (see Kaplan et al. (2019) for an extroversion example), behavioral traits or skills prove to enhance performance criteria are promising research directions.

(ii) Issue identification - [4] **Distributing work in human-robot picking teams with substitutable tasks** System providers and warehouse operators further addressed that allocating or distributing work among humans and robots is an essential topic. For example, one system provider [15SP] raised the question of *"which jobs should I give to robots, and which to my [human] employees?"* Compared to issue [1] in which humans and machines complement each other for quality control, this topic now deals mainly with the potential substitution of human and machine work forces, also leading to novel research opportunities:

[4.1] Do performance differences between robots and humans have an impact on the performance of humans, and, if so, which psychosocial factors influence the deviations?

[4.2] What might be optimal operating policies for distributing tasks among robots and humans when accounting for human preferences and behavior?

[4.3] Who (human or robot) should distribute tasks, and how does this impact psychosocial factors?
[4.4] Under which conditions should human employees work with lower or higher perceived autonomy (in task execution and allocation) when teaming with robots?

(iii) Characterization and (iv) Consequences of issue Pickers and robots share the same zone, shift and customer order in such a cooperation. Humans see and hear the robots performing tasks, and consequently question the nature and allocation of the respective jobs (e.g., which items are picked by robots, and which by humans).

A starting point of a behavioral analysis could be to determine both the preference and performance of humans for each specific type of job to decide on the allocation. For the former, this includes an investigation of tasks by product type or location (Larco et al., 2017) in terms of human desirability and comfort, also to avoid devaluation feelings (Gombolay et al., 2015). In this sense, warehouse managers reported the common phenomenon that employee motivation and performance increased when robots were introduced, but the effect diminished over time. On the other hand, if humans see robots performing the undesirable jobs, psychosocial factors (such as satisfaction or acceptance) are improved. In this case the trade-off with physical, ergonomic job-rotation benefits (see Otto and Battaïa (2017) for an assembly example) needs to be evaluated. For the latter, each human has individual skills (see Matusiak et al. (2017)) that affect performance in different job types. This requires an understanding of the jobs in which human performance is generally lower, and should therefore be transferred to robots (for instance to promote specialization, see Schultz et al. (2003) in a production setup). Further, Sanders et al. (2019) find that humans tend to distribute a picking task to humans rather than to robots, mainly due to trust issues and the fear of financial loss for the human. Also, as humans value it when their preferences are taken into account (Gombolay et al., 2017), it is crucial to analyze the influence on performance criteria depending on whether humans or robots distribute tasks in the warehouse. When humans decide on the allocation, the level of trust in humans and machines has a major impact on the decision-making (Sanders et al., 2019). In the reverse setup (i.e., task assignment from a human or an algorithm), Bai et al. (2021) studied the influence on fairness and efficiency. Their results indicate higher perceived fairness when machines are distributing work, even yielding in a persistent boost of picking performance. It remains open how these findings may be different depending on the level of transparency in the distribution process, a significant research opportunity also mentioned by one expert [10WO]. Moreover, Cragg and Loske (2019) compare different picking technologies and find that the lower the human's experienced work autonomy, the higher the picking efficiency. However, the effects of work autonomy on key performance criteria may have a different degree or even magnitude depending on additional human factors (e.g., mental workload) and individual characteristics (e.g., personality traits) of the subjects. Hence, it is relevant to incorporate such factors in further studies as well.

(ii) Issue identification - [5] **Overcoming mental impoverishment and physical overload at advanced workstations** For parts-to-picker setups, a key issue divulging from practice is how to balance mental and physical workload at advanced picking workstations. Warehouse system providers focused in the past on reducing mental workload and achieved progress in improving ergonomics and safety: "*We were able to reduce the physical strain and also designed the systems in a manner that limits the necessary input of employees via several ergonomic initiatives*," as a system provider [16SP] reported. However, the reduction of mental workload for humans also led to several psychosocial problems in the mid- and long-term. For example, one warehouse operator [14WO] reported: *"Unfortunately we see mental impoverishment of our people at the workstations."* This requires addressing the following research questions:

[5.1] What is the optimal amount of perceived decision-freedom and machine support-level for human operators to avoid mental impoverishment?

[5.2] What is the optimal throughput model to maximize both operational efficiency and psychosocial well-being factors?
[5.3] What are efficient incentive schemes to maximize worker and machine productivity?
[5.4] Does backlog design have an impact on psychosocial factors and performance, and if so, what is the optimal design and why are the underlying mechanisms impacting such setups?

(iii) Characterization and (iv) Consequences of issue In collaborative interactions at advanced workstations, humans receive visual input from screens, light or voice indications, including the number of items to be picked within a certain time frame. Based on the perceived information, humans perform their picking task, often by putting items from one bin to another, and confirming the operations executed either via buttons or voice commands. Fig. 3.2 shows an example. Usually, the standard processes are predefined and no mental effort is required. Physical effort (such as the speed of movement) is high as companies usually want to maximize machine output. Behavioral analyses show that performance criteria suffer from mental impoverishment, including lower accuracy (despite visual support of the workstation such as pick-to-light, see D'Addona et al. (2018) in a manufacturing example) and retention as jobs are increasingly unattractive: "*No one wants to do this job anymore* [14W0]." To counteract reduced attention (situational awareness) or job satisfaction as well as increased boredom or fatigue, managers need to innovate the human-machine interaction and account for mental stimulus (for example by providing more decision autonomy or information, including gamification). Moreover, practitioners (e.g., [1SP] [5C], [7WO]) frequently mentioned mounting performance pressures at workstations. Reported consequences are higher stress and physical overload paired with lower job satisfaction. System performance criteria such as a lower service level (see Kostami and Rajagopalan (2014) for a service operations setting) and increased fluctuation are experienced by interview participants (e.g., [2WO], [15SP], [19WO]). In this sense, Batt and Gallino (2019) find insights on how pick times are reduced when humans are more experienced, proving a great need for higher retention. Another related example is provided by Tan and

Netessine (2014), who discover an inverted-U-shaped relationship between workload and performance of service operators. As considering new ways of balancing the physical and mental workload has proven successful in warehousing (see Kudelska and Niedbal (2020), who find decreased mental and physical workload and improved efficiency with shelve-moving robots) and in other settings (see Delasay et al. (2019), Gombolay et al. (2017), Parasuraman et al. (2008), Proctor and van Zandt (2018) or Teigen (1994) for examples), similar analyses on advanced workstations seem promising. Further, different setups of backlog (or perceived workload) at advanced workstations influence efficiency and even motivation or satisfaction: "*We see differences in our shift performance depending on the backlog of open orders on the display at the workstations,*" stated a warehouse manager [10WO]. In a related setting, Wang and Zhou (2018) show that workers operate faster in dedicated compared to shared backlogs in a supermarket context, mainly due to the social loafing effect. Delasay et al. (2019) define the relationship between backload and skill level as an open research avenue. Performance criteria are additionally influenced by other design elements, such as displaying backlog privately or publicly. Also at advanced workstations, it is critical to model and understand actual human behavior when analyzing such parameters (see Wang et al. (2021) in their conclusion on human interactions with shelve-moving robots).

3.4.4 (Consolidation) & packing: interactions and issues

As order consolidation is not a necessary step in all warehouses, we focus on the packing process in the following. Note that human-machine interactions also exist for consolidation (e.g., with put walls or sorting systems, see Boysen et al. (2022)). While fully automated packing lines exist, many warehouses run on semi-automated solution with significant human-machine interactions.

(i) Description of interactions In this paper, we differentiate between two main interactions in packing. First, robots and human workers jointly work at a packaging line by distributing tasks for each sub-activity, which results in a cooperative setup. Second, humans are supported by wearing virtual and augmented reality (VR/AR) systems in collaborative interactions. We focus on head-mounted-devices (or AR glasses) as a common application. For instance, AR systems are able to support the operator in the multiple-bin-size bin-packing problem to load parcels onto a pallet or into a truck. Another use case is the selection of the most efficient container (often the one that minimizes material use), but still perfectly packs and protects all the items to be shipped. Note that AR glasses are also used for picking (pick-by-vision) and findings may be transferable (see Egger and Masood (2020) for an overview).

(ii) Issue identification - [6] **Forming dyads with robot-assisted packing machines** When deciding on semi-automated packing lines, warehouse managers are faced with the decision on which sub-activity to assign to robots and how to design the interaction among humans and machines. In such hybrid work cells, robots take the role of the helping hand for the humans when packing a container for delivery, leading to open research questions:

[6.1] Which sub-activity should be performed by robots and which by humans based on individual personality types and skills, and what is the allocation mechanism?
[6.2] Do perceptional factors of robots impact mental workload and system performance, and if so, why?

(iii) Characterization and (iv) Consequences of issue In such interactions, humans and robots operate in the same space to finalize orders. Workers see, hear, and exchange information with the packing machines. Their new coworker focuses on routine tasks without showing any fatigue (such as erecting the cartons or sorting the products), while humans excel by performing activities that do not always follow structured patterns (such as troubleshooting (Banerjee et al., 2015) or special labeling). When switching from a robot sub-task to a human sub-task (or vice versa), interactions need to be orchestrated and adjusted to fit both the technical skills of the robot and the natural physical movements of the human. Humans need to anticipate and understand why a robot is reacting and behaving in a particular way. This is crucial to mentally anticipate the next move of the robot to ensure safe standard operating procedures.

Banerjee et al. (2015) conduct human-robot kitting experiments and achieve faster execution times and comparable quality by implementing visual indication when human troubleshooting is required. By letting the robot assume repetitive tasks, physical workload is reduced and task duration times lowered. Maettig and Kretschmer (2019) and Maettig et al. (2019) also study the influence of visual indications in a packaging line by minimizing the perceived information to reduce mental workload and improve quality. As the reduction of mental workload or difficulty may evoke different consequences for different people (Schulz et al., 2018), it still remains open which sub-activity of the packing process should be performed by a robot, depending on personality traits, human knowledge or skill. Further, if we assume that the order of tasks within packing is fixed (due to the line setup), and robots and humans jointly solve a task (e.g., robot erects the carton, humans inserts items, robot seals it), humans prefer to work with robots that are pro-active (they know and prepare which task to do next) and information- or intent-sharing (Baraglia et al., 2016). Humans also favor leading the interaction, except when mental workload is high (Schulz et al., 2018). The overall setup consequently impact psychosocial factors such as job satisfaction as well as both physical and mental workload, leading to interesting research possibilities on the influences on performance criteria (e.g., throughput times of the packing line). Besides that, cooperating (or in some instances also collaborating) in such a close proximity with robots may influence perceptual factors (such as noise levels) in the warehouse, and this consequently needs to be addressed as well. Regarding the above-mentioned multiple-bin-size bin-packing problem, Sun et al. (2021) observe that humans deviate from algorithmic suggestions due to superior information or complexity issues. They install a human-centric intervention that incorporates such anticipated deviations, leading to a reduction of deviations and an improved performance. Future research may explore the observed worker heterogeneity (e.g., in terms of traits or preferences) or the possibility to provide additional information. One possible way for this are AR devices, which we discuss next.

(ii) Issue identification - [7] Solving the quest of augmented reality for packing Mark Zuckerberg, the CEO of Facebook, expects AR glasses to redefine the relationship with technology (CNBC, 2020), and many collaborative AR applications already exist for packing and other activities (see Stoltz et al. (2017) for an overview). However, for extensive implementations and safe interactions, the following research questions need to be answered:

[7.1] How can human factors be improved when operating AR devices, what is the performance impact of such behavioral aspects depending on individual workers, and why?

[7.2] What are optimal operating policies (e.g., which tasks to conduct) for operational activities when incorporating preferences and psychosocial effects of employees wearing AR devices?

(iii) Characterization and (iv) Consequences of issue Using AR, the operator sees through the head-mounted device and receives the respective information on the display. These may be the location where to put an item, or which container to choose based on a pre-selection. The human is required to process the information and to act on given instructions (such as putting items into a bin or erecting the carton). In some cases voice commands, gesture or touch screen input are required, depending on the type of AR support.

Three related experiments offer starting points to find answers to above-mentioned questions. Stoltz et al. (2017) analyze human factors and behavior in a parcel-categorizing task using a head-mounted device, while

Kretschmer et al. (2018) and their follow-up study in Plewan et al. (2021) investigate the performance and usability of an AR head-mounted device for palletization. The authors compare the systems to traditional approaches (such as paper based or tablet methods) in all three setups. Stoltz et al. (2017) encounter a potential ephemeral motivational effect given the novelty of the AR glass and the useful information displayed, linking to reduced mental workload as the decision-making processes are assumed by the machine. In line with this, Kretschmer et al. (2018) find a lower mental and temporal demand and experienced effort, but no significant reduction in perceived workload, which is confirmed by Plewan et al. (2021). Note that the results indicate that workload was lowest for the AR condition despite the missing effect significance. Regarding usability, Kretschmer et al. (2018) and Plewan et al. (2021) report lower scores compared to traditional approaches, resulting in a key challenge for practitioners. Interestingly, performance metrics vary across the studies. While Stoltz et al. (2017) and Plewan et al. (2021) find improvement in quality, time (as the efficiency indicator) is not reduced in Kretschmer et al. (2018), and is even negatively impacted in Stoltz et al. (2017) and Plewan et al. (2021). Concluding, it is evident that AR devices help to increase the quality, but efficiency criteria need to be assessed further. In particular, relationships with perceived mental and physical workload, usability and acceptance seem visible, and are also of highest relevance (Masood and Egger, 2019). Wearing a head-mounted device for a whole shift increases physical workload, and users are visually limited and may be distracted due to visual and audio information. Thus, situational awareness and consequently safety is negatively impacted (see Aromaa et al. (2020) in a related lab experiment). Also, understanding the long-term motivational effect given lower decision discretion and competence requirements needs to be understood, and findings are always dependable on the individual hardware, subjects, and their personalities (see De Vries et al. (2016b)). In any case, there are many possibilities to further conduct field experiments with real warehousing workers to answer above-mentioned questions and assess movement towards the expectations Zuckerberg voiced.

3.5 Theoretical foundations and unifying themes

This section provides an aggregated view on the empirical findings through a behavioral lens. In our systematic framework, we characterized each human-machine interaction by its setup components *space*, *time*, *aim*, and *contact*. We now develop theoretical foundations by discussing which behavioral theory informs potential effects based on each interaction setup component. By combining the theoretical foundation with both our insights obtained in the previous section and further coding of our data sources, we derive a set of four unifying themes (A)-(D) for the warehousing context. Each theme illustrates a common behavioral aspect relevant in human-machine interactions in warehousing across operational activities. The theoretical foundations underpin unifying themes with prevalent behavioral theories to highlight the causalities among the various interconnections (such as interaction setup and human factors). Fig. 3.8 summarizes those connections, and Table 3.1 delineates links to issues, interviews and literature.

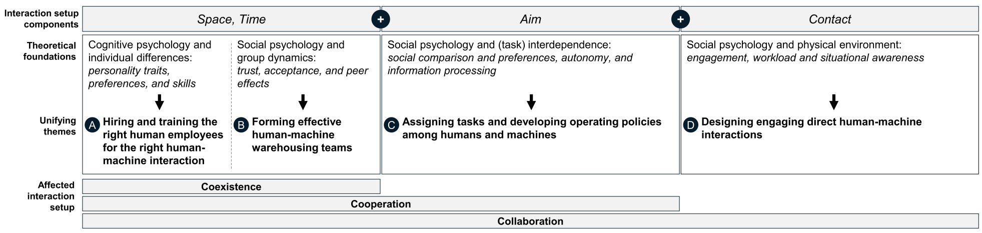


Figure 3.8: Overview of theoretical foundations and unifying themes

Table 3.1: Interconnection among unifying themes, issues, and data sources

Unifying themes		Related issues	Evidence from interviews	Related literature
A	Hiring and training the right human employees for the right interaction	[1], [2], [3], [4], [5], [6], [7]	[2WO], [3WO], [4WO], [5C], [6C], [8WO], [9SP], [11SP], [12WO], [13WO], [14WO], [15SP], [17C], [18SP], [19WO]	Pasparakis et al. (2021), Plewan et al. (2021), Roy and Edan (2018)
B	Forming effective human-machine warehousing teams	[1], [2], [3], [4], [6]	[2WO], [3WO], [4WO], [6C], [8WO], [9SP], [10WO], [13WO], [14WO], [15SP], [16SP], [17C], [18SP], [19WO]	Sanders et al. (2019), Stoltz et al. (2017)
C	Assigning tasks and developing operating policies among humans and machines	[1], [2], [4], [5], [6], [7]	[1SP], [2WO], [3WO], [7WO], [10WO], [11SP], [12WO], [13WO], [14WO], [15SP], [17C], [18SP], [19WO]	Banerjee et al. (2015), Bai et al. (2021), Cragg and Loske (2019), Maettig and Kretschmer (2019), Maettig et al. (2019), Pasparakis et al. (2021), Sun et al. (2021), Roy et al. (2019)
D	Designing engaging direct human-machine interactions	[2], [5], [7]	[2WO], [4WO], [5C], [6C], [7WO], [10WO], [11SP], [13WO], [14WO], [16SP], [17C]	Kretschmer et al. (2018), Kudelska and Niedbal (2020), Plewan et al. (2021), Roy and Edan (2018), Stoltz et al. (2017)

In the following, along the unifying themes, we elaborate on behavioral theories connected to the interaction setup components, and outline which behavioral aspects and mechanisms play a significant role. We further specify our findings from our empirical observations regarding the respective unifying theme. By highlighting which human factors are salient to which consequences, we show causal relationships for research going forward.

(A) Hiring and training the right human employees for the right human-machine interaction As the human-machine interaction is happening in the same *space* and *time*, humans react to the presence of automated machines and robots in warehouses, and hence, theories of cognitive psychology and individual differences of employees inform human behavior within interactions (Croson et al., 2013; Kihlstrom and Park, 2018). Particularly, aspects such as personality traits, preferences and skills vary among humans (Donohue et al., 2020), and thus, play a crucial role in managing efficient warehouse setups. The vast majority of interviewees emphasized the need to hire, train, and employ suitable humans, depending on the operational task at hand (see Table 3.1). The required skills include, but are not limited to, professional (e.g., programming capabilities for warehousing robots), methodological (e.g., trouble shooting skills to resolve workstation blockages), and personal competencies (e.g., eagerness to adapt to adjusted tasks). This preference and competence based view is required to account for the heterogeneity and individual differences of employees, and necessary as job profiles are changing given adjusted or novel activities in human-machine interactions with automated or robotized systems (also driven by the rise of specific types of warehouses such as fulfillment centers for e-commerce, see Boysen et al. (2019)). Examples for salient relationships exist in the moderating effect of specific traits and preferences on efficiency (see Pasparakis et al. (2021) for a picking example) and retention, potentially triggered by differences in human motivation and satisfaction. Further, recognizing individual's skill set is important to understand which worker to deploy for which task, as individual human performance varies even in standardized activities (see Matusiak et al. (2017) for a related picking study). Given high fluctuation rates in general, and large temporary labor needs during peak demand periods, it is crucial to learn how to attract and retain labor for human-machine interactions. Hence, analyzing this first theme certainly leads to important understandings around individual differences that can be utilized to hire and train the right human at the right interaction across warehousing activities.

(B) Forming effective human-machine warehousing teams Having new team mates in the same *space* and *time* triggers human behavior from social psychology and group dynamics. For instance, theories around trust and (technological) acceptance (Glikson and Woolley, 2020) as well as peer-effects (Tan and Netessine, 2019) inform the behavior within the group, and consequently also the outcome of the human-machine interaction. To manage effective human-machine teams, interviewees describe trust and acceptance as key success factors to implement warehouse automation efficiently. This has been further accentuated by several practitioners that reported failed automation attempts, with large negative outcomes on system performance only due to lack of trust and acceptance by humans. Salient factors to consider are in particular perceptual factors (e.g., how robots are perceived and introduced) and their relationship to trust and acceptance of the employees in warehouse interactions. It is important how team or firm loyalty may vary given emerging human-machine setups, and how this moderates efficiency (potentially changed due to peer effects) and particularly retention. Only by including such behavioral mechanisms into optimization efforts will ensure to build efficient teams in the warehouse of the future.

As *space* and *time* are by nature components of coexistence, cooperation, and collaboration setups, themes (A) and (B) are of relevance for all human-machine interactions.

(C) Assigning tasks and developing operating policies among humans and machines Adding the interaction component *aim* sparks further mechanisms from social psychology. In particular, the same *aim* creates (task) interdependence, making the performance of humans and machines dependant on reciprocal actions (Bendoly et al., 2010). Thus, humans compare themselves with the machines, show social preferences (for example in task distribution or job execution), and react to process information and setup choices (see Gombolay et al. (2017) or Loch and Wu (2005) for related examples). The insights from practice regarding task assignment and policy development show that addressing how to best leverage the strengths of humans and machines, how to distribute the workload within human-robot teams, and how to design the workflow (such as communication and operating policies) becomes indispensable. When solving related issues, mental factors are salient as humans think about the tasks, process and setup choices (e.g., information provision). In this way, feedback, transparency and perceived autonomy influence motivation and satisfaction, and moderate the effect on performance in warehouses. This constitutes the pathway to explore different avenues in cooperation and collaboration setups such as the role of above-mentioned factors in substitutable versus complementary tasks.

(D) Designing engaging direct human-machine interactions Direct *contact* provokes further behavioral mechanisms from social psychology and physical environment (Bendoly et al., 2010; Vischer, 2007). Specifically, theories explaining behavioral factors such as engagement (e.g., goal-setting or incentive theories), workload (e.g., speed-accuracy trade-off) and situational awareness inform interactions, particularly in collaboration setups. Insights from the interviews show that, with increasing automation, experts see issues around designing engaging interactions and thus, struggle to create an attractive workplace for human employees in warehouses. While in cooperation setups mental and physical workload are mostly regarded separated, the direct *contact* among human and machines makes the balance of both factors a key relationship to optimize. For example, high physical workload (or speed-up pressure, see Schultz et al. (2003) or Wang and Zhou (2018)) and low mental workload may both reduce efficiency and quality. Hence, finding the optimal equilibrium (e.g., by adjusting decision-discretion) is a key area of future research. Goal-setting theory and incentive theories are starting points to inform more engaging (for higher usability and efficiency) and more sustainable (for higher retention) solutions. On top of that, a salient relationship exists between situational awareness and quality or safety (e.g., see Aromaa et al. (2020) for a related lab experiment), and needs to be taken into account when designing attractive interactions.

To summarize, the unifying themes (A)-(D) provide an aggregated view on the detailed issue discussion in Section 3.4 while the prevailing theories constitute the foundation for human-machine interaction research in warehousing going forward. This offers a cohesive body of knowledge to better understand causalities within human-machine interactions and to ultimately provide more efficient warehousing setups when behavioral factors are influencing the system and its performance. Addressing the research questions for the specific issues will therefore also result in transferable findings to other issues. The main rationale for this generalization materializes from analogous behavioral mechanisms that are triggered through similar *interaction setups* and, in some cases, *systems involved.* For instance, findings on fixed versus floating AMR policies when storing shelves [RQ

2.2] may be transferred to picking as the interaction setup and systems are comparable. Further, results on backlog design studies at advanced workstations for picking [RQ 5.4] are applicable for receiving and inspection as well as packing applications due to the potentially similar setup of the system and interaction. Also, insights on efficient operational policies that incorporate human usability and situational awareness when packing with AR glasses [RQ 7.2] are transferable to picking tasks. These three examples and the comprehensive overview in the Appendix indicate that findings generated for one issue provide opportunities in additional activities.

3.6 Conclusion

Interactions between human operators and automated or robotized systems in the warehouse are developing into a multi-disciplinary field of research. This has recently evolved and gained momentum due to the rapid growth of automation in logistics. As humans still excel in specific tasks due to flexible skills and economic advantages, new issues related to the role of workers in warehousing and in operations of the future emerged. To optimize system design and operations, it has become essential to investigate human-machine interactions in operational warehouse activities. This paper develops the pathway to necessary research within this nascent research area by identifying key interactions, corresponding behavioral issues, theoretical foundations, and unifying themes. We first developed a systematic framework to investigate issues in such interactions, and additionally presented our empirical findings from expert discussions. The developed research agenda unfolds open areas and related questions to better manage human-machine interactions in automated warehouses. The analysis of the warehousing literature revealed significant gaps across the identified issues. In addition to the novelty of the warehousing systems involved, a predominant reason is the research focus on either OM or BS, but interdisciplinary methods are missing to tackle those behavioral issues. It becomes evident that more synergistic approaches among OM and BS are required. Hence, we enriched the discussion to allow a cross-disciplinary perspective, which is in line with the call for interdisciplinary OM and BS research in Moniz and Krings (2016). We elaborated specifically on the type of human-machine interaction setup, and how its component are connected with prevailing theories. This overarching theoretical foundation particularly informs four emerging unifying themes for human-machine interactions in warehousing going forward. To conclude, we outline managerial and theoretical implications, and provide limitations as well as an outlook of our study.

3.6.1 Managerial and theoretical implications

Insights on the identified issues and themes could inspire practitioners when designing and planning modern warehouses. For example, the implications help warehouse systems providers and engineers to design better products (such as incorporating behavioral findings in design and setup choices of advanced workstations) and assist warehouse managers with better decision-making by accounting for human-machine interaction effects (such as hiring employees with a specific skill set or deciding on the type and number of robots for a team). These findings can also enhance project managers' awareness of behavioral issues when drafting implementation projects for warehouse automation (overcoming motivation and acceptance issues, for instance). Ultimately, insights into the issues and themes will facilitate the application of efficient OM models and tools that are grounded in empirical observations and behavioral theories, aimed at increasing system performance via enhanced human-machine interaction and associated factors. Furthermore, we demonstrate that interactions with systems of a robotic nature are prevalent, showing the enhanced relevance of autonomy in intralogistics (see Fottner et al. (2021)), also given their flexibility and scalability. Hence, managers need to prepare themselves and their teams for further human-robot interactions. Moreover, as unifying themes exist across activities, it is crucial to optimize interactions beyond picking (i.e., the activity with the largest cost share), as focusing on one individual activity may create bottlenecks in others, leaving behind untapped opportunities for improving an efficient material and information flow, and making a holistic research approach necessary (see also Boysen et al. (2021) or Van Gils et al. (2018)).

The systematic framework, theoretical foundation, and unifying themes also build a structure to advance human-machine interaction research. They can be applied in other contexts, particularly in both different activity levels and related OM fields. For the former, they can be transferred to the analysis of non-operational warehousing activities (such as interactions with intelligent maintenance software in automated warehouses). For the latter, the concepts remains valid for manufacturing (e.g., collaboration with assembly robots), transportation (e.g., supervising automated truck driving), health care (e.g., interactions with care robots) or other applications in supply chain management (see Perera et al. (2019) for a forecasting example). Consequently, this work opens up a broad variety of relevant topics as human-machine interaction continues to progress in many OM fields.

When finding solutions to above-mentioned issues and unifying themes, blending research of machine-centric OM with human-centric BS by applying the systematic framework and theoretical foundations is vital to establish efficient human-machine interactions in the warehouse. In order to enhance decision-making as well as OM principles and theories for such interactions, it is also important to utilize a variety of methods to address the research questions proposed. While we acknowledge that selected issues may be resolved using a single method, it becomes indispensable to apply an integrated research approach for the majority. This requires utilizing quantitative methods (such as simulation, optimization or analytics) and transferring principles from OM to human-machine interactions. These need to be based on empirical insights using experimental and field research or surveys to capture the actual behavior of agents involved and test existing and nascent theories (DeHoratius and Rabinovich, 2011). Using our developed systematic framework and theoretical foundation, the combination of both lenses will inform OM models, theories and principles (Bendoly et al., 2006), which ultimately enhances system performance to a greater extent. We refer to the 3.6.2 for two examples how future research can take such an integrated path.

3.6.2 Limitations and outlook

The list of research questions in one paper can never be exhaustive. We mitigate this problem by conducting expert interviews to detect the most relevant issues to explore this emerging field. By the design of this research, we have concentrated our efforts on blue-collar, operational activities and have not extended our perspective on white-collar planning tasks (whether tactical or strategic). For instance, in control rooms of automated warehouses, a common issue is overwriting optimal parameters for automated systems by operators. This often happens based on individual human preferences, or unknown information. Consequently, future work could explore issues in these directions. Additionally, we did not focus on integrative topics for human-machine interactions when the systems are in the early phase of technological developments and implementation. However, the systematic framework, theoretical foundation, and unifying themes can also serve to solve such matters including the alignment of human-machine navigation (e.g., how to avoid the dominance of humans at intersections with AMRs), the supervision of robotic systems such as inventory counting drones (e.g., how to deal with low situational awareness), and the integration of robotic exoskeletons (e.g., how to increase the acceptance and usability of such supportive devices). Moreover, due to the novelty of the systems and interactions, it is not yet possible to derive any inference on the long-term implications driven by behavioral mechanisms, which constitutes a further research opportunity.

In conclusion, we see growing opportunities for managerially relevant and theoretically challenging investigations in the field of human-machine interactions in general, and in the context of warehousing in particular. Elon Musk, CEO of Tesla Inc. and one of the strongest advocates of technologically induced change, fittingly said: "*Humans are underrated.*" It was a reaction to over-automation without balancing human and machine skills at Tesla's production facility in California (Edwards and Edwards, 2018). Our contribution will serve to stimulate this line of research and further enhance the blending of novel automated or robotized warehousing systems with human factors and behavior.

Appendix

Appendix A: Research Methodology

In the following, we provide additional information on our research methodology along the steps within our triangulation approach. We start with details on how we developed our *(1) conceptual foundation*, follow with comprehensive information on our *(2) expert interviews*, and finally highlight the procedure of the conducted *(3) systematic literature analysis*.

A.1 Conceptual foundation

By analyzing seminal research, we sketch out fundamental literature and concepts to rely on accepted definitions and relationships, and generate a common understanding across the related research domains of human-machine interactions in warehousing:

- First, machine-centric operations management offers recent works on novel automated and robotized systems (including Azadeh et al. (2019) or Boysen et al. (2019, 2021)), and seminal research on warehousing (such as Gu et al. (2007b) or De Koster et al. (2007)). These sources help to identify which systems and activities result in important interactions and issues.

- Second, we draw on theories from the human-centric behavioral science stream that are relevant for behavioral operations (see discussion papers of Bendoly et al. (2006, 2010), Croson et al. (2013), or Loch and Wu (2005) for instances). These include, but are not limited to, cognitive psychology (such as anchoring or framing), social psychology (such as feedback and control theory, or technological acceptance), experimental economics (such as incentive schemes or nudging), and group dynamics (such as teamwork dynamics). In this way, we incorporate existing behavioral theory from related fields, particularly to uncover behavioral aspects in human-machine interactions (Boyer and Swink, 2008). Additionally, we leverage the heterogeneity of such theories based on individual characteristics such as personal skills, competencies, behavioral traits (e.g., personality types) or demographic aspects (e.g., culture and age). This is reflected at a later stage by varying such factors in the respective research questions. We also screen behavioral work in manual warehouses as part of the analysis (e.g., Batt and Gallino (2019), or Grosse et al. (2015, 2017)).
- Finally, we additionally study foundations of human-machine interaction outside the warehousing domain (e.g., Karwowski (2005); Salvendy (2012); Sanders and McCormick (1993); Schmidtler et al. (2015); Schulz et al. (2018)). This way, we also examine the links of theories and variables within such interactions, and finally, can propose a novel research approach applied to the context of warehousing.

This research step delivers the foundation for the systematic framework in Section 3.3, which denotes the relationships among important building blocks of human-machine interactions in warehousing.

A.2 Expert interviews

To ensure external validity and enhance the practical relevance, we collect primary data and conduct semi-structured interviews (McCutcheon and Meredith, 1993). According to Qu and Dumay (2011), this is particularly suitable when disclosing important facets of human behavior (see Smith et al. (2009) or Wu and Pullman (2015) for similar approaches). We interviewed in total 19 warehouse system providers, warehouse managers and intralogistics consultants. We applied theoretical sampling for our interviews (Eisenhardt, 1989; McCutcheon and Meredith, 1993), which took place between March and June 2020 with ongoing data analysis after each interview. We hosted audio and video conferences that lasted 50 minutes on average. Table 3.A1 provides an anonymous overview of the participants including the order of the interviews. The selection process resulted in a sample that shares internal homogeneity (i.e., experts in human-machine interactions for operational warehousing activities) and external heterogeneity (i.e., experts from different steps of the value chain) to ensure a holistic approach (see Wu and Choi (2005) or Trautrims et al. (2012) for examples). The interview questions (see Table 3.A2) were probing which human-machine interactions exist in the different operational warehouse activities, and which associated behavioral issues the experts observe. All interviews were subsequently transcribed and coded using data analysis software (Miles et al., 2013). At regular meetings, all authors discussed the codes, categories, and findings to set aside subjective impressions and come to an objective meaning of interviewee perceptions to ensure repeatability of our insights (Lincoln and Guba, 1985). Further information on the interview procedure can be found in Table 3.A1 and 3.A2. The empirical findings build the main source for describing the human-machine interactions and identifying the issues. Seven categories (each category represents one issue) were derived from the interpretation of the data. Section 3.4 is structured along the seven issues.

Table 3.A1: Anonymous overview of interviewees and supplementary information

Code	#	Company type	Interviewee role	Warehouse experience years	Gender
1 SP	1	System provider	Managing Director	> 20	Male
2 WO	2	Warehousing operator	COO	10 - 20	Male
3 WO	3	Warehousing operator	Head of Intralogistics	10 - 20	Male
4 WO	4	Warehousing operator	Operations Manager	< 5	Female
5 C	5	Consultancy	Partner	10 - 20	Male
6 C	6	Consultancy	Senior Expert	> 20	Male
7 WO	7	Warehousing operator	Supply Chain Manager	10 - 20	Male
8 WO	8	Warehousing operator	Warehousing Manager	> 20	Female
9 SP	9	System provider	CEO	5 - 10	Male
10 WO	10	Warehousing operator	Warehousing Manager	10 - 20	Male
11 SP	11	System provider	Senior Product Manager	10 - 20	Male
12 WO	12	Warehousing operator	Warehousing Manager	5 - 10	Male
13 WO	13	Warehousing operator	Logistics Manager	< 5	Male
14 WO	14	Warehousing operator	Warehousing Manager	> 20	Male
15 SP	15	System provider	CTO	> 10	Male
16 SP	16	System provider	Senior Product Manager	10 - 20	Male
17 C	17	Consultancy	Senior Expert	> 20	Male
18 SP	18	System provider	Head of R&D	10 - 20	Male
19 WO	19	Warehousing operator	Supply Chain Manager	> 20	Male

Supplementary information on sampling. We started by screening the global top 20 system providers (Modern Materials Handling, 2019). Four of them reported prominently about implementing major warehouse automation projects in the press and on conferences. We invited these four and three participated. We also mirrored the dynamic landscape for warehouse automation and reached out to three innovative smaller providers for novel systems identified through press clippings, conferences, and further references. Two of those joined us for interviews. In the same manner we identified potential warehouse operators and managers as well as intralogistics consultants, and screened recent implementations (see e.g., WEKA (2020)), too. We reached out to fourteen contacts, of which eight operators and three consultants participated. After these 16 interviews many issues and interactions were identified, and the repeatability already increased from session to session. To ensure we achieved information saturation, we conducted three more interviews (system provider, warehouse manager and consultant). As these did not reveal any new insights (i.e., new codes that resulted in interactions or issues identified), we concluded information saturation for the most relevant interactions and issues (Holton, 2012). Nineteen interviews is in-line with recommendations to ensure academic rigor and generalizability (see e.g., Eisenhardt (1989), Guest et al. (2006)).

Table 3.A2: Guiding questions for interviews and supplementary information

#	**Guiding questions**
1	In which areas or activities in the warehouse do you see significant interactions among humans and automated machines or robots?
2	In the identified areas or activities, which (behavioral) issues do you see in the human-machine interaction?
3	In which activities do humans and machines substitute each other; in which activities do humans and machines work complementary?
4	What behavioral influences does the interaction have on humans (e.g., motivation, acceptance, attention)?

Supplementary information on interview procedure and analysis. We deal with the investigation of new structures and processes. Qualitative research is particularly suitable for such settings (Bryman and Bell, 2015). A semi-structured interview approach with open-ended questions has been applied to retrieve relevant information and gain sufficient flexibility, which is appropriate when exploring a rather little-known area of research (Creswell, 2009; DeHoratius and Rabinovich, 2011; Edmondson and Mcmanus, 2007). We referred to the main operational warehouse activities and potential systems involved derived from the theoretical foundations to guide through the discussion. After asking open-ended questions about interactions and issues, probes were informed by potential associated human factors and behavior. Interviews were conducted in German with German-speaking participants and in English with other participants. We based our inductive analysis neither on a deductive logic nor a strict grounded theory approach (Randall and Mello, 2012), as "*data are inextricably fused with theory*" (Alvesson and Kärreman, 2007, p. 1265). The interviews were analyzed in two layers. First, an objective content analysis was conducted focusing on the identification of relevant human-machine interactions. In a second layer, we concentrated on the behavioral issues identified, associated human factors, and the impact on system performance to extract the underlying behavioral aspect in those interactions (Trautrims et al., 2012). After establishing relevant human-machine interactions available from the content analysis in the first layer, the second layer of analysis required deconstruction of the data for the extraction of tacit knowledge from the interviews. The transcripts were rephrased, reflected on and compared to create meaningful categories (Eisenhardt, 1989; Trautrims et al., 2012).

A.3 Systematic literature analysis

To match the identified issues with existing warehousing research, we perform a systematic literature analysis. This ensures a comprehensible and objective process (Snyder, 2019). We utilize a fourfold approach, starting with a keyword-based search on Scopus and Business Source Premier. For the sake of focus, only peer-reviewed articles written in the English language from 2010 or later that conduct experiments in the context of human factors or behavioral issues in interactions with automated and robotized warehousing systems are considered. Initial screening and selection (including eliminating duplicates) is conducted by three team members based on title, abstract and keywords. Subsequently, suitable articles are read and either included (if they match the above-mentioned criteria) or excluded. Second, the reference sections of selected articles were screened to identify further matching work (snowball method). Third, we use Google Scholar to analyze any articles that cited selected research from step one and two to further find matching articles. Fourth, manual searches of leading journals in the field are carried out. As an outcome, we screened a very large number of papers (2,218) to reflect the interdisciplinary nature of the research. For further information on the process we refer to Table 3.A3. Ultimately, we identify 13 articles that are matched to the identified issues in Section 3.4 to mirror state-of-the art research.

To summarize, we utilize the interviews and literature in Section 3.4 within our issue identification approach. The systematic framework in Section 3.3 developed from the conceptual foundation lines up this discussion

Table 3.A3: Keywords utilized in literature review and supplementary information

Area A: **Warehousing**	Area B: **Interaction**	Area C: **Human Factors and Behavior**		
Warehous*	Autonom*	Behavio*	Situational Awareness	Training
Intralogistics	Robot	Human Factor*	Motivation	Supervisi*
Distribution center	Automat*	Human Considerati*	Satisfaction	Moving
	AGV	Ergonomics	Loyalty	Operating
	Human-machine	Psychosocial	Fairness	Learning
	Human-robot	Physical	Confidence	Team structure
	Machine	Decision making	Skills	Team setup
		Cogniti*	Percept*	Diversity
		Safety	Sens*	Resistance
		Bias	Augment*	Commitment
		Heuristics	Mental	Adoption
		Trust	Think*	Stress
		Acceptance	Information process*	Emotion
		Workload	Boredom	Attention

Supplementary information. Any combination of the keywords from the first, second and third area in the abstract, title or keywords qualified for a hit. The keywords from the first area have been chosen to ensure we target warehouses, while the second area mirrored the interdisciplinary nature of human-machine interactions. The third area was used to ensure behavioral experiments and settings were researched. We excluded any paper with the word *data warehouse* in the abstract, title or keywords. Additional manual searches were carried out in the following journals: Management Science, Production and Operations Management, Journal of Operations Management, Manufacturing and Service Operations Management, European Journal of Operational Research. A sample of approximately ten percent of all articles are initially screened by two people in parallel to ensure consistency. No significant deviations in terms of the articles selected could be identified among the three members. In the rare event of different classifications, articles are marked as relevant to avoid missing related research. The search have been updated in December 2021 during the revision to include recently published papers.

Appendix B: Transferability of findings and integrated research approaches

In the following Table 3.A4, we show how findings generated for one issue provide opportunities in additional activities. Subsequently, we highlight two examples for integrated research approaches when tackling the identified issues.

B.1 Transferability of findings

Table 3.A4: Findings on relevant issues may be transferred to other operational activities

Issue		Interaction setup	Systems involved	Operational activities				Related literature
				R&I	S	OP	P	
[1]	Controlling quality using human-machine complementarities	Cooperation	Automated inspection and control system	✓			(✓)	
[2]	Filling shelves with autonomous mobile robots	Collaboration	AMRs with storage function (and potentially lift & seat)		✓	(✓)		Pasparakis et al. (2021), Roy and Edan (2018)
[3]	Building teams in human-robot picking setups	Co-existence	Fully autonomous mobile picking robots		(✓)	✓	(✓)	
[4]	Distributing work in human-robot picking teams with substitutable tasks	Cooperation	Fully autonomous mobile picking robots, Pick assignment machine		(✓)	✓		Bai et al. (2021), Sanders et al. (2019), Cragg and Loske (2019)
[5]	Overcoming mental impoverishment and physical overload at advanced workstations	Collaboration	Advanced picking workstations	(✓)		✓	(✓)	Kudelska and Niedbal (2020)
[6]	Forming dyads with robot-assisted packing machines	Cooperation	Robotic packing systems	(✓)		(✓)	✓	Banerjee et al. (2015), Maettig and Kretschmer (2019), Maettig et al. (2019), Sun et al. (2021)
[7]	Solving the quest of augmented reality for packing	Collaboration	Augmented-reality head-mounted display	(✓)	(✓)	(✓)	✓	Stoltz et al. (2017), Kretschmer et al. (2018), Plewan et al. (2021)

Examined activity ✓, Transferable to (✓); R&I = Receiving & Inspection, S = Storing, OP = Order picking, P = Packing

B.2 Integrated research approaches

As mentioned above, it is important to apply an integrated research approach for many of the identified issues. For example, when finding the optimal team structure [3.1], field experiments are suitable for analyzing behavioral aspects (such as peer effects) and performance metrics with a varying share of robots. These behavioral aspects can then be included in mathematical optimizations (e.g., Solow et al. (2020)) that model performance as a function of the proportion of robots being used by accounting for the actual human behavior. Hence, decision-making on the optimal number of humans and robots including the resource allocation can be empirically enhanced to improve organizational capabilities and system reliability. Moreover, when finding the most efficient operating model for advanced workstations [5.4], behavioral implications of design and setup choices need to be assessed. One could start analyzing the impact of backlog design options (and the underlying reasons) on human performance criteria by conducting field or lab experiments. The results could then be implemented in simulation studies (e.g., using digital twins) on which backlog design options are preferred, even with different personality types. Ultimately, this results in a multi-criteria model that optimizes both throughput and human factors. These were just two examples out of many that show how future research can take such an integrated path.

4 It's in your hands: Elevating performance with goal-setting at the cost of social discord in an intervention-based human-machine interaction study

Co-authors: Andreas Fügener and Alexander Hübner
In submission process to *Journal of Operations Management* as of August 09, 2022

Abstract In course of the digital transformation, activities in operations management contexts are prone to automation. Still, humans play an important role in many settings, and human-machine interactions need to be managed efficiently to ensure smooth operations. In many of those interactions, machines determine the assignment and sequencing of tasks, while human workers mainly execute repetitive and monotonous activities. One downside of such settings is the mental impoverishment of workers which relates to stagnating productivity coupled with undesired effects on human factors such as low satisfaction, self-determination, and perceived fairness. To address these shortcomings, we perform an intervention-based research field study in a semi-automated grocery warehouse, where we enable human workers to decide the number of picks they want to perform at their current workstation. While we observe a 5.6% increase in performance, workers report decreased levels of satisfaction, perceived fairness, and self-determination. Triangulating surveys, focus interviews, and practitioners' discussions revealed that the intervention led to the suspension of informal work arrangements,

resulting in the deterioration of the human factors. Our insights contribute to the growing field of addressing behavioral issues in human-machine interactions, and provide new insights to merits and potential pitfalls of applying goal-setting interventions.

4.1 Introduction

The interaction of human workers and machines is increasingly becoming a central component of operations management research (e.g., Sun et al. (2021)). For decades, expanding automation and robotization has been the focal point across operations contexts to achieve faster throughput times, reduced costs, and higher service levels (IFR, 2020). Still, automated systems are often jointly utilized with humans as manual workers have certain advantages in flexibility and skills, resulting in a variety of human-machine interactions (Lorson et al., 2022; Olsen and Tomlin, 2020). In many of those interactions, machines determine the assignment and sequencing of tasks, while human workers mainly execute repetitive and monotonous activities (see, e.g., Bai et al. (2021); Sun et al. (2021); Wang et al. (2021)).

This particular division of work causes novel behavioral issues within human-machine interactions. Mentally, the machine governs most steps of the working process and dominates the decision-making, whereas simple, repetitive, and physically exhausting tasks need to be completed by the humans. Among others, many companies across operations and logistics experience lower levels of human satisfaction (McKinsey & Company, 2021b), self-determination (Parasuraman et al., 2000), and perceived fairness (Langer and Landers, 2021; Newman et al., 2020; Robert et al., 2020) that can relate to the mental impoverishment of employees (Lorson et al., 2022). Not surprisingly, stagnating operational worker productivity and overall system performance paired with high employee turnover are common consequences (McKinsey & Company, 2021b). While maximizing performance for repetitive and monotonous operational activities plays a major role in many organizations' success (Bernstein, 2012; KC, 2020; Staats and Gino, 2012), little research is available to account for the particularities of human-machine interactions. Thus, existing empirical findings and behavioral theories on managing human factors and worker productivity (see Bendoly et al. (2006, 2010); Croson et al. (2013); KC (2020)) need to be leveraged and extended to solve this behavioral issue for human-machine interactions.

One promising approach is the introduction of goal-setting interventions for human workers. Working towards goals has successfully increased performance across different contexts, subject groups, geographies, and sources of the goal (Corgnet et al., 2015; Goerg and Kube, 2012; Schultz et al., 2010; Van Lent and Souverijn, 2020). Particularly, goal-setting applications also improved human factors (Locke and Latham, 2002) within operations management contexts. For example, Doerr et al. (1996) demonstrate how individual goals improve worker satisfaction. Setting participative goals includes the human worker into the decision-making process of her activities, which has shown to be a source for self-determination (see Deci and Ryan (2000); Deci et al. (2017) for overviews) and a significant factor in the evaluation of fairness (Cropanzano et al., 2008). Given that goal-setting is an important facet within operations management (Bendoly et al., 2010) with potential to solve mental impoverishment and stagnating performance of workers, an extension of the theory towards human-machine interactions of blue-collar workers in operational activities is promising.

Our research aim is to improve the human-machine interaction by introducing a participative goal-setting intervention in a blue-collar operational context. We hypothesize that such intervention improves both system performance and satisfaction, self-determination, and perceived fairness. This leads to the research question: *How is a participative goal-setting intervention impacting performance and human factors within a human-machine interaction for an operational, monotonous activity?*

To approach this, we conduct a field study within an Intervention-Based Research (IBR) approach (Chandrasekaran et al., 2020; Olivia, 2019), which has been successfully utilized in behavioral research questions across operations management contexts (Akkermans et al., 2019; Chun et al., 2022). It particularly provides the opportunity to observe human actions and behavior in monotonous processes over a longer period. We choose order picking within a warehouse as our research context where humans and machines jointly work together to finish the picking task, with human activities being highly monotonous and repetitive. The presence and importance of behavioral issues within the human-machine interaction in this setting will allow us to obtain generalizable insights. To do so, we collaborate with a warehouse to introduce an intervention within a semi-automated order picking zone. We design a participative goal-setting intervention in which pickers are able to choose how many items they want to pick at their current workstation out of a set of five pick numbers (goals).

To facilitate the intervention, we draw upon the "Context-Intervention-Mechanisms-Outcomes" (*CIMO*, see Denyer et al. (2008)) framework, which is frequently used in IBR (e.g., Akkermans et al. (2019); Friesike et al. (2019); Groop et al. (2017)). We analyze the effect of the intervention in a real-world field study with pickers over the course of eight weeks. The picking performance improved by 5.6% during the intervention compared to pre-intervention, suggesting that goal-setting is a suitable approach to increase system performance in monotonous human-machine interactions. This result is replicated across a set of robustness checks, including a Difference-in-Differences analysis considering a control warehouse. However, employee satisfaction, self-determination and perceived fairness deteriorated compared to pre-intervention scores and to a control group that was not part of our intervention. The negative impact on human factors was further explored by triangulating two types of human factor surveys, focus interviews, in-depth process analyses, and discussions with the respective practitioners. We could observe that the goal-setting intervention suspended informal arrangements among workers. Specifically, the intervention affected possibilities for humans to informally organize themselves in their working day and to overrule the machine, with repercussions on satisfaction, self-determination and perceived fairness.

Our research offers theoretical and practical contributions to operations management. First, we demonstrate that the goal-setting intervention indeed has the potential to improve performance in highly physical, operational activities without any kind of monetary incentives, and extend the theory given the potential pitfall of affecting informal arrangements and their influence on human factors. Second, we demonstrate that behavioral aspects of individuals and teams are important and need to be accounted for when designing human-machine interactions and the associated workflows and processes. This is in particular a contribution to the warehousing literature where human factors have been mostly ignored in prior work related to automation (Azadeh et al., 2019; Boysen et al., 2020; Yuan et al., 2019). Third, we illustrate the value of conducting an IBR study in warehousing. Running interventions during the regular course of action in a warehouse setting enables us to draw robust inferences and general insights that normal lab experiments are not able to reveal, such as effects of informal work arrangements as in our study.

The remainder is organized as follows. Section 4.2 details our research methodology. We follow with a detailed explanation of the empirical setting and research design using the *CIMO* framework in Section 4.3. Section 4.4 highlights results, while Section 4.5 discusses those from a managerial and theoretical perspective. We close in Section 4.6 with a conclusion and outlook.

4.2 Research methodology

This study follows an IBR approach (Chandrasekaran et al., 2020; Olivia, 2019) which has been applied to tackle complex behavioral research questions in many operations management contexts such as inventory management (Land et al., 2021), scheduling (Öhman et al., 2021), manufacturing (Hedenstierna et al., 2019) or healthcare (Anand et al., 2021; Chun et al., 2022; Song et al., 2018). As we aim to explore reasons behind potential changes in system performance and human factors, an IBR study is appropriate to identify how individual workers and teams behave and are affected during the intervention (Langley et al., 2013; Olivia, 2019). In this way, following an IBR approach requires further in-depth analyses of observed human behavior, interactions, and processes to better understand and explain unknown phenomena. This supports detecting unexpected findings and triggers abductive reasoning to elaborate on the outcome of the research. We leverage the IBR approach across our study to obtain empirical findings, analyze the outcomes, and enhance existing theories. By doing so, the continuous iteration among empirical evidence obtained within an IBR study and existing behavioral theories enhances the theoretical insights in operations management (Chandrasekaran et al., 2020).

Analyzing repetitive and monotonous operational activities requires to capture agent and team behavior over a longer term during the normal course of business. Working monotony and behavioral issues related to mental impoverishment develop over time and thus require investigations in settings where agents are already active for a longer period. Furthermore, these repetitive and monotonous activities are subject to distinctive characteristics (such as mental and physical impact within the human-machine interaction), workflows and processes (such as developed job routines), and specific team dynamics (such as pre-existing relationships among workers or informal team arrangements for working procedures). Field research allows to effectively investigate such behavioral facets in established team settings and human-machine interactions. It further offers the opportunity to demonstrate external validity and still understand phenomena in detail (Ibanez and Staats, 2018). In this way, an IBR field study offers the chance to move away from the quintessential "ivory tower" syndrome (Van Aken et al., 2016; Van Mieghem, 2013) and also to go beyond expected results as some influencing factors on operational performance and human behavior may not be obvious before. IBR in the field thus can generate new empirical and theoretical insights by closely iterating between practice and theory (Van Aken et al., 2016). Accordingly, a field study with an industry cooperation is well suited for enhancing empirical findings in operations management and behavioral science by observing human actions and behavior in monotonous daily working processes over a longer period of time.

To do so, we choose order picking in warehouses as our research context for three main reasons. First, the controllable and structured environment makes warehouses popular incubators for the development and application of innovative automated and robotized systems (Azadeh et al., 2019; Fragapane et al., 2021). As of now, such systems often only take over part of the operational activity (e.g., order picking), since humans have distinctive characteristics, skills and capabilities that robots are not able to replicate or able to perform cost efficiently (Gombolay et al., 2015; Schäfer et al., 2022; Sgarbossa et al., 2020). As a consequence, humans and machines work alongside each other, resulting in a variety of human-machine interactions and behavioral issues (Lorson et al., 2022; Olsen and Tomlin, 2020). The necessary collaboration of human operators with a growing diversity of machines including resulting behavioral issues thus exposes warehousing as an unique research area for human-machine interactions. Second, the human task of order picking consists of repetitive physical motions such as walking, grabbing, scanning and putting, making it a generic example of a monotonous task where issues related to mental impoverishment and stagnating worker performance are present. While our direct context of research is warehousing, we expect our findings to be generalizable across similar tasks in different applications across operations management. Third, enhancing the value of human factors becomes increasingly relevant in warehousing. Given the ongoing labor shortage and high worker turnover, a recent shift in the perspective on blue-collar labor from an exchangeable resource for completing open jobs to a valuable asset for smooth and sustainable operations with the potential for productivity improvement, was observed in warehousing, but holds true as well as for other operations areas (McKinsey & Company, 2021b).

Table 4.1 summarizes the phases of our IBR study that are highlighted below.

Table 4.1: Research phases and timeline

(1) Pre-intervention *Aug 2020 - Jan 2022*	**(2) Intervention** *Feb - Mar 2022*	**(3) Post-intervention** *Apr - Aug 2022*
• Interview warehouse, provider, warehouse manager, and pickers • Explore and identify behavioral issues in order picking • Determine type of intervention • Program intervention and adjust front-end for login • Beta-testing intervention	• Brief pickers about adjustments in workflow • Conduct intervention • Collect system performance data • Survey human factors including control group • Host focus interviews with pickers	• Analyze the effect of the intervention and report on results • Identify how individual pickers and team leaders participated and are affected • Provide theoretical explanations for outcomes

(1) Pre-intervention A leading warehouse system provider served as business and thought partner to identify relevant behavioral issues in the field. Upon selecting a suitable research site of a grocery retailer and embarking on the project jointly with the respective warehouse manager, we conducted a series of interviews, ran data analyses, and picked on-site with the employee group to gain first-hand experience. After identifying the most prevalent issues, we started analyzing potential approaches and designed an intervention based on feasibility, impact, and research gap. A software engineer programmed the intervention in the back-end of the warehousing software and created a suitable front-end design based on the conceptual input of the research team. Within this process, several beta-tests have been conducted to ensure a smooth execution.

(2) Intervention The intervention was started within a representative period to avoid any operational and load-related distortions (e.g., due to public holidays). Before going live, we had detailed group briefings on the functionalities. We started the intervention with the goal to conduct the study for five weeks, however, were able to run it for eight weeks. During this time, we collected the system performance data in the same manner as pre-intervention to ensure a comparable data foundation. To survey human factors, the pickers filled out weekly questionnaires one week prior and during the implementation. Additionally, we hosted focus interviews five weeks after the intervention. We also run the surveys with a control group of employees in a different part of the warehouse that was not part of our intervention.

(3) Post-intervention After completion of the intervention study, we analyzed the effect of the intervention and elaborated on the outcomes. The conducted focus interviews with pickers and shift leaders as well as discussions with the warehouse manager and department head about the observed findings allowed us to generate tacit knowledge about picker actions and behavioral aspects.

4.3 Research design and empirical setting

The *CIMO* framework of Denyer et al. (2008) is used to develop our study and lines up this section. It is one of the dominantly used research designs in IBR studies (see, e.g., Akkermans et al. (2019); Groop et al. (2017); Ilk et al. (2020); Johnson et al. (2020)), and structures the analysis and implementation of organizational interventions including information on the specific *context (C)* in which the *intervention (I)* is implemented, while elaborating on the *mechanisms (M)* that drive the expected *outcomes (O)* (Friesike et al., 2019). Figure 4.1 shows an overview.

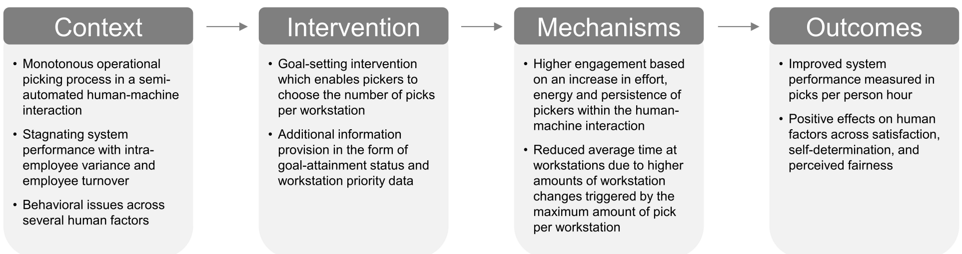


Figure 4.1: CIMO logic lining up Context, Intervention, Mechanisms and desired Outcomes

4.3.1 Context (C) and research problem

Picking in a semi-automated human-machine interaction The field study takes place in a large and modern distribution center of a leading grocery retail chain in Western Europe. The site is jointly run by the retailer and the warehouse system provider. The provider is a global market leader in the design and operation of automated warehouses. The retailer and warehouse operator supplies more than 1,000 stores from the warehouse every day. After the receiving and inspection of incoming pallets, products are stored and subsequently retrieved when required. We consider the retrieval of items with a semi-automated pick-to-light order picking system as the ideal research context. Humans and machines jointly work together to finish the picking task, with human activities being highly monotonous and repetitive. The picking process under investigation operates as follows. When one or more units of a product are needed, a conveyor-based system sends a transport box to the picking area, consisting of two aisles with six advanced workstations each. The system lines up the boxes (with items inside) either to the left, to the right or behind the respective workstation, where displays and lights guide the human picker in retrieving the right amount and the right product (see Figure 4.2 for a layout overview). Items are then placed into a box at the workstation and, once the picker retrieved all necessary items, transported either to another workstation or to a consolidation area. Typical products are cigarettes, drugstore articles or canned food. Employees work usually in three shifts per day (night, early and late shift) and five days per week.

We refer to Figure 4.3 for exemplary images of the picking zone. Usually, nine to ten pickers work in one shift on twelve workstations. To serve all workstations and as the workstation replenishment takes longer than the picking, pickers need to switch between workstations. Human pickers thus frequently change those, depending on the workload distribution or schedule, with an average of every 14 minutes. The assignment of human pickers to the workstations is centrally steered by a "machine". An algorithm, the so-called picker guide, defines the sequence of workstations for each

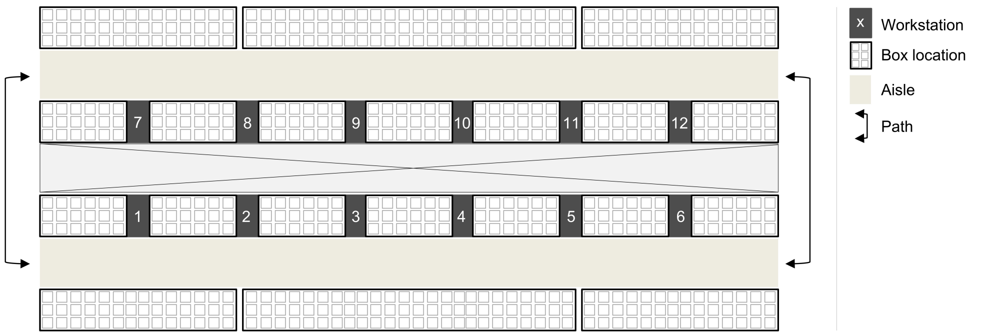


Figure 4.2: Layout of picking area

picker. The priority of each workstation is continuously updated by the algorithm and based on workstation workload, picker proximity or delivery deadlines, among other criteria. After a picker has logged into the assigned workstation, she completes the number of picks defined by the algorithm. In this standard operating procedure, the picker has no decision discretion where or how many items to pick before she needs to move on to the next workstation. The twelve workstations inhibit differences driven by their allocated products and mechanisms. Two workstations deal with "slow-mover" items, leading to larger than average waiting time between arriving boxes. In general, one aisle has lighter products (e.g., deodorants), while the other one deals with the heavier items (e.g., ketchup bottles). These characteristics lead to different human preferences and performances along workstations. Most of the pickers favor to pick in the light gear aisle, especially at the cigarette workstations as a high picking efficiency is easiest to achieve. Contrary, pickers dislike the heavy gear aisle as well as the two slow-mover workstations.

Performance and human factors issues Human workers are exposed to monotonous and repetitive operational activities in warehousing due to the setup of the human-machine interaction. This may cause performance and behavioral issues, particularly where an algorithm determines the job

Figure 4.3: Close-up view (left) and distant view (center) on advanced picking workstations, and aisle view (right)

sequence and amount of work at a workstation such as in our application. Our starting point to comprehensively identify behavioral issues in human-machine interactions were several semi-structured interviews with the warehouse manager, the department head, shift leaders and individual pickers from the retail chain as well as with system designers and management from the warehouse provider. These interviews were conducted via video/audio-conferencing tools and on-site field meetings. Also, one author joined the picker group pre-intervention to experience the activity first-hand. Additionally, we complemented our qualitative findings from the interviews with data analyses on quantitative system data to validate reported issues and completed a literature review. Table 4.2 summarizes the main observations. We discuss the issues along the framework of Lorson et al. (2022) and differentiate between performance and human factors.

Table 4.2: Gaining insights into human-machine interaction problems across sources to identify prevalent issues

Source	Important problem statements	Identified Issues
Warehouse manager	"There is often a general dissatisfaction, it's just a hard job." "Pickers are not involved in any decision-making processes, maybe only in micro decisions. The picker guide algorithm decides for each picker where to pick and when to switch workstations." "As in many other companies in the logistics sector, we have a high employee turnover and need to improve our retention." "Picker have own tally charts to check when someone worked in which aisle. They often feel unfairly treated." "The performance overall did not change much over the last years." "We have individual performance information for each picker, but we typically do not share details with them." "We often see pickers questioning the decisions of the picker guide."	Performance Stagnating performance Human Factors Low satisfaction Little self-determination Perceived unfairness
Department head	"Sometime pickers have to wait for the replenishment at a workstation and that is very boring and frustrating, particularly at the slow mover workstations 6 and 12." "There are conflicts and discussions between pickers, mostly about break time and workstation assignment." "Some colleagues use the log-off function or go to the bathroom many times. When they come back, they choose the workstation they want." "The performance among pickers is different. We have very strong ones, but some also work slower than the average. Overall, performance stays mostly the same."	
Shift leaders	"I want to ensure that workstation 6 and 12 are always occupied not to miss delivery deadlines." "With some colleagues I have to pay attention that they do not spend too much time at the preferred workstations 1 and 2 or in the easy aisle."	
Pickers	"We do not get any reward for picking faster. We only have the fixed salary." "Work is work and it is always the same." "Some colleagues always want to stay longer at workstation 1 and 2 because there you can get many picks."	
Data analyses	Non-uniform picking performance across workstations Measurable deviation of picking performance across pickers High employee turnover No significant differences in the workstation assignment across employees Stagnating overall picking performance over the last years	

The warehouse manager and the department head indicated that the overall picking performance was stagnating. A data analysis of the previous four years confirms these observations. A linear regression indicated a very small (0.02% of the constant) and insignificant (p=0.72) slope over time. Further, the human-machine interaction has not been innovated over the last years, and the work of the human pickers remained unchanged. Please note that other performance metrics such as quality of work within the picking process was not of relevance, as before and during the intervention only a negligible number of picking errors occurred. The same holds true for safety related metrics (De Koster et al., 2011).

As common in the logistics sector, turnover among employees is high. For example, 40% of the employees that were active in the respective study weeks in 2021 were not present in 2022. Related to this is the low satisfaction of the employees. One obvious reason is the monotonous work of the pickers: "*Work is work and it is always the same*", to cite an employee. Also, missing performance feedback was mentioned several times, as pickers do not receive direct, consistent, and comparable feedback on their picking performance. Furthermore, a perceived unfairness among the pickers with respect to the allocation to less or more preferred workstations has been raised as an issue by the pickers and leadership. Many of them verbalized an unequal assignment to the different workstations. However, with a look into the data, this unfairness can be labelled as perceived, but not actual. All pickers have a similar share of assignments across workstations including preferred and unfavorable ones. Little self-determination constitutes a further human factor issue, as the "machine" is taking all necessary decisions within the work task: "*Pickers are not involved in any decision-making processes*", to cite the warehouse manager.

4.3.2 Designing and implementing the intervention (I)

The system under investigation showed issues of stagnating picking performance and human factors related to low job satisfaction, perceived unfairness and little self-determination. The goal of the behavioral intervention was to efficiently tackle these identified issues, while being conscious about the overall technical change effort and the possibility to implement the solution. For this purpose, we defined each variable involved in the underlying issue, and explored options how to improve those. We jointly created with the management team of the warehouse operator and the warehouse provider a list of potential interventions that are suitable to address those issues, and subsequently evaluated them along feasibility of implementation, expected impact on operations, and research gap. As we jointly found evidence that a goal-setting approach may solve the identified issues while being technically minimal invasive, the ultimately selected intervention is based on goal-setting theory and can be simplified as follows:

> *Instead of being extrinsically dictated by the machine and algorithm, human workers can now choose at each workstation login how many items they want to pick at their workstation out of a set of five goals (pick numbers). In this way, pickers are able to indirectly control how long they want to stay at the respective workstation. The goal-setting intervention includes additional information about the progress of picking towards the selected goal (i.e., goal attainment status).*

Design of the intervention A picker is now required to select one out of five possible goals whenever she logs in (either when switching to another workstation, at the start of the shift, or after a break). This is shown on the login display (see Figure 4.4), next to workstation priority information compared to other workstations. Figure 4.A1 in the Appendix compares the pre-intervention and intervention designs. The chosen goal is shown on the picking display, during normal course of operation. Right below, the goal-setting intervention provides a tracker which counts the picks conducted since choosing this goal. This goal-attainment status is particularly important for workers to track their progress and to trigger goal-setting mechanisms (Locke and Latham, 1990). The five goals (25, 35, 45, 65, 105 items to pick) have been jointly determined with the retailer based on historical number of picks at a workstation and some initial tests. The minimum was determined by the warehouse manager to avoid excessive changes. Based on picker feedback, the largest number should be substantially higher to also allow longer times at workstations.

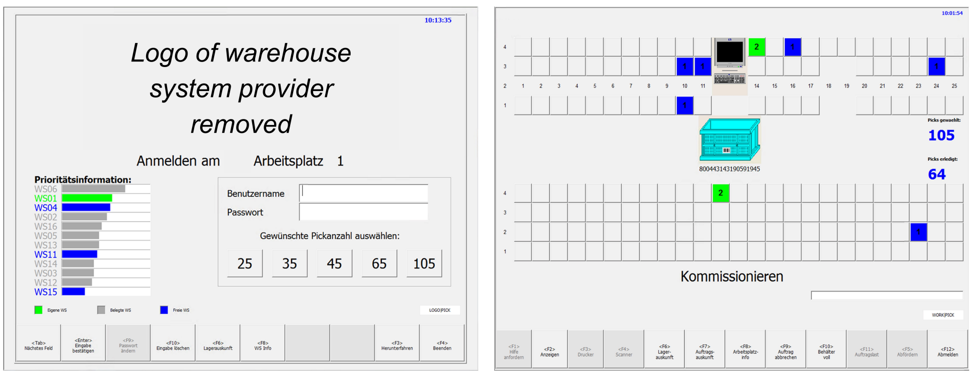


Figure 4.4: Intervention design - login (left) and picking display (right) at the workstations

Implementation of the intervention Figure 4.5 shows the detailed timeline of the implementation. We first explain the final preparation of the intervention and then turn to data collection.

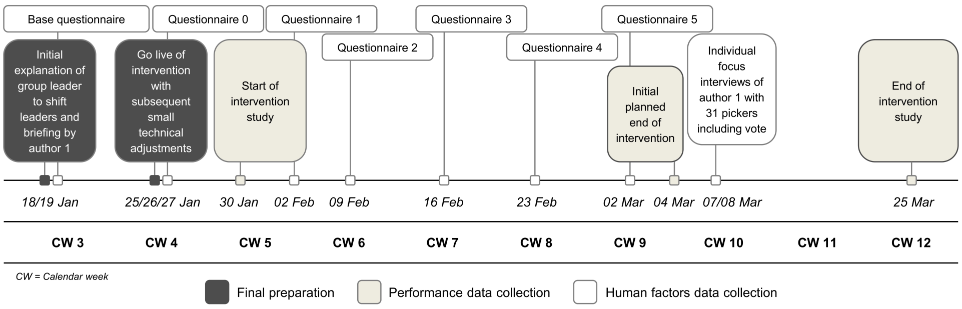


Figure 4.5: Implementation timeline of intervention and data collection

Final preparation. To enable a smooth transition from the old (pre-intervention) to the new (intervention) system, the department head briefed all shift leaders regarding timeline and functionality of the intervention, while one author additionally conducted nine meetings (4-5 pickers each for all three shifts) to emphasize the changes. Overall, 41 pickers were active during the intervention. The intervention was uploaded on the system prior to the start week and finalized with minor corrections of technical bugs and small adjustments.

Performance data collection. The intervention was planned to run for five weeks, but was extended for another three weeks, totaling in eight weeks of intervention time. We automatically collected hourly data retrieved from the warehouse system. These included time stamps and performance values such as (i) workstation ID, (ii) worker ID, (iii) conducted picks, as well as (iv) time spent and (v) pick time at the respective workstation. For the purpose of our study, we regard the conducted picks per person pick hour (i.e., how many items were picked when picking was technically possible) as the main performance indicator. This allows us to exclude any external influences of the technical system and measure only the human action. This also strengthens the accuracy of outcomes when comparing pre-intervention and intervention metrics. These data points are available for 24 months prior and during the intervention.

Human factors data collection. To measure the human factors under investigation (satisfaction, perceived fairness and self-determination), we conducted weekly surveys. We started two weeks before the intervention with a base questionnaire that also included questions to identify personality traits such as regulatory focus (Higgins, 1997) and big five inventory (Digman, 1990), retrieve demographic information (sex, age, height and picking experience), and explore workstation preferences to validate our observations from picker discussions. We followed Larson (2019) to reduce social desirability bias in the questionnaire. For example, pickers were briefed that their employer has no access to individual answers, and that no individual evaluations of their answers will be performed. Also, before filling out the questionnaires, pickers received instructions that answers are not differentiated between correct and wrong, and that they should provide honest self-evaluations. From calendar weeks 3 to 9, we surveyed the pickers each week on the same day. Participants had to choose on a 5-point-likert scale (no, rather no, neither/nor, rather yes, yes) if they were (i) generally satisfied with their work, (ii) working in a self-determined way, which means making decisions, and (iii) believing that the assignment of pickers to the workstations is fair. We additionally surveyed a control group (n=14) within the same warehouse that worked in a different activity and that did not experience any change in their daily work during that time. As the warehouse manager promised his employees to make a democratic vote if the intervention will be implemented after the initially planned five weeks, we hosted individual focus interviews with 31 pickers from the night, early and late shift. In this way, we were able to retrieve individual thoughts and feedback with the absence of group dynamics that the research team could observe during the intervention study. After discussions on (dis-)advantages, pickers had to answer three simple questions. We asked them if they were (i) more satisfied, (ii) more self-determined, and (iii) perceiving a fairer workstation assignment with the old or new system. Additionally, they had to vote with which system they want to continue working: the new system, the new system with adjustments (which they had to define by themselves), or the old system. While the questions about the new system were matched to worker IDs, the vote was confidential. The outcome of the vote was communicated directly the day after by the department head and warehouse manager.

4.3.3 Triggering goal-setting mechanisms (M)

As goal-setting is a well-researched field in other contexts, we can draw on a large body of knowledge to formulate the expected triggered mechanisms. Additionally, the intervention might lead to a change in the average time at a workstation, that is, the time a worker is assigned for picking to a particular workstation. We thus also discuss this potential mechanism.

Goal-setting mechanisms Locke and Latham (2002) deconstruct mechanisms that play a significant role when setting goals. Out of those, three are relevant for our intervention. To start with, goals increase effort and reduce behavior that is irrelevant towards achieving the goal. Second, they energize people, also for physical and repetitive tasks. Third, goals boost persistence, which, for example, enables workers to increase work pace when confronted with tight deadlines. If we apply these three mechanisms to our goal-setting intervention, we expect pickers to increase their effort during the picking process to reach the selected goal, while they may also cut down on behavior that is irrelevant for this (such as chatting with co-workers or logging out in the middle of the picking process). The mechanisms are likely to trigger more energy and persistence, meaning pickers try physically harder to reach the goal as soon as possible. Overall, this can also be denoted as higher engagement as a broader term of these mechanisms.

Reduced time at workstation Additionally, while we did not change any operational processes, we acknowledge a possible indirect mechanism of the intervention which is a change in the average time a picker spends at a workstation. Before the intervention, a change of workstation was only triggered if no more picks were available, or if another workstation was assigned a higher priority by the picker guide. By introducing a maximum goal of 105 as well as different goals to select from, we expect to decrease the average number of picks at a workstation. This leads to an increase of the number of changes per worker per shift and a decrease of the average time at a workstation.

4.3.4 Expecting improved outcomes (O)

We derive the expected outcomes and formulate hypotheses based on the mechanisms and specific setup. We discuss expected outcomes on system performance and the three human factors.

System performance Introducing goal-setting options have successfully increased performance across contexts, subject groups, geographies, or goal sources (Corgnet et al., 2015; Goerg and Kube, 2012; Locke and Latham, 2002; Schultz et al., 2010; Van Lent and Souverijn, 2020). For the latter, research differentiates among external (set by an outsider), internal (self-set), and participative (cooperatively set) goals (Locke and Latham, 2019; Van Lent and Souverijn, 2020). Our intervention comes closest to participative goals: While the system sets the five different choices of numbers of picks, each worker has the freedom to choose any of them without further restrictions. Related research provides clear evidence for expected performance improvement. For example, Latham et al. (1978) and Downing and Geller (2012) show that the introduction of participative goals results in higher performance. While some studies combine monetary incentives with goal-setting (e.g., Doerr et al. (1996) or Friebel et al. (2017)), performance enhancements were also achieved without any financial benefit for the subjects (e.g., Goerg and Kube (2012); Van Lent and Souverijn (2020)). To improve the chances of increasing performance, we apply suitable design elements mentioned in related literature. To name a few, we propose very specific goals (Locke and Latham, 1990), make sure pickers have the skills to reach the goal (Latham and Locke, 2006), and provide goal-attainment status by showing the tracker on the picking display (Jung et al., 2010).

Furthermore, as we outlined before, the goal-setting intervention might trigger a larger number of workstation changes. Interestingly, two important findings in operations management are relevant when estimating the influence on the picking performance. On the one hand, workers may experience a loss of rhythm, which may result in deteriorating performance due to a break in the rhythm of the operations process (see e.g., Schultz et al. (2003) and Staats and Gino (2012)). On the other hand, the more frequent changes may reduce the monotony and boredom of the picking process. This can be seen as an accelerated job rotation for the human picker which may increase the picking performance (Grosse et al., 2015). With a look on the historical data, we identified a negative relationship between the average time spent at a workstation and the picking performance. Hence, we ultimately expect that more frequent changes triggered by our intervention will increase picking performance, supporting the argument based on goal-setting theory. This leads to our first hypothesis:

Hypothesis 1*: The goal-setting intervention increases the picker performance measured in picks per hour.*

Human factors We further hypothesize that our intervention will be beneficial to all three identified human factors. First, goals have been successfully established to increase job satisfaction (e.g., see Doerr et al. (1996) for a production line setting). Specifically, Zhang (2008) defines goal-setting and information on goal-attainment as a source for satisfaction within human-machine interactions. Also, pickers will not experience negative effects of goal-setting approaches (such as dissatisfaction when goals are not reached) as they either pick until the selected goal is achieved or until all available items in the workstation are picked. Hence, pickers are expected to generate a feeling of goal success (Locke and Latham, 1990), which serves as a strong foundation that the intervention increases satisfaction, resulting in:

Hypothesis 2a*: The goal-setting intervention increases picker job satisfaction.*

Further, we expect a positive effect on self-determination given the introduced ability to choose five different pick quantities at each workstation login. Compared to the status quo, in which human employees are purely steered by the assignment algorithm and hence, do not have any possibility of process co-determination, they are now faced with a decision: *Do I want to pick 25, 35, 45, 65 or 105 at this workstation?* This increased freedom of choice enables workers to design an important element of their work, which further boosts the higher engagement into the task itself (see Deci et al. (2017)). Hence, we derive:

Hypothesis 2b*: The goal-setting intervention increases picker self-determination.*

Finally, the intervention might also affect the perceived (un)fairness. Now, pickers are able to influence the distribution of workstations by deciding on the picking goal (and consequently on the time spent) at the respective workstation. For example, if a picker thinks that she is currently being sent too much to a workstation that she does not prefer, she may select small goals (e.g., 25 or 35) to reduce her time at that workstation. Thus, she is able to change the assignment of workstations in a way she perceives to be fairer. As individuals often evaluate the fairness of an event by taking into account the assignment process (Cropanzano et al., 2008), the intervention should increase the perceived fairness of pickers, leading to:

Hypothesis 2c*: The goal-setting intervention increases the perceived fairness of workstation distribution for pickers.*

4.4 Outcomes of the intervention

After having established the triggered mechanisms and expected outcomes, we turn now to the observed intervention results and highlight first the impact on performance, and then on the human factors.

4.4.1 Impact on system performance

A two-sample t-test is applied to investigate the effect on human performance (i.e., picks per person pick hour). The application of several linear regression models including different control variables allow obtaining robust results. Additionally, we conduct a Difference-in-Differences (DiD) analysis including data from a comparable warehouse.

Change of performance during intervention period We compare the eight weeks of the intervention in 2022 with the respective eight weeks in 2021 (calendar weeks 5-12) to ensure a like-for-like comparison (see similar approaches in Kaipia et al. (2017)). Comparing the same weeks avoids any demand distortions due to seasonality, public holidays, and other factors that are especially important in grocery retail. Focusing on week level also ensures all working days and shifts are covered within one aggregation point. Within this time, no technical or organizational changes have been conducted that may have resulted in performance and workload changes. The workload in the respective weeks of both years was comparable since the demand for the grocery retail sector in general, and for our warehouse under investigation in particular, stayed similar. There was no demand shift between these two years caused by the Covid-19 pandemic or other external factors. Furthermore, the order structure and the workload across different workstations did not deviate between the two periods under investigation.

Figure 4.6 plots the performance from the beginning of the year (starting with week 2 to avoid distortions from the holiday season) until the end of the intervention in week 12. The pick performance during the intervention exceeds clearly the pre-intervention performance. A two-sample t-test (weeks 5-12) shows also a significant improvement after introducing the intervention. The difference between the means before and during the intervention is 24.70 (465.58-440.89) picks per person pick hour, and these means are statistically different from each other ($p<0.01$). This results in a performance improvement of 5.60%, which is also indicated by higher picking performance after introducing the intervention when conducting a Mann-Whitney test ($p<0.01$). These findings also hold if we use other temporal aggregation of our data (e.g., day or shift level). It is astonishing that the intervention achieved such a high and significant performance improvement in a system that suffered under stagnating performance over years.

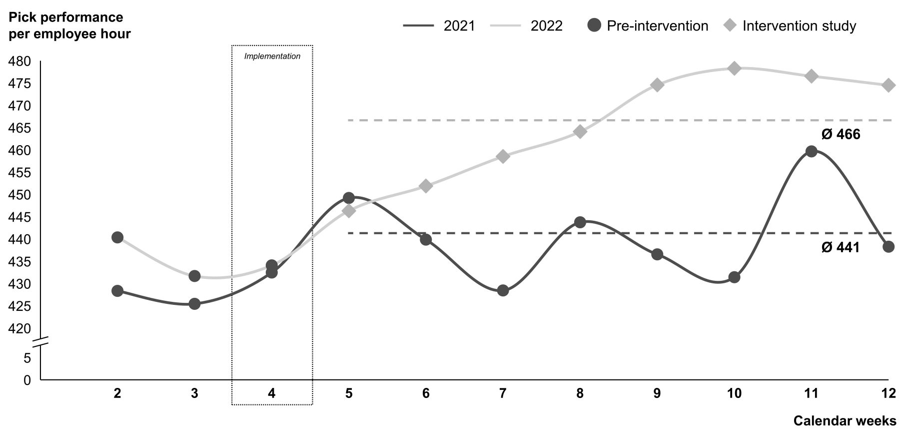


Figure 4.6: Performance comparison on week level between pre-intervention and intervention study

Robustness checks: Regression models on hourly data Because the performance of human pickers might be influenced by several variables, a regression analysis is applied for robustness checks (see also Donohue et al. (2018)). We estimate the effect of the intervention on the picking performance first with a linear OLS regression and add the following five control variables successively:

(i) *Load factor.* The actual load of the system may play a crucial role on the performance of groups and individuals (Delasay et al., 2019). In our application, the workload is determined by how many picks are required to fulfill all orders in the system. The received orders are known for each day, and hence, we can assign each data point the load factor of that respective day. The average load is around 53.000 picks per day with a standard deviation of 7.340 picks.

(ii) *Time at workstation.* The average time spent at a workstation changes during the intervention as pickers select their maximum number of picks at each workstation before they move on to the next station. We observed a reduction of the average time at a workstation by 12% from 14.39 minutes (pre-intervention) to 12.62 minutes (during intervention). This results in approximately 4 more workstation changes per picker during a shift.

(iii) *Workstation ID.* Different products are picked at twelve workstations. This may result in different difficulty levels to pick the respective products, resulting in distinctive human performances (and preferences) for each station. For this reason, the variable workstation ID is included as a fixed effect to remove any potential influences due to these different characteristics.

(iv) *Worker ID.* There is a substantial variation among workers in terms of performance and behavior. Also, given the high employee turnover, only 29 out of 42 employees from 2021 were active during the intervention period in 2022. To control for individual worker effects, we include the worker ID as a fixed effect in our model.

(v) *Calendar week.* To handle seasonality effects, we include the respective calendar week of the data point as a fixed effect.

The pick performance for each pick hour is the dependent variable to capture the effects of the time spent at workstation, worker ID and workstation ID.

To gain a more robust indication, we consider two different time horizons. First, and similar to the initial t-test, we regard data from the intervention study and the respective same eight weeks from 2021 in regression set (1). For regression set (2), we include data from calendar week 13 in 2020 until the end of the intervention.

Regression set (1) We compare again the eight weeks during the intervention with the same calendar weeks in the previous year. This results in 23,381 data points for 2021 and 27,618 for 2022. Note that the larger number of observations for 2022 results mainly from the increased number of workstation changes. Table 4.3 provides an overview of the six different models. The main independent variable (i.e., the intervention dummy variable) as well as the F-statistic show significant values ($p<0.01$) for all models. Model 5 results show that our intervention improves performance by almost 22 additional picks per person pick hour. This is in line with the observed improvement in Figure 4.6. Running the linear regression set (1) replicates the positive impact of our intervention on picking performance.

While load does not seem to have a significant influence on the picking performance, the time at workstation is significant across all three models in which we included the variable. The negative sign confirms our expectation that shorter times at a workstation and consequently a higher number of workstation changes improve the picking performance across pickers. While this effect is significant ($p<0.01$), the magnitude on the overall picking performance is rather small. In Model 5, for example, increasing the time at workstation by 1 hour reduces the expected picking performance by 92.55 picks per hour for pickers. Translating this to the observed reduction of the average time at workstation of 1.77 minutes (difference between 14.39 and 12.62, see above), the time at workstation effect is expected to improve picking performance by approximately 3 picks per person hour. That means one significant, but small part of the improvement can be explained by the more frequent changes of workstations.

Table 4.3: 8 weeks in 2022 compared with 8 weeks in 2021 – Dependent variable: Pick performance

	Base	Model 1	Model 2	Model 3	Model 4	Model 5
Intervention	18.73***	18.87***	16.34***	12.32***	22.01***	21.96***
	(1.94)	(1.95)	(1.95)	(1.82)	(2.20)	(2.20)
Load		0.00	0.00	0.00	0.00	-0.00
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Time at workstation			-89.21***	-112.80***	-92.47***	-92.55***
			(5.32)	(5.16)	(5.02)	(5.02)
Workstation ID				✓	✓	✓
Worker ID					✓	✓
Calendar week						✓
Constant	467.66***	458.97***	479.38***	615.08***	630.73***	624.70***
	(1.42)	(8.48)	(8.55)	(8.63)	(9.87)	(10.13)
Observations	50,999	50,465	50,465	50,465	50,465	50,465
R^2	0.0018	0.0019	0.0074	0.1384	0.1930	0.1938
Adjusted R^2	0.0018	0.0018	0.0073	0.1381	0.1918	0.1925
F-statistic	93.15***	47.37***	125.48***	578.65***	158.54***	145.90***

Note: Standard errors (SE) are in parentheses; *** p < 0.01

Regression set (2) In order to capture a larger time horizon of two full years, we extend the data set and include hourly data from calendar week 13 in 2020 to calendar week 12 in 2022, with the last eight weeks being our intervention period. This gives 304,039 pre-intervention data points. Table 4.4 reports on the results.

The intervention variable and the F-statistic are significant ($p< 0.01$) across all six models, and after introducing the fixed effects, the model fit improves (higher R^2 value). Regression set (2) also confirms that the intervention improves picking performance in a similar amount. The load factor shows significant values ($p<0.01$) for Model 3 and 4, but not for Model 5, which shows the highest model fit. Similar to regression set (1), time at workstation is significant ($p<0.01$) in all three involved models and the R^2 value is enhanced when introducing fixed effects of workstation ID and worker ID, with only little improvement when adding the fixed effect of the calendar weeks.

Table 4.4: 8 weeks in 2022 compared with 97 prior weeks – Dependent variable: Pick performance

	Base	Model 1	Model 2	Model 3	Model 4	Model 5
Intervention	27.25***	27.42***	25.61***	20.32***	21.69***	20.78***
	(1.42)	(1.43)	(1.43)	(1.30)	(1.38)	(1.88)
Load		0.00	0.00	0.00***	0.00***	0.00
		(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Time at workstation			-84.77***	-97.91***	-82.81***	-83.26***
			(2.00)	(1.89)	(1.85)	(1.84)
Workstation ID				✓	✓	✓
Worker ID					✓	✓
Calendar week						✓
Constant	459.14***	458.52***	476.32***	595.63***	602.81***	602.10***
	(0.41)	(2.33)	(2.36)	(2.54)	(3.56)	(4.59)
Observations	331,657	330,718	330,718	330,718	330,718	330,718
R^2	0.0011	0.0011	0.0065	0.1384	0.2209	0.2245
Adjusted R^2	0.0011	0.0011	0.0065	0.1381	0.2206	0.2241
F-statistic	369.58***	4183.58***	724.63***	5128.59***	699.50***	514.57***

Standard errors (SE) are in parentheses; ***$p < 0.01$.

Difference-in-Differences analysis with control warehouse To substantiate the robustness of the performance improvement based on our intervention, we compare the performance development to a control warehouse using a DiD approach (see e.g., Anand et al. (2021); Chun et al. (2022); Patel et al. (2021) for similar analyses in IBR). The warehouse system provider was able to share data of the eight weeks in 2021 and 2022 from a similar warehouse that is comparable due to the identical picking automation, workstation setup and operating principles and hence, the identical human-machine interaction. The warehouse is operated by the same retailer with a similar size and product portfolio, the same function (i.e., central warehouse), as well as shift and data system. We run the analysis using weekly data points. Figure 4.7 plots the results of the DiD regression within a linear-trends model to have equalized values for both warehouses in calendar week 5 in 2021 for better comparability. Both warehouses have a very similar development over the eight weeks in 2021. However, the introduction of the intervention leads to a significant increase in performance, which cannot be identified in the control warehouse. In fact, the regression output shows an average increase during the eight weeks of 36.07 picks per hour compared to the control warehouse ($p<0.01$).

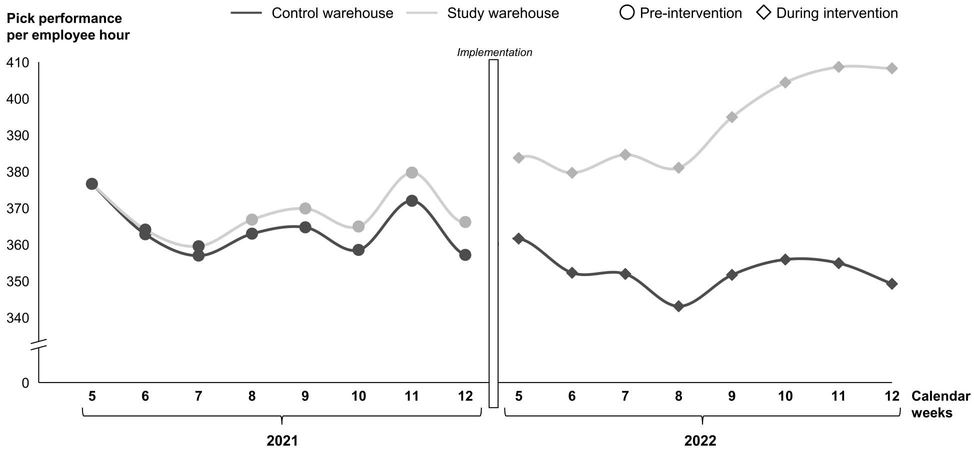


Figure 4.7: Linear-trends model of DiD approach with a similar control warehouse

We have gathered evidence across a t-test, a Mann-Whitney test, two sets of linear regressions with six models each, and a DiD approach that our intervention indeed significantly improves to a large extend the picking performance. Hence, we find a clear support for Hypothesis 1. Picker statements from the focus discussion mentioned several aspects that are related. For example, many appreciated the goal attainment tracker to understand when the next workstation change is coming. Moreover, the number of pickers that preferred the higher number of workstation changes exceeds the ones that did not like the additional walking distance. We also confirmed that the number of workstation changes increased. While the reduction of the time spent at a workstation has a significant impact on performance, the magnitude was low.

4.4.2 Impact on human factors

This section analyzes the impact on satisfaction, self-determination, and perceived fairness. We structure this section along the three human factors, but start with a description on how we obtained the insights. In fact, we retrieved observations across two types of surveys (see Section 4.3.2). First, pickers filled out a questionnaire each week (starting one week prior to the implementation). We compare the weekly survey scores from calendar week 3 (pre-intervention) with the the average of calendar weeks 5-9 (intervention). We discard the score of calendar week 4 as the technical implementation including minor adjustments took place during this period. Second, we hosted 31 individual focus interviews with a subsequent survey (denoted as final consideration) on whether the three human factors have improved due to the intervention study. Figure 4.8 provides an overview of the survey results. We also surveyed the control group in the same warehouse. They showed neutral or positive developments over the course of time and thus, does not contradict any of the mentioned implications.

Surprisingly, there is an overall deterioration of the human factors. To explore this unexpected result, we apply an abductive logic (Chandrasekaran et al., 2020; Sætre and Van den Ven, 2021; Van den Ven, 2007). Particularly, we follow Olivia (2019) and explain the underlying reasons how we ended up in a different situation as expected by describing and interpreting how pickers participated and reacted individually, and within the team to the intervention. The impact on the team interaction requires also to analyze the role of the shift leaders as well as formal and informal arrangements within the team. In the following, we first report on the primary data collected across the two surveys. Then, we provide insights on how and why those outcomes emerged. Specifically, we triangulate the primary data with insights obtained in the focus interviews as well as from in-depth discussions with the department head, warehouse manager, and warehouse provider (we refer to this below as triangulation of sources). Deliberating with the different stakeholders about the unexpected results gave us the opportunity to obtain additional insights in working behavior and decision discretion that was not obvious before (i.e., observations were previously unknown to or have been underestimated by the warehouse manager and operator). To better understand unexpected pickers reactions, we further investigated the boundaries and degree of freedoms in the workflow during and before the intervention by an in-depth process analysis.

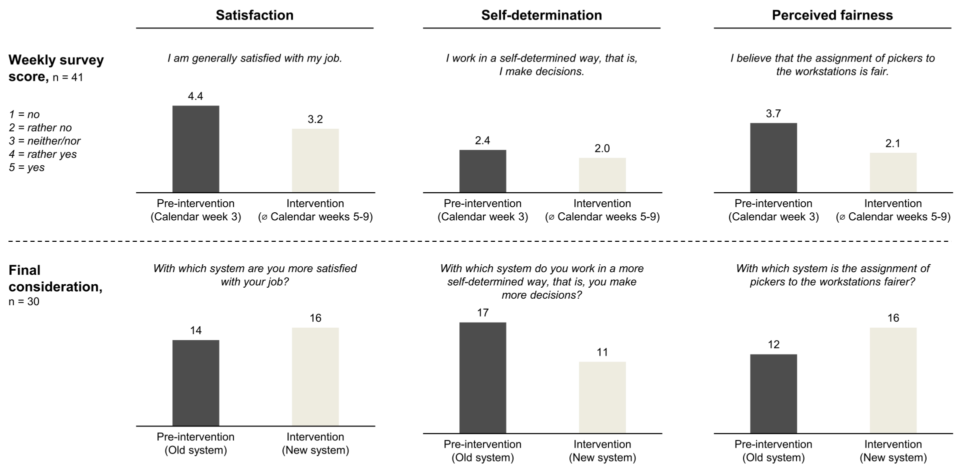


Figure 4.8: Outcome across two surveys for human factors

Satisfaction Figure 4.8 shows a decreasing satisfaction of pickers from pre-intervention to intervention. This contradicts our hypothesis, as we expected satisfaction to be improved by the opportunity of setting own goals. We can establish several reasons that we identified during and after the intervention by triangulating our sources. To start with, the goal-setting intervention affected informal arrangements among the team members, which was the main contributor to the general satisfaction decrease. In fact, the intervention revealed underestimated workarounds that the pickers used to overrule the automated picker guide algorithm in many instances in the pre-intervention phase. The extend of informal arrangements became only transparent through the intervention. The manipulation of the picking sequence (i.e., the pickers logged out manually and changed the workstation instead of following the proposed sequence and assignment) was organized by the pickers themselves and their shift leaders along three possible ways. The first informal arrangement was to daily alternate between the hard and easy aisle. The aisle alternation was ensured by overruling the proposed workstation sequence of the algorithm by the shift leaders and the pickers. For example, if one picker was sent to the hard aisle while being assigned to the easy aisle for the respective day, she manually logged herself out after a couple of picks in the hard aisle and went back to the easy aisle. Second, pickers changed bilaterally when one picker perceived that a colleague spent too much time on the most preferred workstations (i.e., 1 or 2). This was executed by manually logging out of the system and switching the workstation. Third, shift leaders tried to ensure that the slow-mover workstations 6 and 12 were occupied during the entire working hours in order to avoid a sudden large number of open orders for these workstations. For this purpose, each picker had to spend approximately 1.5 hours per shift at such a workstation, and the shift leader managed manually the changes and overruled the picker guide. All three options were taken to bypass the proposed workstation sequence of the picker guide algorithm. However, the overruling of the proposed workstation sequence was suspended with the intervention as pickers now changed workstation only when prompted by the machine, either due to achieving the picking goal or in case of an empty workstation. Despite the aim of the goal-setting intervention to enlarge the participation and decision power of pickers, they themselves perceived the intervention as an suspension of self-invented options to overrule the system and existing informal agreements. The perceived suspension of own-created options can be manifested as the main reason for the general dissatisfaction of pickers. Further explanations that were identified through process insights are aggregated to general resistance against process changes, doubts because of initial technical issues, and lower perceived picking performance. In terms of process changes, pickers frequently mentioned that they did not appreciate the increased number of workstation changes given the extra walking time and distance as well as a lower perceived picking performance. Further, some pickers also reported that given the change of the occupancy at workstations 6 and 12 (recall that pre-intervention, always one picker was assigned there all time), the products at those workstations were picked with delay. However, no evidence for any increase in delays is given. In terms of technical issues, some blockages of transport boxes as well as delays when logging in at a workstation occurred in the beginning of the intervention. The research team fixed both issues within the first days.

However, the lower left part of Figure 4.8 shows a slightly different view than the strong deterioration on the upper left part. In fact, 16 out of 30 pickers (53%) ticked off that their job satisfaction is higher with the new system (i.e., during the intervention) in their final consideration. To understand this difference, we turn to a shift level analysis of the satisfaction score, where each of three shift types (A, B, C) inhibit different group dynamics. Figure 4.9 highlights a clear distinction of shift A compared to shift B and C. While for A, average satisfaction scores are even slightly higher than pre-intervention, for B and C these values decrease. The lower part of Figure 4.9 shows a similar picture. For shift A, most of the pickers (82%) believe that the intervention improved satisfaction, while this value is only 40% and 33% for shift B and C, respectively. Hence, we can determine that the intervention had different implications on the general satisfaction depending on the respective shift. To understand this observation, we further draw on derived in-depth process insights. Specifically via the focus interviews with individual pickers, we found out that shift leaders and other opinion leaders from shift B and C emphasized their negative view on the intervention. These group dynamics led to very similar, negative results of the questionnaires in many instances. In shift A on the other hand, where some disagreements between the pickers and the shift leader existed, higher satisfaction scores were often chosen, potentially as a defiance response towards their shift leader.

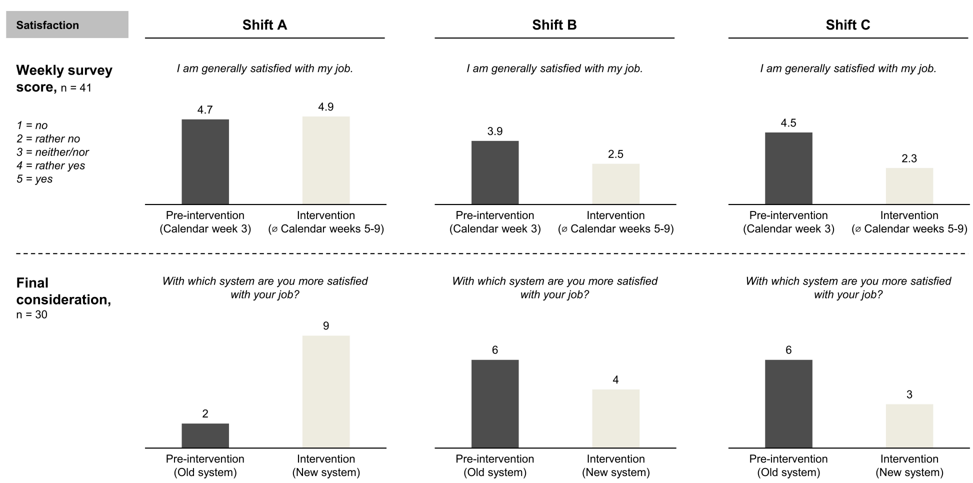


Figure 4.9: Shift level analysis for satisfaction across two surveys

Summing up, we can establish that the intervention led to an overall decrease of satisfaction. The main reason for this negative development was the suspension of options to overrule the system. The low satisfaction scores were particularly driven by pickers in shift B and C, potentially resulting from unfavored group dynamics. Further reasons are disliked process changes, technical issues at the beginning, and perceived lower performance.

Self-determination Figure 4.8 shows that the pickers perceive self-determination slightly lower during the intervention, despite it was particularly expected to increase this human factor, due to the possibility for pickers to self-determine goals. The above-mentioned suspension of options to overrule the system also led pickers to experience a lower decision discretion. More specifically, pickers mentioned that the adherence to the set goals did not allow them to make workstation changes on their own. Before the intervention, this was possible in cases of bilateral agreements when someone spend much time at favorable workstations, or when pickers were sent to a workstation on the opposite aisle. Consistently, only about 40% of pickers voted in their final consideration that they work in a more self-determined way during the intervention compared to the old system. This was confirmed by the focus interviews and the data analysis of the selected goals, both revealing a high tendency for selecting the highest goal and social peer pressure (not to choose low goals at unfavored workstations). This keeps the positive effect on perceived self-determination rather low. Also the department head and the warehouse manager confirmed this tendency, who enforced that pickers often chose the largest goal, and expected colleagues to do the same. Contrary to the satisfaction observations, a significant difference among the shifts did not exist. Hence, although pickers received decision power through the participative goal-setting intervention, in-depth process insights obtained afterwards show that the suspension of managing workstation changes on their own diminished the positive effect stemming from goal-setting theory. In the presence of unfavored group dynamics and consistently selected high goals across participants, the goal-setting approach may go along with a feeling of social pressure which ultimately resulted even in a decrease of perceived self-determination.

Perceived fairness Also the perceived fairness has deteriorated as a result of the intervention (see upper right part of Figure 4.8). This is again contrary to our expected outcome as we initially hypothesized that the intervention would increase the perceived fairness of the workstation assignment by providing options to influence those assignments. 17 pickers directly mentioned in the focus interviews that they experienced a higher unfairness due to the intervention. Specifically, many of those felt that colleagues tend to choose smaller goals at the unfavored workstations, which reduced the perceived fairness of the new system for them. Interestingly, the actual share of small goals across workstations identified in the data was overall rather low, showing that pickers had an overly pessimistic assumptions of colleagues' choice of goals. In other words, there seemed to be some kind of mistrust among the group of pickers as many expected their colleagues to choose small goals at unfavored workstations such as 6 or 12. Moreover, the above-mentioned suspension of informal arrangements also influenced the perceived fairness. Before the intervention, pickers often informally organized themselves to avoid spending too much time at unfavored workstations. In this way, pickers and shift leaders overruled sometimes the picking sequence with aisle changes (when pickers were assigned to another aisle), bilateral changes (when pickers felt that colleagues spent too much time at favored workstations), and the self-adjusted fixed assignment to unfavored workstations for a given time. These agreements led to a supposedly fairer workstation distribution for the individual pickers and for the team as whole. This means that pickers preferred the re-assignment to be done either by the shift leader or by themselves, instead of following the proposed sequence of the automated picker guide. Now, given the intervention, these self-established, informal re-assignment practices have been not available anymore. Because workstation changes and sequences were now solely determined by the automated picker guide after goal achievement (compared to having the decision discretion to leave the workstation earlier by simply logging out), many pickers felt that they are spending too much time at unfavorable workstations. This reduced the perceived fairness. Interestingly, we could not find evidence in the time spent at workstations that would confirm these assumptions of the pickers. For instance, the time spent in the hard aisle with respect to the easy aisle did not change significantly on average.

Similar to the outcome of satisfaction, the vote in the final consideration shows a different picture than the deterioration of the perceived fairness observed in the weekly survey during the intervention. This can be explained by the clear differences between shifts (see Figure 4.10). While the weekly survey score for shift A did not change much (from 2.8 to 2.5), pickers in shift B and C filled out much lower values during the intervention period (from 4.1 to 2.3 and from 4.0 to 1.6, respectively). This trend can also be seen in the final consideration (lower part of Figure 4.10). While 80% of the pickers in shift A favored the new system, more than half of the workers in shift B and C believe that the assignment of pickers is fairer in the old system, where the shift leader had a greater influence on workstation assignment. Many pickers in shift A mentioned in the focus interviews that they actually perceive a fairer workstation distribution with the new system, specifically at the favored workstations. Additionally, many of the workers stated that the increased number of workstation changes is also beneficial for the overall fairness. However, even some pickers in shift A reported that colleagues often choose small goals at unfavored workstations, potentially explaining the lower weekly scores. The intervention also led to overly pessimistic assumptions of colleagues' actions (e.g., overestimating the share of small goals at unfavored workstations). For shift B and C, more than half of the pickers were focusing on this particular unfairness in their focus interviews, while only seven out of 18 mentioned the positive fairness aspects.

To summarize, pickers from shift A favored the intervention in terms of perceived fairness, while pickers from shift B and C did not. The latter two groups were decisive for the overall deterioration, mainly driven by the perceived unfairness due to the suspension of the informal arrangements and the perceived choice of small goals from colleagues at unfavored workstations.

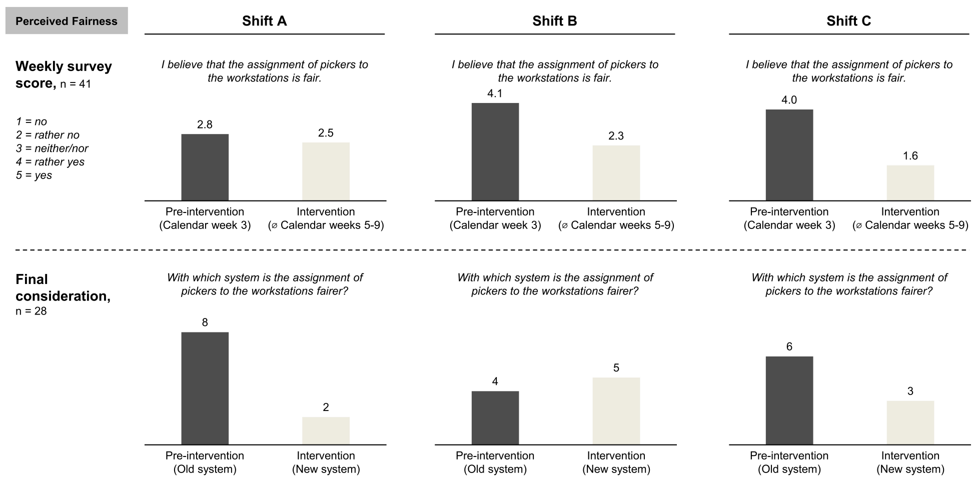


Figure 4.10: Shift level analysis for perceived fairness across two surveys

The surveys show that satisfaction, self-determination, and perceived fairness deteriorated during the intervention compared to pre-intervention. Thus, we cannot confirm Hypotheses 2a-c. The in-depth analyses of observed human behavior, interactions, and processes via further interviews, process analyses, and discussions with workers and leadership enabled us to better understand and explain unexpected findings. We applied abductive reasoning to establish the suspension of manual workarounds as the main reason why the human factors under investigation decreased. Because the intervention diminished possibilities for the pickers to informally organize themselves in their working day, negative repercussions on satisfaction, self-determination, and perceived fairness emerged. Additionally, a shift level analysis showed that shift A actually favored the intervention, resulting in higher satisfaction and perceived fairness scores for those pickers. However, the majority of shift B and C experienced negative developments of these human factors driven by group dynamics among pickers and shift leaders, leading to the overall negative outcome.

4.5 Discussion, implications and way forward

The purpose of this research is to overcome behavioral and system performance challenges in a semi-automated human-machine interaction for an operational, monotonous activity. Our IBR of introducing a participative goal-setting policy results in a performance improvement of human workers. However, these efficiency gains come at the cost of social discord, that is, lower satisfaction, self-determination, and perceived fairness, as prior informal agreements to self-manage working processes are suspended. In this section, we discuss our findings along theoretical and managerial implications. Finally, we sketch out limitations and future areas of research.

4.5.1 Implications for theory

Our study makes multiple empirical contributions to the literature. To start with, we establish a positive impact of goal-setting on worker productivity in highly operational, monotonous tasks within human-machine interactions. Given the stagnating picking performance over years, this constitutes an substantial contribution to the improvement of the operations in an area that is the most cost intensive part in warehousing (De Koster et al., 2007). Our study thus joins the rank of successful demonstrations of the power of goal-setting theory (Locke and Latham, 2019). However, it remains unique in elevating the performance for a repetitive, monotonous, and operational task within a human-machine interaction in the absence of any kind of monetary incentive for blue-collar workers. While goal-setting in operations management has often been applied for many tasks involving white-collar employees with a higher degree of task complexity or variety (see Latham et al. (1978); Friebel et al. (2022); Linderman et al. (2006)), the application for blue-collar workers is rather scarce, and novel for human-machine interactions. Our research sheds light into the optimization of such kind of activities in operations management when other well-known mechanisms may be not applicable (such as managing task complexity or task variety, see KC (2020) for other related mechanisms). The utilized performance metric (i.e., picks per person hour) allows us to center our understanding of the performance improvement on the behavior of the human worker. We can derive that despite receiving no higher financial compensation nor any type of reward, human workers are getting incentivized to speed up compared to pre-intervention (i.e., having no goal). It is demonstrated that the goal-setting mechanisms triggered a higher achievement motivation, leading to superior performance of human workers (see Locke and Latham (2019) for similar results). Figure 4.6 shows that the effect did not diminish over time and thus, we may assume that the effect based on goal-setting could be sustainable.

Furthermore, we reveal the existence, and demonstrate the importance of informal arrangements among workers which generates important theoretical insights into human behavior and its implications on worker productivity. On the one hand, we can conclude that these informal agreements are essential for human workers with respect to satisfaction, self-determination, and perceived fairness. Hence, the suspension of these informal agreements leads to social discord for individuals and groups. Under the assumption that an organization wants to avoid this, the suspension of informal agreements is not advisable, and these may be incorporated into the operational policies when goals are implemented. Hence, not respecting informal agreements and team structures may be an additional pitfall of the goal-setting theory (Latham and Locke, 2006). Building up on this, we find evidence that goal-setting approaches may trigger additional pressure to behave according to the group's expectations. For example, the many comments regarding the choice of goals at unfavored workstations show that workers expected colleagues to make decisions that are conform with the groups' wishes. On the other hand, a better performance does not necessarily go along with a higher satisfaction or enjoyment. As goal-setting theory is based on achievement motivation, and not intrinsic motivation (Locke and Latham, 2019), workers do not have to love their job while improving their results. Our study goes in the same direction and reports a higher performance paired with lower human factors scores. Hence, one may question if there is any cost of social discord, at all? Due to the IBR approach we learned that human workers engaged prior to the intervention in activities that were not aligned with the firm's goal to pick as fast as possible. For example, the informal arrangements did not have any justification for a higher performance, but rather for individual self-optimization. One could thus argue that lower human factor scores can be accepted, if human behavior is now more aligned with the ultimate firm goal. However, this comes at a big caveat as the true cost of the social discord will be only visible long-term, potentially in even higher employee turnover.

Taking a deeper look into the establishment of these informal arrangements shows that they originally evolved as a reaction to the decisions of a "machine". However, our findings may be extended to settings in which operational decisions are made by humans. In particular, we find evidence that people particularly disliked the loss of being able, or having autonomy, to influence their working day (i.e., in our case the workstation distribution). Similar, the focus interviews revealed that in both periods (pre-intervention and during intervention), people showed signs of distrust towards the decisions of both the "machine" and the shift leaders. Additionally, this is extended to the perception of the goal choices of colleagues (recall the overly pessimistic assumptions of colleagues' choice of goals at unfavored workstations). This has implications for self-determination and perceived fairness of the employees, as it seems that humans tend to deviate from fixed operating structures and perceive them unfair whether these policies were made by humans or machines.

Another contribution comprises the demonstration of the importance to understand human factors and behavior within human-machine interactions. A key issue when humans and machines are collaborating is the design of engaging interactions in order to achieve an efficient performance level (Lorson et al., 2022). Goal-setting theory proved to be a solid approach to produce cognitive achievement levels (Zhang, 2008). While we do not claim that goal-setting mechanisms are the only way to improve human-machine interactions, they certainly showed a significant and positive performance impact on human workers. This can inform a wide range of human-machine interaction studies in warehousing, as the human behavior has so far widely been neglected (e.g., Wang et al. (2021)).

Finally, we contribute to the broader operations management literature of using IBR studies. Implementing an intervention during the normal course of action delivered a variety of theoretical insights that researchers using lab experiments would have missed (such as the differences between shifts based on pre-existing team dynamics). Furthermore, abductive reasoning allowed us to capture the reasons for the contradictions to our Hypotheses 2a-2c formulated in Section 4.3.4, showing the potential of IBR to detect unexpected findings. While we acknowledge the informal arrangements may differ from case to case, their existence whenever humans and machines form teams is likely. This fact and given that order picking is a generic task, make our findings transferable to other fields where the operational productivity of human workers is important. We hope that our study encourages more operations and warehousing scholars to utilize IBR to tackle behavioral issues, crucial to innovate human-machine interactions in the future.

4.5.2 Implications for practice

Our findings have managerial implications for the development, design, and execution of human-machine interactions, including their integration into the operations of a firm. The increasing number of cooperations between human operators and automation establishes human-machine interactions as a centerpiece of operations management (Olsen and Tomlin, 2020). Despite all the technological progress of automated and robotized systems, we show that humans, their actions, and behavior will continue to play a significant role in the operational efficiency of firms. While industry and engineering (such as in our application) mainly dealt with the design and technical functionalities of automated systems, the role of human factors and behavior is still underrepresented in management decisions. For example, stagnating worker performance and high employee turnover are obvious problems nowadays in operations and logistics (McKinsey & Company, 2021b). This may be also traced back to a limited consideration of human factors of blue-collar workers in the past, particularly in highly monotonous interactions with advanced automated and robotized systems. Our intervention demonstrates that considering the implications of human behavior is decisive for the effectiveness of human-machine interactions. It thus becomes indispensable to anticipate human behavior with all its facets, and to put it at the center of companies' operations when striving to increase system performance.

Behavioral implications need to be considered not only in the design of human-machine interactions, but also in the management of related workflows, particularly across operational activities on the shop floor level. For example, we found that there is a significant and negative relationship between the average time spent at a workstation and worker productivity. Given the nature of repetition and monotony in the task at hand, we find evidence that operating policies should be comprehensively defined to reduce boredom, or to increase task variability, of blue-collar workers.

Our intervention further establishes that integrating human factors into such systems does not need to be invasive in current operations, and can be achieved by managers with rather simple methods which proved to be effective in related domains. Practitioners can thus innovate human-machine interactions with an easy to implement participative goal-setting option, which yields in a significant productivity gain for repetitive and monotonous tasks.

Moreover, the deterioration of satisfaction, self-determination, and perceived fairness shows the necessity to acknowledge informal work arrangements, structures, and decision discretion within operational groups. This calls for a deeper understanding of individual workers and groups, team dynamics, and informal arrangements. Managers, system engineers, and planners will need to be mindful about of such tacit knowledge and implicit team rules, and include these in operational policies and design principles of human-machine interactions when suitable. For example, the picker guide algorithm which steers the workstation distribution may incorporate the informal agreements of aisle change by fixing each picker for one day to a specific aisle.

Implications for study warehouse Subsequently to our IBR study, the retailer's warehouse management and the warehouse provider became more concerned that human factors matter in the design and execution of the semi-automated picking. In fact, the warehouse provider understood the value of integrating behavioral aspects in the design of the workflow and algorithm that regulates the assignment and sequence among pickers. The retailer's management further build up on the behavioral learnings and introduced immediately an upper limit on the number of picks at each workstation to enforce faster workstation changes, reduce monotony, and increase fairness. Moreover, the senior management teams of both firms appreciated that minimal invasive interventions can already achieve a significant performance impact (in our case of 5.6% higher picking performance). Combining all these considerations, at our suggestion, both practice partners will try to enhance operating procedures and include human factors in the design and operations of the workflows in the warehouse, particularly for human-machine interactions. This will aim the path for expanding and testing further non-monetary incentives for workers to create a more attractive environment for repetitive and monotonous tasks, and to increase the performance levels like in our intervention.

4.5.3 Limitations and future research

Research in human-machine interactions with monotonous and repetitive tasks is still nascent due to the very recent implementation and growth of automated and robotized solutions (Lorson et al., 2022). Our IBR study builds the starting point for future research at the intersection of operations management and behavioral science in this area. We now discuss limitations and future avenues for research along the context of our study and beyond.

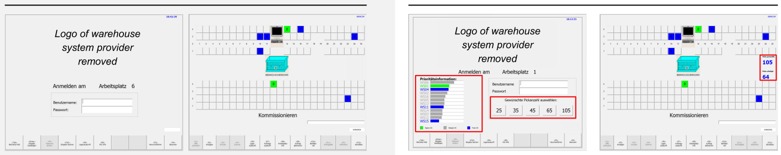
As the goal-setting intervention included five possible goals, we could not analyze if the number of different goals or their respective values may have an influence on the worker productivity. The same holds true for the impact of other design options (e.g., external set goals) or interaction effects between selected goals and worker productivity. Future research can explore such directions. Moreover, we established that human factors mainly deteriorated because informal agreements were suspended. To further disentangle reasons for this finding, future studies can explore if the deterioration is based on the change of the process (i.e., loosing autonomy or decision-discretion, see Dietvorst et al. (2018)) or outcome (i.e., having the subjective feeling of spending more time at unfavored workstations). Furthermore, another limitation is our sample size (n=41 pickers). Given the nature of this field study, we were limited in the maximum size of participants in the respective human-machine interaction. It would be interesting to replicate our findings in a randomized controlled trial setting with several treatment and control entities (i.e., installing the intervention across multiple warehouses with a similar amount of control warehouses). Additionally, one could argue that the performance increase of the human workers is partly based on demand effects. However, we mitigated this potential impact by avoiding at any time during the intervention the communication of our research hypotheses or questions regarding picking performance (see Eckerd et al. (2021)). Our study further focused on the aggregate effect of the goal-setting intervention on performance. Future endeavors can explore worker heterogeneity (by defining subgroups based on skills, personality traits, or other criteria) in terms of performance and how individual pickers reacted differently to the intervention. Finally, our study is limited to the eight weeks intervention period. Further research may investigate the long-term impact on worker productivity and human factors as well as worker retention. Future IBR studies can make use of archival data to analyze such an issue (see, for example, Oliva and Watson (2011)).

There are also further related research opportunities beyond the context of our study. Despite that order picking in a semi-automated warehouse is prototypical for many other human-machine interactions across different operations areas, the transfer of the findings and implications should be tested also, for example, in a manufacturing (e.g., production line with automated robots) or service operations (e.g., call centers with automated task assignment) setup. Furthermore, it would be interesting to see how goal-setting interventions work in human-machine interactions with operational, but less monotonous and repetitive tasks (e.g., in healthcare operations). Finally, we have concentrated our efforts on blue-collar, operational activities and have not extended our perspective on white-collar planning tasks that need to deal with repetitive activities. For instance, in control rooms of operations, a common issue is overwriting optimal parameters for automated systems by human operators. This often happens based on individual human preferences, or unknown information. Consequently, future work could explore issues in these directions and leverage findings obtained in this study.

4.6 Concluding remarks

We performed an IBR study to explore solutions to a practical problem faced by a warehouse operator within a human-machine interaction at order picking. Based on goal-setting theory, we proposed a new operating policy, which enables human pickers to participate in the decision how many items they want to pick at the specific workstations. Our research strategy aimed to improve both human performance (conducted picks per pick hour) and human factors (satisfaction, self-determination, and perceived fairness). While we find support for our hypothesis on increasing human performance with the goal-setting intervention, we did not observe a positive effect on human factors. We triangulated sources and observed underlying processes to detect that the effects on informal arrangements were the main reason for the deterioration of human factors. Reflecting on the design of efficient operational human-machine interactions shows the necessity to incorporate human behavior into decision models. While deviations of human actions from expected outcomes are nothing new (Boudreau et al., 2003), their appearance are still very recent (Roels and Staats, 2021), and crucial for the performance in human-machine interactions (Lorson et al., 2022). We hope to stimulate further studies on human behavior when interacting with novel automated and robotized systems. While we acknowledge that many research methodologies are suitable to uncover behavioral issues in worker productivity (Bendoly et al., 2010), applying an IBR approach for a real-world field study allowed us to both solve a complex practice issue and deploy operations management theory. This way, interesting and surprising findings (such as the effects on informal arrangements) can be revealed to enrich existing theories. Thus, we encourage scholars to adopt IBR in future studies on human-machine interactions in warehousing, and beyond.

Appendix



	Pre-Intervention	Intervention study
Determination of pick number	Automatically determined by the picker guide without any decision discretion by the picker	Pickers choose pick number out of five possible options which is determined as the goal before switching to the next workstation
Trigger of workstation change	No information for the picker when workstation change is triggered Picker guides decides based on algorithm	Workstation change is triggered when picker reaches selected goal In case the picker finished all available picks at the workstation, or no transport boxes are available anymore, the goal is also achieved
Goal attainment	No information on goal attainment status	Goal attainment tracker on picking display that automatically counts the conducted picks in relation to the selected goal

Figure 4.A1: Comparison between pre-intervention and intervention study (display changes highlighted in red)

5 Finding the right one: Decision support for selecting cost-efficient order picking solutions

Co-authors: Fabian Schäfer and Alexander Hübner
In submission process to *IISE Transactions* as of August 09, 2022

Abstract Enabled via recent advances in technology coupled with the advent of new systems providers and decreased price points, automated and robotized order picking solutions (e.g., pick assisting autonomous mobile robots) have evolved as a surging market. Such innovative picking technologies aim at reducing labor costs, using available space more efficiently, and increasing throughput rates. As implementation projects and the variety of solutions rise, managers face the decision of which ones to select for their specific warehouse and products. However, comprehensive decision models for this strategic problem are lacking in pertinent literature. We propose an innovative mathematical optimization model that selects and sizes order picking solutions and assigns them products as well as warehouse spaces. Expert interviews are used to comprehensively identify the decision-relevant costs and constraints. In particular, we minimize setup, module, labor and error costs while adhering to characteristics related to the area (e.g., available space), technology (e.g., throughput, handling capabilities of certain products) and product (e.g., physical dimensions). We conduct a case study and complement our findings with numerical experiments. We find significant cost reduction potential of up to 57% by selecting a mix of different order picking solutions. Further analyses highlight the need to retain human workers and to account for maximum labor capacity.

5.1 Introduction and motivation

Expanding automation and robotization has been the focal point of operations in recent years (IFR, 2020). Enabled via advances in Internet of Things devices and artificial intelligence coupled with the advent of new providers and decreased price points, one surging change in operations evolved to be in the arena of warehousing. In fact, the size of the warehouse automation industry increased 12% annually from 2014 to 2019, and is predicted to reach double its size in 2026 compared to 2019 (Statista, 2022). A search on an independent comparison platform delivers more than 200 results for warehousing robots from more than 80 different providers, showing the wide range of automated and robotized options for warehouse operators (Lots of Bots, 2022). Such solutions are deployed to achieve faster throughput times, reduced costs, higher pick quality, more efficient space utilization, improved ergonomics, and lower dependence on human workers to cope with ongoing labor shortages (Azadeh et al., 2019; Pazour et al., 2014). Many such automation initiatives start at order picking since it is the most labor, and cost-intensive warehouse activity (De Koster et al., 2007; Boysen et al., 2019). Warehouse operators therefore invest heavily in a wide range of semi- and fully-automated order picking solutions (OPSs). Examples include shuttle-based storage and retrieval systems or a variety of pick-assisting autonomous mobile robots. This trend of implementing automated OPSs will doubtless continue in the future. In a survey of Banker (2020), almost 80% of the respondents stated that their organization is *likely* or *very likely* to invest in warehouse automation within the next three years. This requires diligently selecting the most appropriate OPS technology that suits to the warehouse capacity and product portfolio (De Koster et al., 2007; Marchet et al., 2015; Van Gils et al., 2018).

Selecting OPSs on a large scale among hundreds of variants constitutes a novel problem as in the past only manual pickers have been used, or individual (semi-)automated picking technologies have been applied for restricted parts of the assortment. The increasing maturity and the advent of a large variety of advanced, flexible, and cost-efficient OPSs allow their application to the entire warehouse and its assortment, but make this selection problem a much more complex matter. However, no comprehensive decision support exists yet (Azadeh et al., 2019; Van der Gaast and Weidinger, 2022). In fact, OPS selection in practice has up to now largely been based on gut feeling and planners' experiences (Gu et al., 2007a; Pazour et al., 2014; Van der Gaast and Weidinger, 2022). This cannot result in optimal decisions as manifold developments and challenges need to be considered, including:

- *Intensifying cost pressure and labor shortages:* The growth of e-commerce and the demand for faster and cost-efficient deliveries make warehouses the focal point of companies (Hübner et al., 2015, 2019; Boysen et al., 2019, 2021). Warehouse operators typically suffer from cost pressure and labor shortages (Instawork, 2022). This makes it important to find cost minimal solutions and manage the capacity of the human workforce, which continues to be required for activities such as supervision, inventory replenishment or semi-automated picking.
- *Skyrocketing number of novel OPSs:* A growing number of automated and robotized OPSs are becoming available. These OPSs differ in many dimensions such as picker-to-parts or parts-to-picker setups, investment and operating costs, throughput time, technical capabilities or pick quality.
- *Enlarging product diversity and assortments:* Driven by ever growing online sales, wholesalers and retailers are offering larger assortments. The resulting diversity of products makes it challenging to find matching OPSs, as each machine is typically limited in its technical product processing capabilities (in contrast to humans, who have greater flexibility) and throughput specifications. As a result, multiple OPSs, including manual solutions, are often implemented in the same warehouse to cope with the large variety of physical product properties.
- *Increasing importance of space utilization:* The application of multiple OPSs requires dividing the picking zone into separate units as usually only one OPS type can be operated in one area. This means, OPS selection and space assignment need to be performed simultaneously.

Space efficiency is decisive given that OPSs differ in terms of space utilization, and also because available land for warehousing is becoming increasingly scarce and cost intensive (Prologis, 2022).

We introduce and formalize this novel decision problem to incorporate the recent developments and to answer several calls for research (e.g., Boysen et al. (2019); Davarzani and Norrman (2015); Jaghbeer et al. (2020)). Choosing suitable OPSs can be classified as a simultaneous selection (of the appropriate OPSs) and assignment problem (of products and warehouse spaces to the OPSs selected). We derive the conceptual background of the novel problem by conducting expert interviews and review related literature in Section 5.2. This builds the foundation for formalizing a cost minimization model in Section 5.3. We additionally conduct a case study using proprietary cost data from a business partner, evidence substantial savings potential and apply numerical experiments to generate managerial insights in Section 5.4. Section 5.5 concludes the study and provides future areas of research.

5.2 Conceptual background and related literature

The application of different picking technologies is an innovative concept both in practice and in academia (Boysen et al., 2021). Section 5.2.1 derives the context and structure of the novel selection and assignment problem. After clarifying the decision problem at hand, a review of related literature highlights existing research in Section 5.2.2. Finally, Section 5.2.3 connects the practice requirements and literature, resulting in the identification of the research gap.

5.2.1 Description of the novel problem and setting

This section delineates the novel problem based on multiple data sources. First, we draw on essential elements identified in warehousing literature on OPSs (e.g., Azadeh et al. (2019); Boysen et al. (2019, 2021); De Koster et al. (2007); Rouwenhorst et al. (2000)) and transfer relevant insights to our problem context. The goal is to capture generalizable data on the selection problem, decision-relevant costs and managerial-relevant constraints. For this purpose, we additionally interviewed eight experienced warehouse planners (see Table 5.A1 in the Appendix) that frequently work on the OPS selection task, either from a provider or an operator perspective. Further insights from a case study with an industry partner are incorporated into the description of the problem setting.

Overview of the selection and assignment problem

Planning problems in warehousing can be classified into strategic (e.g., material flow design or warehouse management system implementation), tactical (e.g., storage assignment, zoning or order consolidation) and operational decisions (e.g., batching, workforce planning or pick assignment) (Davarzani and Norrman, 2015; Van Gils et al., 2018). Our underlying decision problem belongs to strategic planning as it defines the technological equipment for a long-term investment horizon (Marchet et al., 2015; Vanheusden et al., 2022). In our setting, the warehouse planner needs to perform the **(1) selection and sizing of the OPSs**, **(2) assignment of products to OPSs**, and **(3) assignment of warehouse picking space units to OPSs**. Figure 5.1 illustrates the selection and assignment problem. It shows that not all products can be processed by every OPS (e.g., bulky products can only be processed manually). However, each product needs to be assigned to a possible OPS and each OPS selected needs to be assigned to a space in the warehouse. The following sections delineate the specific associated decisions (1) to (3).

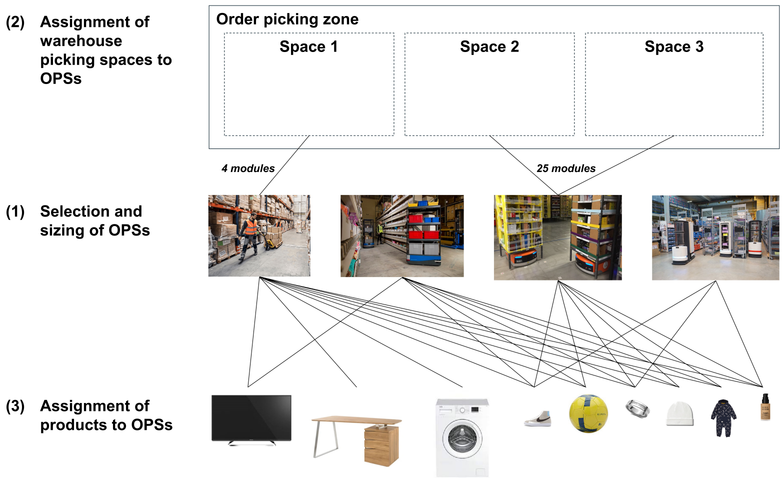


Figure 5.1: Simplified illustration of the OPS selection and assignment problem

(1) Selection and sizing of OPSs OPSs are differentiated between manual, semi-automated and fully automated OPSs. Note that manual picking is also included in our analyses and constitutes a potential OPS. Table 5.1 classifies semi- and fully automated OPSs. We highlight the most common ones and refer for extensive reviews on the particular technicalities to Azadeh et al. (2019), Boysen et al. (2019, 2021) and Fragapane et al. (2021), for example.

Table 5.1: Overview of common semi- and fully automated order picking solution types

Picker/Robot-to-parts	Parts-to-picker/robot
• Fully autonomous picking robots (FAPRs) • Pick-supporting autonomous mobile robots (AMRs) • Pick-supporting automated guided vehicles (AGVs)	• Advanced picking workstations • Shelf-moving robots (SMRs) • Aisle-/grid based shuttle systems • Automated storage & retrieval systems (AS/RS) • Carousels, vertical lifts, dispensers

OPSs are differentiated by the material flow into picker/robot-to-parts and parts-to-picker/robot configurations (Boysen et al., 2021). In the former, the picker (or in the case of full automation, the robot) moves to the storage area to retrieve the products, while in parts-to-picker designs the products are carried to the picker (or the robot) by a transportation system. In fully automated robot-to-parts setups, FAPRs fulfill the process without the help of humans (Fottner et al., 2021). They navigate through the picking area, stop at the respective pick location and retrieve the product. After performing several picks, the robot moves to a repository or consolidation spot. In semi-automated picker-to-parts setups, AMRs or AGVs support the picking process by performing the transportation of goods, while the human removes the product from the shelf and puts it in the bin carried on the robotic device. In simplified terms, AGVs are steered centrally and follow predefined paths, while novel AMRs move flexibly within a given area and decentralize the decision-making processes (Fragapane et al., 2021). Novel OPSs have emerged particularly within this classification (Pasparakis et al., 2021). In automated parts-to-picker/robot solutions, goods are transported to advanced picking workstations where piece picking is performaned manually or in a fully automated manner using a robotic arm (Azadeh et al., 2019; Füßler and Boysen, 2019). Parts are supplied via conveyor systems, SMRs (Lamballais et al., 2020; Wang et al., 2021), aisle-based (Ekren, 2017) or grid-based shuttle systems (Zaerpour et al., 2017), AS/RSs (Roodbergen and Vis, 2009) as well as carousels/vertical lifts (Meller and Klote, 2004) or dispenser solutions (Pazour and Meller, 2011).

Technological processing restrictions hinder any OPS from picking any product, and using certain OPSs for certain products may be more cost-efficient (Boysen et al., 2021). As each OPS has different advantages (e.g., manual picking can handle a large variety of products but usually goes along with higher pick errors, while SMR-assisted picking has one of the highest throughput rates), a mix of OPSs is required. In this regard, Boysen et al. (2021) describe the combination of a parts-to-picker bulk solution for fast-moving products with a picker-to-parts solution for the slow movers, selecting the most suitable OPSs depending on demand. Another example is the fashion retailer Zalando, which implemented FAPRs for shoes (alongside an existing manual picking setup and shuttle system), differentiated by the physical property of the products in this case (TGW, 2016; Magazino, 2021). Usually only one type of OPS technology is applied within one warehouse (e.g., not multiple different FAPRs of different providers), but different OPSs (e.g., FAPRs and AGVs) may be used in parallel in different areas of the warehouse.

When selecting OPSs, each of them can be sized in terms of their picking capacity (e.g., using one or multiple FAPRs). Sizing is done by defining the number of modules per OPS.

(2) Assignment of warehouse picking spaces to OPSs Each OPS needs to be installed at a certain location within the warehouse, that is, specifying the warehouse space units where the OPS operates. The total picking zone in the warehouse is limited and divided into different space units, which may differ in terms of size. A number of adjacent spaces may be combined to one joint area for one OPS. That means an area covers one or more space units, different areas may exist in a warehouse, and only one OPS operates in each area. For example, a leading German grocery retailer has three distinct areas within a central warehouse for three different OPSs, which are all separated from each other. While the AS/RS is located at the heart of the warehouse and sized with several high-bay storage racks, a semi-automated picking system with various workstations is placed next to it, while a manual picker-to-parts setup with multiple aisles is located close to the packing zone to avoid long travel distances.

(3) Assignment of products to OPSs OPS selection requires the assignment of products to OPSs. As denoted above, each product may be picked by different OPSs, but not all products can be technically and cost-efficiently picked by every OPS. This makes it necessary to perform the OPS selection on an individual product level. If the assignment is done without acknowledging the individual physical properties (such as dimensions) of the products, the selected OPS may not be able to handle all the assigned products in the end. For instance, this may be the case for the assignment of voluminous products to AMR-assisted solutions, as robots typically have smaller stockpiling options than manual picking with forklifts. OPS-to-product assignment is therefore always a trade-off between different criteria.

Summarizing problem definition The warehouse stores a set of products with given product characteristics (e.g., product dimensions). To pick these products, there is a set of OPSs that the warehouse planner can choose from, and can scale the size of the OPS by selecting multiple modules. Each OPS has defined characteristics in terms of throughput rates, processing capabilities, pick quality, space requirements and costs. That means the related product characteristics and the technological capabilities of the OPSs determine whether a product can be picked both physically and cost efficiently by a certain OPS. Furthermore, the total warehouse picking space is given in different available space units and each OPS has a certain space requirement. The selected OPSs need to be assigned to the different spaces, while a number of spaces may be grouped to an area if the same OPS is utilized. In summary, warehouse managers need to simultaneously select the optimal mix of OPSs and their modules, assign both the products to OPSs and the OPSs to spaces to find feasible and cost-minimal solutions.

Decision-relevant costs in OPS selection and assignment

Given the novelty of the problem, it is imperative to identify decision-relevant costs. Overall, the OPS expenses are composed of *(i) setup*, *(ii) module*, associated *(iii) labor*, and *(iv) error costs*:

1. *Setup costs* may include one-time expenses for technical implementation, overhead and software as well as services required along the procurement process (Usher et al., 2001). The former need to be incorporated to ensure compatibility with existing systems (such as warehouse management systems). Some integrated OPSs may also require changes of up- and downstream warehouse processes. For instance, when installing

large fixed systems like an AS/RS, the entire storage setup needs to be adapted.

2. *Module costs* refer to any expense that can be attributed to the investment in one unit (i.e., one module) of each OPS. Multiple modules can be selected to define the total size of an OPS. The major share of the costs is the purchase price (or rental price in the event of a leasing contract) for one AMR, for example. Other essential elements are maintenance, energy (particularly for automated and robotized systems) and ancillary costs that occur during operation.

3. *Associated labor costs* are necessary to consider for all OPSs, and depend on the OPS itself. These are particularly workers' wages for the support during picking (e.g., picking with AMRs or from SMRs) and for the additional pre- and post-processing activities such as inventory replenishment or some additional tasks after the picking.

4. *Error costs* occur for picking inaccuracies detected downstream in the supply chain. The OPSs have different error rates. Higher error rates and lower pick quality of an OPS lead to an increase of incorrect orders, and consequently extra handling effort (e.g., external returns processing).

The costs typically occur with different frequencies and time periods (e.g., setup costs are one-time expenses while module costs also have running expenses that occur over a certain time horizon). To account for this, the combination of the costs is widely established in warehousing academia (Rouwenhorst et al., 2000), for example, through the consideration of annual costs (Pazour et al., 2014). The one-time setup costs are transferred to annual costs by dividing them across the expected lifetime. In these cases, an annual present worth economic factor is applied to account for the different time periods. The module costs are available on an annual level (e.g., annual energy costs, annual depreciation). Labor and error costs can be totaled to annual levels as well.

Related constraints in OPS selection and assignment

To ensure feasible solutions, constraints are identified along three domains (see Figure 5.2).

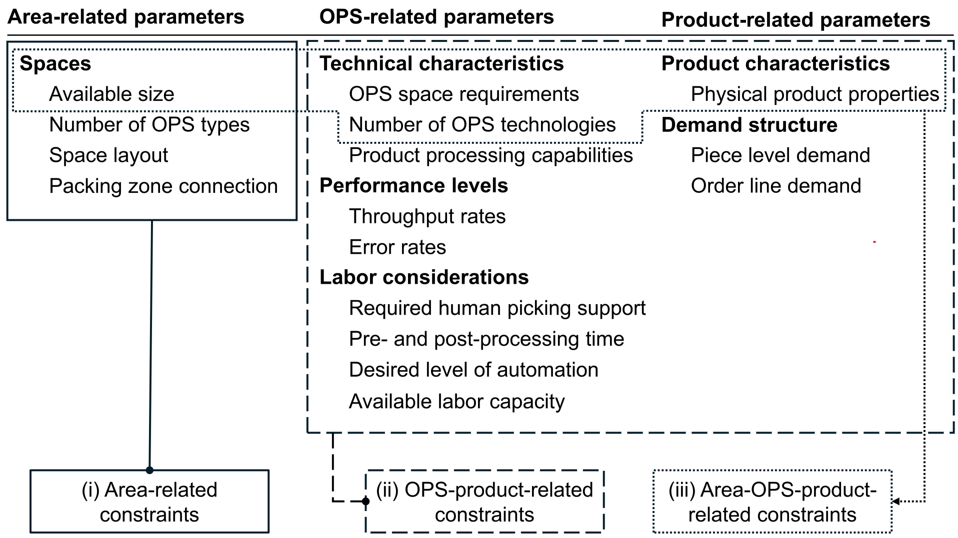


Figure 5.2: Parameters and constraints identified for the selection and assignment problem

(i) Area-related constraints Only one OPS is operated in a distinctive area of the picking zone. As a result, the *number of OPS types* in one area needs to be restricted. If multiple different OPSs are used within a warehouse, the picking zone needs to be separated into different areas with a number of associated space units. Space units can only be merged if the same OPS is installed given individual characteristics and specific setup requirements (e.g., for storage racks, product diversity or human-machine interactions). Moreover, at least one space unit needs to be directly connected to the packing and shipping zone (or any other type of consolidation spot) to reduce travel time and ensure the material flow. The OPS selection problem is thus also constrained by the individual *space layout* and the *packing zone connection*. To illustrate these two factors, Figure 5.3 shows both (i) feasible and (ii) non-feasible combinations for this area-related constraint. For example, the violation on the lower left layout occurs as Area 1, consisting of Space 1 and 2, does not have direct access to the packing and shipping zone. The violation on the lower right materializes as only edge-to-edge combinations are possible (diagonal or corner-to-corner connections between spaces are not sufficient because they do not allow OPS intermobility in an area).

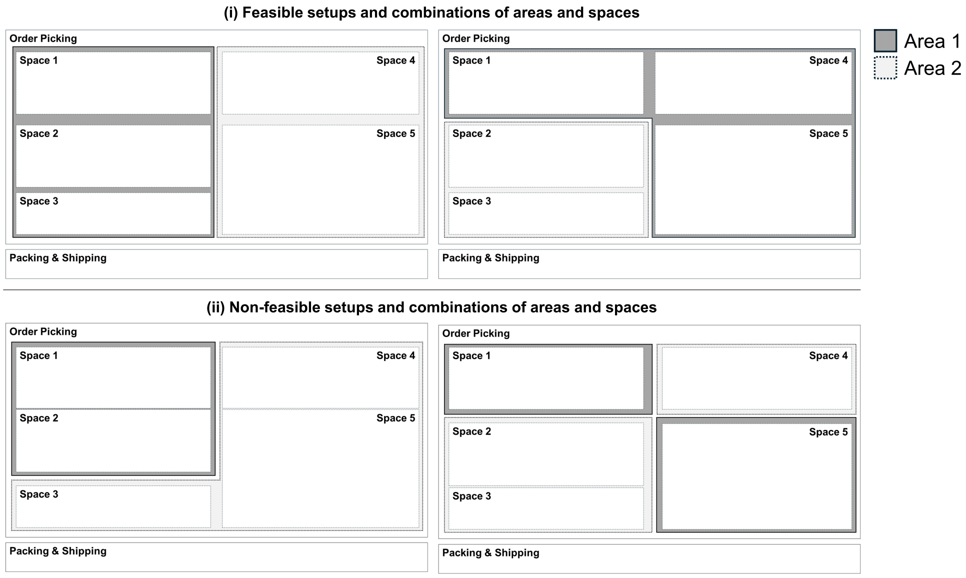


Figure 5.3: Illustration of area constraints including feasible and non-feasible space combinations

(ii) OPS-product-related constraints OPSs are classified into different OPS types (see Table 5.1) and come with distinctive technical characteristics, performance levels and associated labor considerations. Products within a warehouse are characterized by their product characteristics (such as dimensions and other physical properties) and their underlying demand structure. To fulfill a customer order, each product must be retrieved by an OPS. However, especially when dealing with automated OPSs, not every product can be picked by every OPS. For example, not all OPSs are able to deal with every physical product property. These may be product dimensions such as length, width, height, volume or weight or other product characteristics such as shape, unit load, packaging material, hardness, and auxiliary requirements (e.g., if the product requires bin transportation or is bulky, fragile, hazardous, refrigerated or perishable). *Physical product* *properties* of each individual product need to be matched with the *product processing capabilities* of the OPS. That means ensuring that OPSs are physically able to retrieve the products (Mohsen, 2010) as the process would otherwise result in many additional manual activities (Binos et al., 2020; Marchet et al., 2015; Yoon and Sharp, 1996).

Both *piece level demand* and *order line demand* determine the workload and thus *throughput rates* on piece and order line level need to be included (Rouwenhorst et al., 2000). Experts also frequently mentioned the average downtime of OPSs, in particular of robotized ones (e.g., for maintenance or battery charging), which reduces the actual productive time and the average throughput rate. The capacity required for each OPS is obtained by matching the demand rates of assigned products and the throughput rates (Pazour et al., 2014). The possibility of assigning further products to an OPS also depends on the remaining OPS capacity. A further performance criterion is the *error rate* of each OPS. This represents the pick quality and is expressed as the share of wrongly picked (and shipped) order lines.

As most of the OPSs need additional manual support (Lorson et al., 2022), labor requirements need to be considered. Each OPS has a *required human picking support* (assessed via the amount of manual input needed to run one module of an OPS). Picking with robotized AMRs, for instance, typically requires one human worker for three or four robots (Trebilcock, 2020). Additionally, OPS-specific replenishment (e.g., the restocking effort required) and rework activities (e.g., correcting the potential errors of an OPS detected in a quality check) result in *pre- and post-processing time* of human operators. For example, Pazour et al. (2014) assume that replenishment takes about a quarter of the picking time. Furthermore, our interviewees particularly emphasized that labor shortages had evolved to be a main reason for the growing number of automation projects and constitute a relevant limiting factor when running warehouses, often resulting in a *desired level of automation.* Companies thus tend to decide on a specific share of products or orders to be picked by automated machines or robots.

Also, managers are forced to consider the maximum *available labor capacity* within the perimeter of the warehouse.

(iii) Area-OPS-product-related constraints The best use of the spatial area can be made by combining area-, OPS- and product-related factors. OPS automation within a warehouse is naturally constrained by the total *available size*, i.e., available spatial capacity expressed in the amount of square meters or cubic meters available. This is particularly important when either picking areas are being upgraded or the space under consideration is pre-determined (e.g., automation in existing warehouses). The total space required for one OPS depends on the number of modules selected. Each module has different *OPS space requirements* in terms of operating space, determined through technical characteristics such as necessary aisle width or maximum lift height, and the necessary storage space for assigned products. The latter is determined by the *physical product properties* (which are in this case the physical dimensions) and the inventory for each product. Hence, warehouse managers need to ensure that assigned products and their resulting storage volume do not exceed the maximum OPS space capacity available for product storage as well as the total *available size*. Finally, usually only one specific technology of an OPS type is applied within one warehouse (e.g., not multiple different FAPRs of different providers), restricting the *number of OPSs technologies* of an OPS type in a warehouse.

5.2.2 Related literature on OPS selection

This section relates literature to the decision problem identified above. While there is a wide range of publications dealing with the design and control of individual OPSs (see, e.g., Van der Gaast and Weidinger (2022)), there is only a nascent and small field of research concerning OPS selection (Boysen et al., 2019; Azadeh et al., 2019). We concentrate on literature published in the last decade since the advent of advanced automation and robotized technologies. A first related OPS selection model was introduced by Shen et al. (2010). They compare manual picker-to-parts and semi-automated parts-to-picker setups (i.e., AS/RS and a carousel solution). Other OPSs like SMRs, AGVs or FAPRs are not evaluated. They minimize costs by reducing throughput time and apply a genetic algorithm to allocate orders to OPSs. While the authors consider the number of space units necessary for the respective OPS, they only partly acknowledge physical product properties (i.e., the space devoted to the respective product and its length). It therefore cannot be guaranteed that the technical characteristics of the selected OPSs match the product dimensions and characteristics. Although manual picking is included as a potential OPS, not all labor considerations (e.g., for pre- and post-picking activities) are modeled. Ekren and Heragu (2012) compare the performance of two OPSs (namely an aisle-based shuttle system with a crane-based AS/RS). They simulate different scenarios in which the number of modules of the two OPSs are varied and cost implications assessed. They simplify the scenarios by allowing both OPSs to process all products. Space assignment is only partially considered by including the available size via constraints for the number of aisles, tiers and bays per aisle, and a maximum warehouse capacity. They only focused on pallets, making the study hard to generalize for other types of warehouses where product-level considerations are necessary. Pazour et al. (2014) developed the first decision model to minimize setup, module, labor and error costs while simultaneously allocating product groups to selected OPSs. By doing so, they consider space capacities of one module in number of products, acknowledging the importance of differences among OPSs, but not of the total available size. Although the authors compare different solutions (A-frame and parts-to-picker workstation setups) with a variety of manual picking solutions, they do not include novel OPSs. Furthermore, physical properties are not considered. Specifically, the number of products one OPS module can hold is not based on physical product properties (such as dimensions or other characteristics), but only on an average. Also, products are clustered into different demand groups, neglecting the fact that similar demanded products may have totally different physical product properties. Hence, necessary OPS-product-related constraints are not considered. Bozer and Aldarondo (2018) analyze an SMR setup and a miniload aisle-based shuttle solution using a simulation with the objective of minimizing the number of modules based on different performance criteria. They do not directly assign products or spaces to the OPS as they simulate each OPS separately. While they categorize the products into small, medium and large (depending on their size), physical characteristics or exact dimensions as well as product level information are lacing, and demand is assumed to be equal for each category.

5.2.3 Summary and research contribution

The pertinent literature is scarce, but serves as a starting point for the advanced OPS selection and assignment problem. A review shows that only two optimization approaches and two simulations are currently available. However, not all decisions and constraints identified for the problem at hand have yet been considered. Section 5.2.1 reveals the necessity of performing OPS, product and space decisions simultaneously. While all of the contributions identified covered at least the selection and sizing of the OPSs, the assignment of products and spaces to the selected OPSs is only partially covered. Each OPS requires a dedicated space, and different OPS cannot operate technically and efficiently within one area. The separation of the order picking zone into dedicated areas for each OPS selected and the assignment of space units to the area constitutes an open research gap. Only Shen et al. (2010) also determine also the space units required, but do not include the particularities of the warehouse layout (see Figure 5.3). The assignment of products to OPSs based on physical product properties and the product processing capabilities of the OPSs is not specified in any current contribution. Shen et al. (2010) and Bozer and Aldarondo (2018) only partially consider some product characteristics, whereas Pazour et al. (2014) make product assignments on a product group level where the groups are based on demand data and not product properties. This shows that research did not consider the assignment of the items to the OPS on an individual product level. This bears the risk that products are allocated to an OPS, but the OPS may not be capable of processing it in the end and non-feasible solutions are obtained as a result. Furthermore, only the simulation model of Bozer and Aldarondo (2018) covers novel OPSs, (i.e., SMRs) while Shen et al. (2010) and Pazour et al. (2014) are the only ones who compare more than two different OPS types in their model. FAPRs or AMRs are not considered in any model. Two of the contributions are restricted to the selection of two OPSs (Ekren and Heragu, 2012; Bozer and Aldarondo, 2018). The comparison with manual picking is included in Shen et al. (2010) and Pazour et al. (2014), while only Pazour et al. (2014) diligently considered costs for manual labor tasks in semi-automated, pre-, and post-picking activities. Also, none of the existing research considered labor availability.

As it is an imperative to include the selection and assignment decisions as well as important characteristics to mirror novel developments and necessary constraints, existing literature on the OPS selection problem is limited in its practicability and comprehensibility. Table 5.2 highlights these findings and the research gap.

Table 5.2: Related research for the strategic OPS selection

		Objective function and decisions				**Important characteristics**				
	Solution		Selection	Assignment		No.	Labor	Labor	Physical	Product
Authors	**approach**[1]	Costs[2]	OPSs	Product	Space	OPSs[3]	consid.[4]	avail.[5]	properties	level data
Shen et al. (2010)	Opt		✓	(✓)	✓	3	(✓)		(✓)	
Ekren/Heragu [2012]	Sim	S-M	✓		(✓)	2				
Pazour et al. (2014)	Opt	S-M-L-E	✓	(✓)	(✓)	3	✓			
Bozer/Aldarondo [2018]	Sim	(S-M)	✓		(✓)	2*	(✓)		(✓)	
This paper	Opt	S-M-L-E	✓	✓	✓	4*	✓	✓	✓	✓

✓= included; (✓) = partially included;
[1] Opt = Optimization, Sim = Simulation;
[2] Costs regarded for setup (S), module (M), labor (L), and error (E) costs; () indicate indirect considerations, blank other objectives
[3] Maximum number of different OPSs types in numerical study; * incl. novel OPS
[4] Consideration of manual picking as an OPS and costs of manual support required for semi-automated-, pre- and post-picking tasks
[5] Availability of manual labor

5.3 Development of the decision model

This section formalizes the selection and assignment problem and develops an MIP model to minimize total costs while adhering to area, OPS-product and area-OPS-product constraints. It defines the optimal technology mix of OPSs ($o \in O$) and assigns warehouse areas ($a \in A$), warehouse space units ($s \in S$) and products ($j \in J$) to the technologies selected. The related decisions made by the warehouse planner are the determination of the number of modules of an OPS (expressed by the integer number of modules selected for an OPS n_o), the assignment of a product to an OPS (expressed by binary assignment variable $x_{j,o}$), and the assignment of an OPS to a dedicated warehouse area and space (expressed by the binary assignment variable $z_{a,s,o}$). Because OPSs can be fully automated or require some manual support, we divide the set of OPSs O into fully automated OPSs (O^+) and human-supported OPSs (O^-) in the following, such that $O^+, O^- \subseteq O$, $O^+ \cup O^- = O$ and $O^+ \cap O^- = \oslash$. Each product $j \in J$ has certain product specifications and each OPS $o \in O$ has certain technology specifications. We define the subset $O(j)$ that denotes an index set of all OPSs o that are compatible and can pick the product j. Table 5.3 summarizes the notation.

Table 5.3: Notation

Sets and indices	
A	Set of areas with $a = 1, ..., \|A\|$
E	Set of OPS types with $e = 1, ..., \|E\|$
J	Set of products with $j = 1, ..., \|J\|$
O	Set of OPSs with $o = 1, ..., \|O\|$, with subset O^+ for fully automated OPSs and subset O^- for human-supported OPSs
$O(j)$	Subset of OPSs $O(j) \subseteq O$ depending on product $j \in J$, if OPS $o \in O$ is able to pick product $j \in J$
S	Set of space units with $s = 1, ..., \|S\|$
Area-related parameters	
$b_{s,s'}$	Binary; 1 if space unit $s \in S$ is a bordering neighbour (i.e., adjacently located) to space unit $s' \in S$, 0 otherwise
g_s	Space size of unit $s \in S$ [in space units]
p_s	Binary; 1 if space unit $s \in S$ is located next (has access) to the packing zone, 0 otherwise
OPS-related parameters	

Continued on next page

Table 5.3 – *Continued from previous page*

c_o^{err}	Annual downstream error costs per module of OPS technology $o \in O$ depending on the number of order lines processed [currency units]
c^{lab}	Annual labor costs of one human worker [currency units]
c_o^{mod}	Annual total costs per module of OPS technology $o \in O$ [currency units]
c_o^{set}	Annualized, one-time setup costs of OPS $o \in O$ [currency units]
f_o	Space-volume utilization coefficient of an OPS $o \in O$, which is dependent on the maximum operating height of the OPS
k_o	Integer number of modules of an OPS $o \in O$ that one human worker may operate simultaneously when supporting picking
$t_{e,o}$	Binary; 1 if OPS $o \in O$ belongs to OPS type $e \in E$, 0 otherwise
α_o	Degree of automation of an OPS $o \in O$, with $0 \leq \alpha \leq 1$ (0 if manual and 1 if fully automated)
β^{min} (β^{max})	Minimum (maximum) share that should be picked by automated OPS
δ_o	Post-processing time of one order line at OPS $o \in O$ [in time units]
ϵ	Maximum number of different OPS types in one area
λ_o	Pre-processing time of one unit at OPS $o \in O$ [in time units]
$\overline{\rho}_o$ ($\underline{\rho}_o$)	Throughput rate of OPS $o \in O$ [in order lines per period] ([in units per period])
Ψ	Maximum integer number of available human workers [in number of workers]
Product-related parameters	
d_j (D_j)	Total demand of product $j \in J$ [units per period] ([order lines per period])
m_j	Volume-coefficient of the product $j \in J$, depending on the dimensions (i.e., width, length and height) and the space needed to store the inventory of the product
Decision variables	
n_o	Integer number of modules selected for OPS $o \in O$
$x_{j,o}$	Binary; 1 if product $j \in J$ is assigned to OPS $o \in O$, 0 otherwise
$z_{a,s,o}$	Binary; 1 if space unit $s \in S$ is located in area $a \in A$ which is served by OPS $o \in O$, 0 otherwise
Auxiliary variables	
u_o	Integer number of human workers required to support picking of semi-automated OPS $o \in O^-$
v_o	Integer number of human workers required for pre- and post-picking tasks of OPS $o \in O$
w_o	Binary; 1 if OPS $o \in O$ is selected, 0 otherwise
$y_{a,o}$	Binary; 1 if OPS $o \in O$ is assigned to area $a \in A$, 0 otherwise

Objective function Equation (5.1) represents the objective function to minimize total annual decision-relevant costs TC. It consists of four parts. The first part accounts for one-time setup cost c_o^{set} of each OPS $o \in O$ if an OPS is utilized. The binary variable w_o is equal to 1 if the OPS $o \in O$ is applied and otherwise 0. The related costs are implementation costs independent of the number of OPS modules that are selected. To account for an annual cost level, the one-time costs are offset by an annual given present worth economic factor. The second part comprises the annual module operating costs c_o^{mod} of one OPS module. It contains annual depreciation (e.g., purchase price of one module) and volume- and product-independent operating costs (e.g., maintenance, insurance and other module costs that occur during operation). The variable n_o denotes the integer number of modules selected. The third part includes the constraint that human workers may be necessary to support and execute operational activities that depend on the OPS o. The labor costs are denoted by c^{lab}. The related number of workers required is expressed by the integer variable u_o (denoting the number of employees required to support the picking process) and the integer variable v_o (denoting the number of employees required for pre- and post-picking tasks). For example, these costs arise for additional OPS-specific replenishment, picking and packing time. The final part represents error expenses which occur when incorrectly picked products are identified and claimed after they have left the warehouse. The downstream error costs per order line are denoted by c_o^{err} and depend on the OPS o. They may arise once a product j is assigned to a specific OPS o, $x_{j,o} = 1$. Since the correction processes are usually outsourced to external service providers, this constitutes a separate cost element.

$$Min\ TC = \sum_{o \in O} \left(c_o^{\text{set}} \cdot w_o + c_o^{\text{mod}} \cdot n_o + c^{\text{lab}} \cdot (u_o + v_o) + \sum_{j \in J} c_o^{\text{err}} \cdot D_j \cdot x_{j,o} \right) \tag{5.1}$$

(i) Area-related constraints Constraints (5.2) to (5.8) ensure that areas and OPSs are properly connected to the picking space. The total picking space of the warehouse is divided into space units $s \in S$. The space units can be grouped into flexibly large areas $a \in A$. Not all available spaces $s \in S$ and areas $a \in A$ have to be used. Constraints (5.2) ensure that each area opened, indicated by $y_{a,o} = 1$, can only be served by one OPS.

Constraints (5.3) ensure that each space unit $s \in S$ can only be assigned to a maximum of one area $a \in A$. In Constraints (5.4), an area $a \in A$, space $s \in S$, and OPS $o \in O$ combination $z_{a,s,o}$ is excluded as long as no OPS is assigned to an area and $y_{a,o} = 0$. If more than one space unit $s \in S$ is assigned to an area, then at least one edge-to-edge connection of a space unit $s \in S$ to another space unit within the same area is required. This is maintained by Constraints (5.5). Diagonal (corner-to-corner) connections between space units are not sufficient since an area-wide movement of the OPSs cannot be guaranteed. To ensure that at least one space unit $s \in S$ within an area $a \in A$ has access to the packaging area (denoted by $p_s = 1$), Constraints (5.6) is introduced. Additionally, Constraints (5.6) state that only areas $a \in A$ served by an OPS $o \in O, y_{a,o} = 1$ can be occupied by space units. To avoid symmetrical solutions, we use areas with a lower index first and impose Constraints (5.7). From a practical point of view, a lower index might represent better access to the packaging area. These constraints also have the efficiency advantage of eliminating permutations and thus preventing an exponentially increasing solution space. The binary variables are denoted in Constraints (5.8).

$$\sum_{o \in O} y_{a,o} \leq 1 \qquad \forall a \in A \qquad (5.2)$$

$$\sum_{a \in A} \sum_{o \in O} z_{a,s,o} \leq 1 \qquad \forall s \in S \qquad (5.3)$$

$$z_{a,s,o} \leq y_{a,o} \qquad \forall a \in A, s \in S, o \in O \qquad (5.4)$$

$$\sum_{s'' \in S} b_{s,s''} \cdot z_{a,s'',o} \geq z_{a,s,o} + z_{a,s',o} - 1 \qquad \forall a \in A, s \in S, s' \in S : s' \neq s, o \in O \qquad (5.5)$$

$$\sum_{s \in S} \sum_{o \in O} p_s \cdot z_{a,s,o} \geq \sum_{o \in O} y_{a,o} \qquad \forall a \in A \qquad (5.6)$$

$$\sum_{o \in O} y_{a,o} \geq \sum_{o \in O} y_{a+1,o} \qquad \forall a \in \{0, 1 \ldots |A| - 1\} \qquad (5.7)$$

$$y_{a,o}; z_{a,s,o} \in \{0, 1\} \qquad \forall a \in A, s \in S, o \in O \qquad (5.8)$$

(ii) OPS-product-related constraints Constraints (5.9) ensure that each product $j \in J$ has to be picked by exactly one OPS o that is capable of processing this product, ensured by the subset $O(j)$. Constraints (5.9) thus allow products to be assigned only to an OPS that is capable of processing this product. For instance, a manual worker cannot pick a heavily packed pallet without assistance or an SMR cannot pick a product that exceeds the maximum shelf space of its rack. Constraints (5.10) ensure that capacity is sufficiently available and not exceeded, which can be scaled via the number of modules deployed n_o to pick the period demand d_j of the products $j \in J$ that are assigned to it. Similarly, Constraints (5.11) determine the capacity needed with regard to the number of order lines D_j of a product j. The actual picking process may involve human-supported activities before, during or after the picking that all depend on the selected OPS $o \in O$. Constraints (5.12) guarantee that sufficient workers v_o are available for any manual work that may arise for tasks before picking (denoted by λ_o for processing time of tasks before picking) and after picking (denoted by δ_o for processing time of tasks after picking). The first term on the left-hand side represents the capacity required for preparing the picking (e.g., for inventory replenishment), calculated by the demand d_j divided by the rate λ_o required for one unit at OPS $o \in O$. The second term comprises the post-processing time that depends on a post-picking effort δ_o for each order line D_j of a product processed by OPS o. For some semi-automated OPS ($o \in O^-$), it may be necessary to employ extra human workers u_o to support the picking activity. Constraints (5.13) determine the associated number of employees required, whereby k_o represents a parameter for the number of modules n_o that one person may operate simultaneously of a certain OPS $o \in O$. Constraints (5.14) are intended to meet potential labor shortages where the total number of available workers is represented by the parameter Ψ. Constraints (5.15) enable a certain minimum (maximum) level of picking automation β^{min} (β^{max}) to be enforced, if there is a desired limit on the degree of automation due to labor law requirements, for example. The parameter α_o represents the degree of automation of each OPS $o \in O$. Constraints (5.16) define the domains of the decision variables.

$$\sum_{o\in O(j)} x_{j,o} = 1 \qquad \forall j \in J \tag{5.9}$$

$$\sum_{j\in J} \frac{d_j}{\underline{\rho}_o} \cdot x_{j,o} \leq n_o \qquad \forall o \in O \tag{5.10}$$

$$\sum_{j\in J} \frac{D_j}{\overline{\rho}_o} \cdot x_{j,o} \leq n_o \qquad \forall o \in O \tag{5.11}$$

$$\sum_{j\in J} \frac{d_j}{\lambda_o} \cdot x_{j,o} + D_j \cdot \delta_o \cdot x_{j,o} \leq v_o \qquad \forall o \in O \tag{5.12}$$

$$\frac{n_o}{k_o} \leq u_o \qquad \forall o \in O^- \tag{5.13}$$

$$\sum_{o\in O} (u_o + v_o) \leq \Psi \tag{5.14}$$

$$\beta^{min} \cdot \sum_{o\in O}\sum_{j\in J} x_{jo} \leq \sum_{o\in O}\sum_{j\in J} \alpha_o \cdot d_j \cdot x_{j,o} \leq \beta^{max} \cdot \sum_{o\in O}\sum_{j\in J} x_{j,o} \tag{5.15}$$

$$x_{j,o} \in \{0,1\}; \quad n_o, v_o, u_o \in \mathbb{N}_0 \qquad \forall j \in J, o \in O \tag{5.16}$$

(iii) Area-OPS-product-related constraints Constraints (5.17) relate the number of modules used n_o to the variable w_o for the selection of certain OPS $o \in O$, where M represents a large coefficient (“Big M”). Constraints (5.18) ensures that each selected OPS $o \in O$ can only be operated in one area $y_{a,o} = 1, a \in A$, with $y_{a,o}$ expressing the area assignment variable for all areas $a \in A$. As products are uniquely assigned to one OPS and each OPS to an area (cf. Constraints (5.2) and (5.9)), the products are directly assigned to an area. Constraints (5.19) ensure that only selected OPSs $n_o > 0$ can be assigned to an area $a \in A$. Each OPS $o \in O$ belongs to exactly one OPS type $e \in E$ (e.g., AMR, AGV). Only a maximum number ϵ of different OPSs belonging to one type of OPS can be applied within the warehouse (see Constraints (5.20)). This keeps the technical complexity within the warehouse manageable for the operator. To accommodate the products $j \in J$ in the designated areas $a \in A$, the available area volume (total of the space volumes assigned to the area) must be greater than the total stored product volumes (Constraints (5.21)). The defined area volume (left side of Constraints (5.21)) is calculated from the assigned spaces $s \in S$ with their square meters g_s and the space-volume utilization coefficient f_o of the OPS $o \in O$, whereas the demand volume (right side of Constraints (5.21)) is composed of the demand d_j and a volume coefficient m_j. Constraints (5.22) define w_o as a binary variable.

$$n_o \leq M \cdot w_o \qquad \forall o \in O \qquad (5.17)$$

$$\sum_{a \in A} y_{a,o} = w_o \qquad \forall o \in O \qquad (5.18)$$

$$\sum_{a \in A} y_{a,o} \leq n_o \qquad \forall o \in O \qquad (5.19)$$

$$\sum_{o \in O} t_{e,o} \cdot w_o \leq \epsilon \qquad \forall e \in E \qquad (5.20)$$

$$\sum_{a \in A} \sum_{s \in S} g_s \cdot f_o \cdot z_{a,s,o} \geq \sum_{j \in J} d_j \cdot m_j \cdot x_{j,o} \qquad \forall o \in O \qquad (5.21)$$

$$w_o \in \{0, 1\} \qquad \forall o \in O \qquad (5.22)$$

5.4 Numerical experiments and case study

This section will first present a case study where the model has been applied to optimize the OPS selection and assignment for an e-commerce warehouse. The second part derives managerial insights using sensitivity analyses and further data sets that are based on the case study setting. All numerical experiments were conducted on Windows 10 64-bit with Intel Core i7-8550U processor and 16-GB memory. The tests are implemented in Python 3.10 and solved with Gurobi 9.5.1.

5.4.1 Case study

Data sources Area- and product-related data as well as labor considerations are obtained by a case study with a Western European e-commerce warehouse (e.g., product and demand data) and a warehouse planning company (e.g., OPS data). In the representative warehouse, 180 manual workers are required to conduct the picking, while an additional 32 are necessary for the pre- and post-picking tasks, totaling 212 manual workers. The warehouse runs two shifts a day for eight hours each. The related manual OPS and its costs constitute the benchmark. Information about different OPS systems are retrieved from a warehouse planning company. The partner company specializes in robotic implementations and has access to comprehensive information across multiple OPS providers. This enabled us to retrieve technical, performance and financial data across a variety of OPSs. Please note that some specific costs and performance data of OPSs are subject to confidentiality and proprietary information. In the following, we first describe the data along area-, product- and OPS-related factors. We then analyze the solution of the decision model.

Area-related data of the e-commerce warehouse The order picking zone of the warehouse under investigation consists of six equally sized space units with $g_s = 1,000$ square meters each (which is the space for both storing and picking). Three space units are next to each other with direct access to the packing area with $p_1 = p_2 = p_3 = 1$, and three arranged in the same manner just behind the first row with $p_4 = p_5 = p_6 = 0$. Hence, space unit 1 may be combined with 2 and 4, space unit 2 with 1, 3 and 5, and so forth.

Product-related data of the e-commerce warehouse Daily piece level demand d_j (in order lines D_j) per product varies from a minimum of 1(1) to a maximum of 1,325(982), with a median of 6(2). The average order size amounts to 1.96 items, typical for e-commerce warehousing (Boysen et al., 2019). The demand data of the products results in a product-demand curve of 17/80, meaning that 17% of the 8,522 products (i.e., 1.399) are responsible for 80% of the total demand (i.e., 198,912 out of 250,186). In terms of order lines, 15% of the products are responsible for 80% of the total order lines (i.e., 101,539 out of 127,705). The median product is 10cm long, 8cm wide, 10cm high, and weighs 0.88kg. 11% of the products are round, while 89% are rectangular shaped. None of the products need to be picked with a corresponding box or pallet, and only a few are classified as bulky. Half of the products have cardboard packaging, while a quarter have plastic or no packaging, respectively. Only 9% of the items are fragile, while none are hazardous, perishable or refrigerated. These physical product properties (such as dimensions and other product characteristics) have been available for most of the products. In cases where information is lacking, the data set has been supplemented with simulated data based on representative product sets.

OPS-related data from warehouse planner We obtained data for four OPS types that are relevant for e-commerce warehousing. Manual picking, AMR-assisted picking, SMRs with advanced workstations, and FAPRs are usual OPSs in the context of our application. For example, AMRs and SMRs are particularly suitable as they are able to handle small orders and large assortments (Boysen et al., 2019). Also, robotic solutions are highly scalable, a major point raised in the expert discussions. Only one OPS type can operate in one area (i.e., $\epsilon = 1$). We obtained detailed technical, performance and cost data on manual picking, 5 AMRs, 1 SMR and 1 FAPR. We consider a floor space utilization (rack height) of 0.55(2.7m) for manual picking and AMR-assisted picking, while SMRs have a better utilization but lower racks (0.65(1.8m)) (see also Roser (2021)). For FAPRs, the values obtained are 0.60(1.80m). Further, about 70% of the rack space volume can be utilized by products. This results in a total space-volume utilization coefficient (f_o) of 1.04 for manual picking, 1.04 for AMR-assisted picking, 0.82 for SMR-assisted picking and 0.76 for FAPRs. In terms of product dimensions and characteristics, the OPSs naturally differ, with manual picking being the most flexible (100% of products can be picked) and FAPRs being the least flexible (43% can be retrieved). The hourly throughput rate $\underline{\rho}_o$ in units ($\overline{\rho}_o$, in number of order lines) is 175(88) for manual picking, 245(123) for AMR-assisted picking, 350(175) for SMRs and 84(42) for FAPRs. Related pre-picking times (λ_o) are represented by replenishment rates. These are 1,000 units per hour for all OPSs, except for SMRs with 900. In the case study there is no further post-picking effort required and hence $\delta_o = 0, \forall o \in O$. The observed error rates are 0.4% for FAPRs, 0.35% for manual picking, 0.2% for AMRs, and 0.15% for SMRs, while one wrongly delivered order line results in a customer return. This extra process for each error that occurs is penalized with 10 cost units (EHI, 2019). The setup costs c_o^{set} for SMRs are 100,000 cost units, irrespective of the number of modules. For other solutions, setup costs are included in module costs as they depend on a single module and not the entire OPS. Total module costs c_o^{mod} consist of purchasing costs, annual operating costs and annual maintenance costs. The purchasing costs are 40,000 cost units per AMR, 25,000 per SMR, and 80,000 per FAPR. For (semi-)automated picking, annual operating costs (e.g., for energy spending) vary between 5,000 and 8,000 cost units. Annual maintenance costs are in the same range for all (semi-)automated OPSs. One human is able to work with three AMRs and four SMRs, respectively (i.e., $k_o \in \{3; 4\}$; see also e.g., Barry (2022) or Trebilcock (2020)). For manual picking, module costs consist of the annual labor wage (27,300 cost units) and maintenance costs (1,000 cost units) as well as one-time costs for hiring and training one human employee (4,000 cost units). The labor costs for one worker for pre- and post-picking tasks are equivalent to the labor costs for one manual picker. To take into account the costs on an annual basis, the minimum attractive rate of return (MARR) is used to calculate the annual given present worth economic factor, which also considers the underlying planning horizon in terms of years. This is then used as the depreciation coefficient for the setup and module costs. The MARR for the case study is set to 6% and spans a time horizon of five years. To obtain a best possible view on potential solutions, we do not apply labor restrictions (i.e., $\Psi = \infty$) nor a desired level of automation (i.e., $\beta^{min} = 0; \beta^{max} = 1$) in the base case.

Analysis of the case study solution An optimal OPS selection and product and area assignment leads to a total cost reduction of 57% compared to the status quo (i.e., manual picking only, see Figure 5.4). This represents annual savings of approximately 4.2 million cost units. The cost-optimal solution for this case study is a combination of SMRs and manual picking. Sometimes, warehouse operators prefer solely picker-to-parts solutions instead. When only picker-to-parts OPSs become available, this results in a savings potential of 42% (by mainly utilizing AMR-assisted picking).

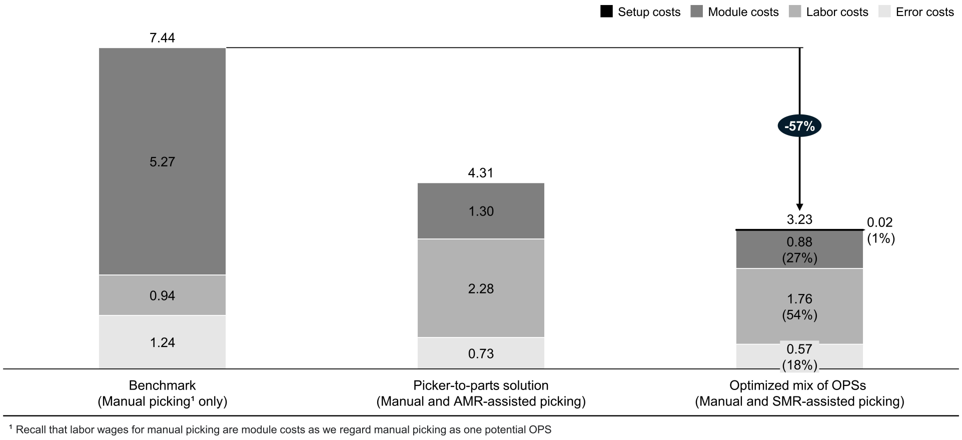


Figure 5.4: Comparison of annual costs, in million cost units

Figure 5.5 illustrates the optimal OPS selection and space assignment. Two OPSs across two areas and four spaces were selected: SMRs are assigned for Area 1, which combines Spaces 1, 2 and 4. Manual picking is selected for Area 2, which only consists of Space 3, manual picking is selected. Spaces 5 and 6 are not allocated, as the spatial volume is not required to store and retrieve the products.

Table 5.4 summarizes the results. The optimal number of modules (n_o) for Area 1 is determined as 45 SMRs, while Area 2 is equipped with four human pickers per day. About 98% of the products are assigned to SMRs, while only about 2% remain for manual picking. 24 employees in total support the picking activities (u_o) in Area 1 per day to jointly retrieve the products with the robots. For pre- and post-picking tasks (v_o), 34 workers per day for Area 1 and two per day for Area 2 are necessary. While 212 human operators were required prior to the implementation of robots, now only 64 per day are needed, resulting in a manual workforce reduction of 70%.

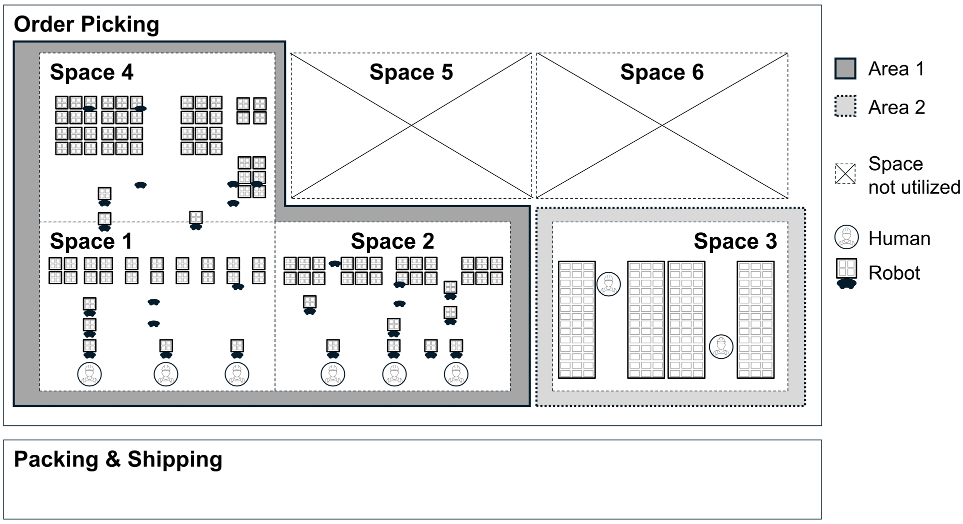


Figure 5.5: Illustrative assignment of OPSs to areas and spaces in the case study

This proves the necessity of diligently modeling the labor costs for both picking with semi-automated OPSs and working on pre- and post-picking activities. These costs are responsible for the highest expenses compared to setup, module and error costs. In the case example, the associated labor costs for the optimized mix of OPS amount to 54% of the total costs.

Table 5.4: Overview of the solution structure of the optimized mix versus the benchmark

		Area 1	Area 2	Benchmark
Decision variables	n_o	45 SMRs	4 human pickers	180 human pickers
	$x_{j,o}$	8,342 products	180 products	8,522 products
	$z_{a,s,o}$	Space 1, 2 & 4 for SMRs	Space 3 for MP	Space 1 & 2 for MP
Auxiliary variables	u_o	24 workers	-	-
	v_o	34 workers	2 workers	32 workers

MP = Manual picking

For Area 1, where SMRs were selected, 72% of the space is utilized. For Area 2, this is only 16%, indicating that SMRs better utilize the available space in this setting. Interestingly, the model did not decide to choose only two spaces for SMRs. Although space utilization overall would have been improved, the higher numbers of human pickers for Area 2 would have resulted in higher costs. Turning to the individual product assignment, more voluminous items are assigned to manual picking. The average volume for a product stored and retrieved in Area 1 is $0.002m^3$, while this value totals $0.01m^3$ in Area 2. Given that SMRs carry racks that are split into shelves (Lamballais et al., 2020), this result makes sense as humans can generally put retrieved products on a picking cart, allowing for a larger flexibility of product dimensions. However, no products were automatically assigned to manual picking given volume capabilities. Instead, only 11 products are retrieved via manual picking given bulky product characteristics which SMRs cannot handle. This also shows that it was necessary to open at least one space for manual picking, and proves the necessity to include product processing capabilities of OPSs as well as physical product properties in optimization approaches. 50% of the products assigned to SMRs are from slow-moving items. This accords with a common advantage of transferring slow-moving products from manual to automated picking: Long travel times with little reward (i.e., small number of picks) are avoided by automating the many slow-moving products (see also Pazour et al. (2014)).

5.4.2 Run time evaluation based on case study data

We performed run time analyses based on case study data to investigate the solvability of the model. To increase the problem scope, the number of products is varied while the remaining parameters remain constant or are adjusted proportionally. To determine the number of products, we randomly draw from the total of 8,522 products with equal probability for all products, including lay-back. To ensure that demand and spatial volume increase proportionally with the number of products, we place theses values in the same ratio as with the original total demand and volume. The sizes of the individual spaces increase proportionally with the number of products. All other parameters remain unchanged. As for the number of products, we start with 4,000 and increase each by a factor of 5 until we reach a total of 100,000 products, resembling a very large warehouse (Logistics

Management, 2018). For each number of products, we drew 20 samples and obtained $3 \times 20 = 60$ further data sets. Table 5.5 summarizes the results. For 100,000 products, a maximum run time of 741.3 seconds is required. Given the strategic nature of the problem, it can be confirmed that the exact solution of the MIP implemented in Gurobi is efficient with regard to run time and solvability of the problem.

Table 5.5: Run time analysis of 20 samples for each product set, in seconds

	Run time (seconds)		
Number of products	min	max	median
4,000	4.0	26.0	7.0
20,000	48.2	91.4	58.6
100,000	236.6	741.3	390.7

Summing up, our case study shows that the proposed model produces a feasible solution that significantly reduces total costs. Second, we prove that the underlying problem and varying problem sizes can be efficiently solved with Gurobi. The results are consistently optimal and can be generated time-efficiently.

5.4.3 Further numerical experiments

In this section we run four further numerical experiments to understand the robustness of solutions and obtain managerial insights. In particular, we vary total demand, area constraints, manual labor-related data, and the performance and cost data of a fully automated OPS.

Demand variations Many warehouses are also likely to see changes in their total demand (Boysen et al., 2019). As a result, we analyze the impact of varying total demand and order lines. We apply four different multipliers for the piece-level and order line demand per product to the data of the case study for this purpose. We further increase space capacity proportional to focus on the implications of the demand. Table 5.6 summarizes the results of this analysis.

Table 5.6: Implications across variations of total demand and order lines

Demand multiplier	Selected modules		Cost share compared	
	SMRs	Human pickers	to base case[1]	to status quo[2]
0.50	22	2	52%	23%
0.75	35	2	76%	33%
1.00	45	4	100%	43%
1.25	56	4	122%	53%
1.50	68	4	146%	63%

[1] Base setting in case study (i.e., demand multiplier 1.00)
[2] Setting before optimization approach with manual picking only

If the cost share compared to the status quo (i.e., manual picking only) is analyzed, it becomes apparent that significant cost savings are generated even with a demand multiplier of 1.5. The same OPSs (i.e., SMRs and human pickers) are selected across all variations, emphasizing the robustness of our case study solution. The number of modules for SMRs changes almost proportionally to the demand multiplier. This seems logical due to the constant throughput rate for all of the potential OPSs and as the major parts are module and labor costs that scale with demand. When looking at the selected modules for manual picking, we see results that resemble the low utilization of manual pickers in our case study. Specifically, reducing the total demand by 0.75 and 0.50 is sufficient to remove one human picker. Also, increasing total demand by 1.25 or 1.50 does not lead to an additional picker compared to the base case because the additional throughput can be handled with the existing capacity. On the other hand, the costs compared to the base case react similarly to the number of modules as module costs decrease or increase depending on the number of modules purchased. Slight deviations from a linear relationship exist due to the number of modules (i.e., number of pickers) for manual picking. Although human workers are responsible for a lower number of products compared to the automated SMRs, they are necessary given their flexibility in being able to handle all products.

Importance of area constraints We run further numerical experiments to understand the magnitude of the area constraints, in particular the available size and the number of space units. We first reduce the available size by half (recall that originally 6.000 square meters were available), and continue to decrease it throughout six sets of analyses. At the same time we allow one, two, three and six space units respectively for the total available space, equally distributed in size. We focus on eight potential OPSs. Alongside the selected manual picking and SMR (SMR1), we introduce six additional SMRs (SMR2-7). We generate their technical data (e.g., throughput) by multiplying the original value with a randomly drawn uniformly distributed number between 0.8 and 1.2. For each of the OPSs product processing capability (e.g., capability to process fragile products), we set the probability to 90% that they can process one specified product characteristic. Table 5.7 reports that the optimal solution is very sensitive to both available size and number of space units, showing the dependence of available space in the OPS selection task and the importance of incorporating area constraints. First, the partitioning of the available square meters into different quantities of space units influences the choice of OPS technologies. Second, if the available space drops below the space requirement of the best possible OPS configuration, the selected OPS technologies will change.

Table 5.7: Implications of changes across available size and space distribution[1]

	SMR types and modules selected across different space units			
Size in sqm	**1 space**	**2 spaces**	**3 spaces**	**6 spaces**
3,000	35x SMR4	15x SMR2, 20x SMR4	29x SMR2, 6xSMR6	29xSMR2, 6x SMR6
2,700	35x SMR4	15x SMR2, 20x SMR4	29x SMR2, 6xSMR6	29xSMR2, 6x SMR6
2,400	35x SMR4	15x SMR2, 20x SMR4	29x SMR2, 6xSMR6	29xSMR2, 6x SMR6
2,100	manual only	18x SMR4, 18x SMR6	23x SMR4, 12x SMR6	23x SMR4, 12x SMR6
1,800	*inf.*	*inf.*	*inf.*	*inf.*

sqm = square meters; inf. = infeasible solution
[1] Note that the number of human workers (modules for manual picking) are not displayed

Implications of labor shortages The logistics sector has an enormous issue hiring and retaining suitable workers, particularly for labor intensive activities such as order picking (Logistics Management, 2021). Hence, we analyze the impact of increasing labor shortages in the following. For this purpose, we both increase the labor costs (we focus on labor wages only in the following) (+15%; +30%) and decrease the available labor (90%; 80%). The baseline for the latter is the figure of 64 human pickers required for two shifts in our case study.

Table 5.8 and Table 5.9 confirm that the ongoing labor shortages are costly for businesses, even when automation is introduced. When only 90% labor is available on the market (in our case study this would mean only 58 manual pickers may be utilized), the warehouse would need to introduce FAPRs to complete all picking jobs. This would result in a 33% cost increase compared to the optimized solution of our case study. When comparing the cost implications between increasing labor costs and increasing labor shortages, it further becomes evident that managers are advised to focus on the retention of employees as losing labor (and not rehiring the same number of staff) is very costly. In our case study, increasing labor costs of 15%(30%) would only result in an 8%(16%) overall cost increase. However, not having enough labor increases costs more dramatically (33% - 48%). Also, manual pickers are necessary to fulfill all customer orders due to necessary pre- and post-picking activities and particular physical product properties. When reducing the available labor to 80%, there are not enough employees available for these tasks and no feasible solutions are obtained.

Table 5.8: Cost changes with varying labor costs and availability

	Available labor		
Labor costs	100%	90%	80%
0%		33%	*inf.*
+15%	8%	40%	*inf.*
+30%	16%	48%	*inf.*

inf. = infeasible solution

Table 5.9: Optimal OPS and module number when varying labor costs and availability

	Available labor		
Labor costs	100%	90%	80%
0%	2 MPs, 45 SMRs	1 MP, 39 SMRs, 35 FAPRs	*inf.*
+15%	2 MPs, 45 SMRs	1 MP, 39 SMRs, 35 FAPRs	*inf.*
+30%	2 MPs, 45 SMRs	1 MP, 39 SMRs, 35 FAPRs	*inf.*

inf. = infeasible solution; MP = manual pickers

The quest for full automation From a theoretical standpoint, the numerical experiment on labor availability proves the necessity of including manual labor in the OPSs selection problem. From a managerial perspective, despite the introduction of automated and robotized systems in warehousing, retaining labor should be a priority for businesses. On the other side of the scale are fully automated robotic implementations. Based on the proprietary data analyzed, FAPRs are still too expensive, too inefficient in operations and too inflexible in terms of physical characteristics to be employed cost efficiently for our case study. However, many companies and research institutes are investing to achieve fully automated robotic solutions for piece-level picking (see also Correll et al. (2018)). In the following, we analyze this issue and estimate when FAPRs are a solid choice in terms of cost, throughput and product processing capabilities. For this purpose, we increase the reported throughput and decrease the cost data (we only focus on the purchase cost of one module in the following) of FAPRs until they are selected (see Table 5.10). If the overall cost level were kept the same, FAPRs would need to increase the throughput rate by a factor of four until they were included in the optimal set of OPSs. The same holds true if an FAPR only cost 75% of the current reported module cost. Also, when reducing the module costs to 50% or even 25%, FAPRs are only selected when the throughput rate is increased by a factor of 3 or 4, while factors of 1 and 2 (irrespective of the cost changes) do not trigger any change in the optimal solution. This shows that the technology is currently too immature to be implemented on a larger scale.

Table 5.10: Optimal OPS selection for varying module costs and throughput of FAPRs

	Throughput rates of FAPRs ($\overline{\rho}_o, \underline{\rho}_o$)	
Module cost of FAPRs (c_o^{mod})	3× higher	4× higher
100%		1 MP, 40 SMRs, 6 FAPRs
75%		1 MP, 40 SMRs, 6 FAPRs
50%	1 MP, 40 SMRs, 8 FAPRs	1 MP, 40 SMRs, 6 FAPRs
25%	1 MP, 40 SMRs, 8 FAPRs	1 MP, 39 SMRs, 9 FAPRs

Only changes compared to case study solution are displayed; MP = Manual picking

5.5 Conclusion

Contribution and managerial implications Boosted by novel technological developments and new players in the industry, a skyrocketing number of automated and robotized OPSs have become available on the market with distinctive characteristics. This calls for more comprehensive decision support when implementing automation and robotics in warehouses. To assist warehouse managers in this novel OPS selection problem, we identify decision-relevant costs and constraints based on a variety of sources (e.g., literature review, interviews, case study). Using a set of selected novel solutions such as AMRs, SMRs and FAPRs, we find significant cost reduction potential (up to 57%) compared to manual picking for a representative case study in an e-commerce warehouse. We further show that despite ongoing automation efforts, retaining labor is still important until fully automated solutions are ready to be implemented cost efficiently across a variety of products. Our model and findings contribute to the strategic decision-making process of selecting the most appropriate OPSs. The numerical study makes it clear that physical characteristics of products and space constraints should not be neglected to produce feasible solutions. This requires simultaneous OPS selection and space and product assignment. Additionally, analyses of automated or robotized OPSs should ensure a holistic approach, including the option of manual picking. On the managerial side, we provide efficient decision support for the selection of suitable OPSs for warehousing managers, system providers and planners. The model can be utilized to inform the decision process of practitioners across a variety of warehousing environments, and gives clear cost minimization potential. The model is highly adaptable to specific business cases, can be altered based on company preferences (e.g., degree of automation), and is applicable for a variety of warehousing types. These can be centralized distribution centers, but also micro or nano fulfilment centers.

Further areas of research This paper introduces a novel decision problem and serves as a starting point for future research. For example, the model takes a strategic standpoint. Future research can build on this and expand the model to achieve a more integrated, hierarchical planning approach by also including tactical and operational planning issues (Shen et al., 2010; Van der Gaast and Weidinger, 2022; Van Gils et al., 2019). Particularly, slotting or batching policies (see Muter and Öncan (2021)) may be integrated. Reiteration in the hierarchical planning of our solution is naturally still necessary (such as allocating spaces to manual pickers, including the length of pick waves, see (Dallari et al., 2009)). The same reasoning is valid for the assumed performance rate of each OPS. While we take an average performance rate that is based on reported data, we acknowledge that the actual throughput depends on many other criteria (e.g., batch size, see Russell and Meller (2003)). Furthermore, we base the decisions on static and stationary demand. A valuable research path will be to investigate the impact of demand variations, seasonality and demand trends on technology selection. We mitigated this issue by varying demand across different scenarios, showing robust solutions with our model. A further validation of the results with advancing technologies that impact OPS capabilities and throughput appears interesting. These may be a decline in OPS prices with a simultaneous throughput increase based on Moore's law (see also Bogue (2016)) or improved space utilization coefficients required to satisfy the characteristics of micro and nano fulfilment centers (Eriksson et al., 2019).

In conclusion, the ever-growing range of automated and robotized OPSs is a current and significant challenge for warehouse managers. As literature on any feasible model that directly and quantifiable compares the OPSs is lacking, we developed a decision model to support the strategic planning going forward. Our contribution will serve to stimulate more research in this context and enhance the importance of analyzing novel automated and robotized solutions from both managerial and theoretical perspectives.

Appendix

Table 5.A1: Anonymous overview of interview participants

Code	#	Company type	Interviewee role	Warehouse experience years
1C	1	Consultancy	Partner	10-20
2WO	2	Warehouse operator	Director Logistics Development	10-20
3C	3	Consultancy	Expert	5-10
4SP	4	System provider	Manager Picking Solution Design	>20
5C	5	Consultancy	Partner	>20
6SP	6	System provider	Head of Dynamic Systems	10-20
7SP	7	System provider	Product and Sales Manager	10-20
8SP	8	System provider	Group Leader Sales Engineering	10-20

The eight interviews were conducted between November 2020 and January 2021. Additionally, several discussions and interviews with our partner company were held throughout 2021.

6 Conclusion and outlook

This doctoral thesis deals with recent challenges in warehousing, particularly the optimization of human-machine interactions and the selection of suitable OPSs. It equally supports practitioners and researchers in planning and optimizing warehousing operations. Each individual contribution (Chapter 3 – 5) concludes by summarizing theoretical and managerial insights. Additionally, within each of those chapters, potential extensions and future research opportunities are detailed. Aggregated findings and a joint outlook are thus outlined in a more broader context in the following.

6.1 Summary of findings and contributions

The application of **automated and robotized systems for operational activities is transforming warehouse operations**. Compared to traditional, manually operated warehouses, automated systems influence operating policies across warehousing activities. In this way, novel strategic (e.g., which automated solutions to deploy), tactical (e.g., which items to store at which location of the shuttle system), and operational (e.g., which worker to staff at which picking workstation) problems arise.

In a similar manner and magnitude, resulting **human-machine interactions exist across activities and have a significant impact on human behavior and system performance.** In fact, Chapter 3 established behavioral issues across all operational activities, making the optimization of such interactions an imperative for warehouse managers and scholars.

However, human factors and behavior are often disregarded when establishing novel automated solutions for human-machine interactions. Through many statements in experts interviews and the scarcity of literature, it gets clear that both scholars and practitioners alike often miss to incorporate behavioral aspects of human-machine interactions, potentially resulting in undesired and negative impacts on human factors and system performance.

As a starting point to tackle this problem, Chapter 3 established a systematic framework, identified behavioral issues, and developed **four unifying themes including theoretical foundations**. In this way, causal relationships among the various interconnections (such as interaction setup and human factors) and salient variables (such as decision-discretion or motivation) guide future researchers and managers to understand behavioral consequences in their optimization efforts. Moreover, the research agenda helps operations management, and in particular warehousing, researchers to identify potential effects on behavior, while scholars from the field of behavioral science and human factors are encouraged to consider warehousing as an interesting area of application.

Applying one such theory in Chapter 4, that is **goal-setting, proves to engage workers in repetitive activities and to be a source of performance improvement**. In fact, human picking performance is elevated by 5.6% in a human-machine interaction within a semi-automated order picking activity through a goal-setting intervention. As such, goal-setting mechanisms can be established to have a positive effect on human engagement, leading to superior picking results without any kind of monetary incentives. However, the **goal-setting intervention diminished possibilities for humans to informally organize themselves in their working day, with repercussions on satisfaction, self-determination and perceived fairness**. This unexpected, but explainable outcome shows the importance to observe and understand human factors in human-machine interactions. Additionally, the often-advocated support for humans from automated and robotized systems may lead to undesired human behavior because workers prefer to have some kind of decision autonomy.

Chapter 5 turns then to a novel warehousing planning problem by **introducing and formalizing the innovative OPS selection and assignment problem**. In this way, the developed decision-support establishes important decision variables and constraints. Among others, it is demonstrated that the assignment of both spaces and products to OPSs is imperative given the large variety of product properties and the importance of space efficiency. **Applying the decision model results in cost savings up to 57%**, but also ensures a viable and efficient assignment across spaces, OPSs, and products, which previous research has neglected. Furthermore, although the application of automated and robotized systems will continue in the future, retaining labor is established as a key factor to manage operating costs.

6.2 Future areas of research

The findings of this dissertation lead to numerous further research opportunities across three fields.

First, Chapter 3 detailed 18 research questions that guide the way for subsequent studies on behavioral issues in human-machine interactions. In this way, scholars can profit from the developed theoretical foundations and unifying themes to design and implement their research. Additionally, those also serve as starting points for the exploration of behavioral issues beyond operational activities, specifically for tactical and strategic tasks in the warehouse.

Second, while the conducted study in Chapter 4 answered one of the above-mentioned research questions, it also developed new pathways for future endeavours. Particularly, the identified informal agreements, which are determined to be the main reason for the deterioration of the human factors, may be integrated into further operating policies. This serves as a pool for future studies not only in the setting of the semi-automated picking system, but for many interactions where humans and machines collaborate with each other.

Third, the introduction of the novel OPS selection and assignment model in Chapter 5 opens future areas of research as well. For example, the optimization effort can be extended with a larger range of OPSs, different types of warehouses, or stochastic demand. Additionally, integrative approaches that combine this strategic problem with tactical or operational facets (such as batching policies) may yield to further insights.

The development and implementation of automated and robotized warehousing systems indeed transform the way warehouses work. What has been seen in the past as a pure cost center of firms is now a focal point in the value contribution of companies. The large variety of human-machine interactions and order picking solutions deliver significant challenges to establish an optimized material and information flow. This doctoral thesis has both delivered valuable insights and set the pathway for more research towards impactful perspectives on the warehouse of the future.

Acknowledgements I want to tremendously thank Alexander Hübner for supervising my doctoral journey. Not only his dedication and academic acumen are unparalleled, but also his mentoring and kindness. I truly appreciated the collaboration and look back with joy on the multiple workshops, working sessions, and celebrations. I also want to thank in this context Andreas Fügener for providing excellent guidance on my academic path as well as for contributing on an exceptional level as a co-author in two articles. Also, thank you to Fabian Schäfer for patiently joining me as a co-author in my third article.

Further, I want to express my thankfulness towards all practitioners involved in the expert discussions and cooperations for their valuable time and insights. Particularly, the dedication of the warehouse manager involved in the second contribution has to be mentioned. My acknowledgements also extend to editors and reviewers for helpful ideas on improving the submitted papers.

I am also very thankful for the generous support of the Hanns-Seidel-Stiftung that included a scholarship for outstanding students from funds of the German Federal Ministry of Education and Research (Bundesministerium für Bildung und Forschung).

Finally, I want to express my deepest gratitude to Anna for always believing in me and for always being by my side. I am beyond grateful to always felt the support and encouragement from my whole family including my parents, siblings, and grandparents. Each and every one of you made my doctoral journey possible, and I thank you all from the bottom of my heart.

Bibliography

Akkermans, H., Oppen, W., Wynstra, F., Voss, C., 2019. Contracting outsourced services with collaborative key performance indicators. Journal of Operations Management 65 (1), 22–47.

Alvesson, M., Kärreman, D., 2007. Constructing mystery: Empirical matters in theory development. Academy of Management Review 32 (4), 1265–1281.

Anand, G., Chandrasekaran, A., Sharma, L., 2021. Sustainable process improvements: Evidence from intervention–based research. Journal of Operations Management 67 (2), 212–236.

Aromaa, S., Väätänen, A., Aaltonen, I., Goriachev, V., Helin, K., Karjalainen, J., 2020. Awareness of the real-world environment when using augmented reality head-mounted display. Applied ergonomics 88, 103145.

Azadeh, K., De Koster, R., Roy, D., 2019. Robotized and automated warehouse systems: Review and recent developments. Transportation Science 53 (4), 917–945.

Bai, B., Dai, H., Zhang, D., Zhang, F., Hu, H., 2021. The impacts of algorithmic work assignment on fairness perceptions and productivity. Academy of Management Proceedings 2021 (1), 12335.

Bainbridge, H., 2002. Best Practice for the Procurement and Conduct of Non-destructive Testing. Health and Safety Executive, Bottle, UK.

Banerjee, A. G., Barnes, A., Kaipa, K. N., Liu, J., Shriyam, S., Shah, N., Gupta, S. K., 2015. An ontology to enable optimized task partitioning in human-robot collaboration for warehouse kitting operations. In: Proceedings of SPIE, Next-Generation Robotics II; and Machine Intelligence and Bio-inspired Computation: Theory and Applications IX. International Society for Optics and Photonics, Bellingham, WA.

Banker, S., 2020. Automation is the future of warehousing. `https://www.forbes.com/sites/stevebanker/2020/07/31/automation-is-the-future-of-warehousing/?sh=3b70633230f4` [08.08.2022].

Baraglia, J., Cakmak, M., Nagai, Y., Rao, R., Asada, M., 2016. Initiative in robot assistance during collaborative task execution. In: 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI). IEEE, New York, pp. 67–74.

Barry, B., 2022. 4 considerations for employing picking robotics and automation in fulfillment. `https://www.fcbco.com/blog/4-considerations-for-employing-picking-robotics-and-automation` [01.08.2022].

Bartholdi III, J. J., Hackman, S. T., 2020. Warehouse & distribution science. release 0.98.1. Supply Chain and Logistics Institute.

Batt, R. J., Gallino, S., 2019. Finding a needle in a haystack: The effects of searching and learning on pick-worker performance. Management Science 65 (6), 2624–2645.

Bendoly, E., Croson, R., Goncalves, P., Schultz, K., 2010. Bodies of knowledge for research in behavioral operations. Production and Operations Management 19 (4), 434–452.

Bendoly, E., Donohue, K., Schultz, K. L., 2006. Behavior in operations management: Assessing recent findings and revisiting old assumptions. Journal of Operations Management 24 (6), 737–752.

Bernstein, E. S., 2012. The transparency paradox. Administrative Science Quarterly 57 (2), 181–216.

Binos, T., Adamopoulos, A., Bruno, V., 2020. Decision support research in warehousing and distribution: A systematic literature review. International Journal of Information Technology & Decision Making 19 (03), 653–693.

Bogue, R., 2016. Growth in e-commerce boosts innovation in the warehouse robot market. Industrial Robot: An International Journal 43 (6), 583–587.

Boudreau, J., Hopp, W., McClain, J. O., Thomas, L. J., 2003. On the interface between operations and human resources management. Manufacturing & Service Operations Management 5 (3), 179–202.

Boyer, K. K., Swink, M. L., 2008. Empirical elephants–Why multiple methods are essential to quality research in operations and supply chain management. Journal of Operations Management 26 (3), 338–344.

Boysen, N., De Koster, R., Füßler, D., 2021. The forgotten sons: Warehousing systems for brick-and-mortar retail chains. European Journal of Operational Research 288 (2), 361–381.

Boysen, N., De Koster, R., Weidinger, F., 2019. Warehousing in the e-commerce era: A survey. European Journal of Operational Research 277 (2), 396–411.

Boysen, N., Füßler, D., Stephan, K., 2020. See the light: Optimization of put–to–light order picking systems. Naval Research Logistics 67 (1), 3–20.

Boysen, N., Konrad, S., Weidinger, F., 2022. Efficient order consolidation in warehouses: The product-to-order-assignment problem in warehouses with sortation systems. IISE Transactions 54 (10), 963–975.

Boywitz, D., Schwerdfeger, S., Boysen, N., 2019. Sequencing of picking orders to facilitate the replenishment of A-Frame systems. IISE Transactions 51 (4), 368–381.

Bozer, Y. A., Aldarondo, F. J., 2018. A simulation-based comparison of two goods-to-person order picking systems in an online retail setting. International Journal of Production Research 56 (11), 3838–3858.

Brand, A., Allen, L., Altman, M., Hlava, M., Scott, J., 2015. Beyond authorship: Attribution, contribution, collaboration, and credit. Learned Publishing 28 (2), 151–155.

Bryman, A., Bell, E., 2015. Business Research Methods, 4th Edition. Oxford University Press, Oxford, UK.

Cantor, D. E., Jin, Y., 2019. Theoretical and empirical evidence of behavioral and production line factors that influence helping behavior. Journal of Operations Management 65 (4), 312–332.

Chandrasekaran, A., Treville, S., Browning, T., 2020. Editorial: Intervention–based research (IBR)—What, where, and how to use it in operations management. Journal of Operations Management 66 (4), 370–378.

Chun, Y., Harris, S. L., Chandrasekaran, A., Hill, K., 2022. Improving care transitions with standardized peer mentoring: Evidence from intervention based research using randomized control trial. Journal of Operations Management 68 (2), 185–214.

CNBC, 2020. Mark Zuckerberg just made a bold claim: We're going to get a 'breakthrough' in tech glasses this decade. https://www.cnbc.com/2020/01/09/zuckerberg-expects-breakthrough-augmented-reality-glasses-this-decade.html [03.07.2022].

Corgnet, B., Gómez-Miñambres, J., Hernán-González, R., 2015. Goal setting and monetary incentives: When large stakes are not enough. Management Science 61 (12), 2825–3096.

Correll, N., Bekris, K. E., Berenson, D., Brock, O., Causo, A., Hauser, K., Okada, K., Rodriguez, A., Romano, J. M., Wurman, P. R., 2018. Analysis and observations from the first Amazon picking challenge. IEEE Transactions on Automation Science and Engineering 15 (1), 172–188.

Cragg, T., Loske, D., 2019. Perceived work autonomy in order picking systems: An empirical analysis. IFAC-PapersOnLine 52 (13), 1872–1877.

Creswell, J. W., 2009. Research Design: Qualitative, Quantitative, and Mixed-Method Approaches, 3rd Edition. Sage Publications, Thousand Oaks, CA.

Cropanzano, R., Paddock, L., Rupp, D. E., Bagger, J., Baldwin, A., 2008. How regulatory focus impacts the process-by-outcome interaction for perceived fairness and emotions. Organizational Behavior and Human Decision Processes 105 (1), 36–51.

Croson, R., Schultz, K., Siemsen, E., Yeo, M. L., 2013. Behavioral operations: The state of the field. Journal of Operations Management 31 (1-2), 1–5.

D'Addona, D. M., Bracco, F., Bettoni, A., Nishino, N., Carpanzano, E., Bruzzone, A. A., 2018. Adaptive automation and human factors in manufacturing: An experimental assessment for a cognitive approach. CIRP Annals 67 (1), 455–458.

Dallari, F., Marchet, G., Melacini, M., 2009. Design of order picking system. The International Journal of Advanced Manufacturing Technology 42 (1-2), 1–12.

Davarzani, H., Norrman, A., 2015. Toward a relevant agenda for warehousing research: Literature review and practitioners' input. Logistics Research 8 (1), 1–18.

De Koster, M. B. M., Roy, D., Lim, Y. F., Kumar, S., 2020. Managing autonomous and IoT-driven intralogistics operations: Call for papers. Production and Operations Management.

De Koster, R., Le-Duc, T., Roodbergen, K. J., 2007. Design and control of warehouse order picking: A literature review. European Journal of Operational Research 182 (2), 481–501.

De Koster, R. B., Stam, D., Balk, B. M., 2011. Accidents happen: The influence of safety-specific transformational leadership, safety consciousness, and hazard reducing systems on warehouse accidents. Journal of Operations Management 29 (7-8), 753–765.

De Vries, J., De Koster, R., Stam, D., 2016a. Aligning order picking methods, incentive systems, and regulatory focus to increase performance. Production and Operations Management 25 (8), 1363–1376.

De Vries, J., De Koster, R., 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 54 (8), 2260–2274.

De Vries, J., De Koster, R., Stam, D., 2016c. Safety does not happen by accident: Antecedents to a safer warehouse. Production and Operations Management 25 (8), 1377–1390.

Deci, E. L., Olafsen, A. H., Ryan, R. M., 2017. Self-determination theory in work organizations: The state of a science. Annual Review of Organizational Psychology and Organizational Behavior 4 (1), 19–43.

Deci, E. L., Ryan, R. M., 2000. The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. Psychological Inquiry 11 (4), 227–268.

DeHoratius, N., Rabinovich, E., 2011. Field research in operations and supply chain management. Journal of Operations Management 29 (5), 371–375.

Delasay, M., Ingolfsson, A., Kolfal, B., Schultz, K., 2019. Load effect on service times. European Journal of Operational Research 279 (3), 673–686.

Denyer, D., Tranfield, D., van Aken, J. E., 2008. Developing design propositions through research synthesis. Organization Studies 29 (3), 393–413.

Dietvorst, B. J., Simmons, J. P., Massey, C., 2018. Overcoming algorithm aversion: People will use imperfect algorithms if they can (even slightly) modify them. Management Science 64 (3), 983–1476.

Digman, J. M., 1990. Personality structure: Emergence of the five-factor model. Annual Review of Psychology 41 (1), 417–440.

Doerr, K. H., Mitchell, T. R., Klastorin, T. D., Brown, K. A., 1996. Impact of material flow policies and goals on job outcomes. Journal of Applied Psychology 81 (2), 142–152.

Donohue, K., Katok, E., Leider, S. (Eds.), 2018. The Handbook of Behavioral Operations. Wiley online library. John Wiley & Sons, Inc, Hoboken, NJ, USA.

Donohue, K., Özer, Ö., Zheng, Y., 2020. Behavioral operations: Past, present, and future. Manufacturing & Service Operations Management 22 (1), 191–202.

Downing, C. O., Geller, E. S., 2012. A goal-setting and feedback intervention to increase ID-checking behavior: An assessment of social validity and behavioral impact. Journal of Organizational Behavior Management 32 (4), 297–306.

Eckerd, S., DuHadway, S., Bendoly, E., Carter, C. R., 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 67 (2), 261–275.

Edmondson, A. C., Mcmanus, S. E., 2007. Methodological fit in management field research. Academy of Management Review 32 (4), 1246–1264.

Edwards, H., Edwards, D., 2018. How Tesla "shot itself in the foot" by trying to hyper-automate its factory. `https://qz.com/1261214/how-exactly-tesla-shot-itself-in-the-foot-by-trying-to-hyper-automate-its-factory` [25.07.2022].

Egger, J., Masood, T., 2020. Augmented reality in support of intelligent manufacturing – A systematic literature review. Computers & Industrial Engineering 140, 106195.

EHI, 2019. Studie: Versand- und Retourenmanagement im e-commerce 2019. EHI study, available at `https://www.ehi.org/produkt/studie-versand-und-retourenmanagement-im-e-commerce-2019-pdf-version/`.

Eisenhardt, K. M., 1989. Building theories from case study research. Academy of Management Review 14 (4), 532–550.

Ekren, B., Heragu, S. S., 2012. Performance comparison of two material handling systems: AVS/RS and CBAS/RS. International Journal of Production Research 50 (15), 4061–4074.

Ekren, B. Y., 2017. Graph-based solution for performance evaluation of shuttle-based storage and retrieval system. International Journal of Production Research 55 (21), 6516–6526.

Eriksson, E., Norrman, A., Kembro, J., 2019. Contextual adaptation of omni-channel grocery retailers' online fulfilment centres. International Journal of Retail & Distribution Management 47 (12), 1232–1250.

Flick, U., von Kardorff, E., Steinke, I. (Eds.), 2004. A Companion to Qualitative Research. Sage Publications, London, UK.

Forbes, 2021. Meet the billionaire robot overlord reinventing Walmart's warehouses. `https://www.forbes.com/sites/amyfeldman/2021/12/13/meet-the-billionaire-robot-overload-reinventing-walmarts-warehouses/?sh=77850f7878f4` [08.08.2022].

Fottner, J., Clauer, D., Hormes, F., Freitag, M., Beinke, T., Overmeyer, L., Gottwald, S. N., Elbert, R., Sarnow, T., Schmidt, T., Reith, K. B., Zadek, H., Thomas, F., 2021. Autonomous systems in intralogistics – State of the art and future research challenges. Logistics Research 14 (1).

Fragapane, G., de Koster, R., Sgarbossa, F., Strandhagen, J. O., 2021. Planning and control of autonomous mobile robots for intralogistics: Literature review and research agenda. European Journal of Operational Research 294 (2), 405–426.

Friebel, G., Heinz, M., Krueger, M., Zubanov, N., 2017. Team incentives and performance: Evidence from a retail chain. American Economic Review 107 (8), 2168–2203.

Friebel, G., Heinz, M., Zubanov, N., 2022. Middle managers, personnel turnover, and performance: A long–term field experiment in a retail chain. Management Science 68 (1), 211–229.

Friesike, S., Flath, C. M., Wirth, M., Thiesse, F., 2019. Creativity and productivity in product design for additive manufacturing: Mechanisms and platform outcomes of remixing. Journal of Operations Management 65 (8), 735–752.

Füßler, D., Boysen, N., 2019. High-performance order processing in picking workstations. EURO Journal on Transportation and Logistics 8 (1), 65–90.

Gino, F., Pisano, G., 2008. Toward a theory of behavioral operations. Manufacturing & Service Operations Management 10 (4), 676–691.

Glikson, E., Woolley, A. W., 2020. Human trust in artificial intelligence: Review of empirical research. Academy of Management Annals 14 (2), 627–660.

Glock, C. H., Grosse, E. H., Abedinnia, H., Emde, S., 2019. An integrated model to improve ergonomic and economic performance in order picking by rotating pallets. European Journal of Operational Research 273 (2), 516–534.

Goerg, S. J., Kube, S., 2012. Goals (th)at work. Goals, monetary incentives, and workers' performance. Preprints of the Max Planck Institute for Research on Collective Goods 2012/19.

Gombolay, M., Bair, A., Huang, C., Shah, J., 2017. Computational design of mixed-initiative human–robot teaming that considers human factors: Situational awareness, workload, and workflow preferences. The International Journal of Robotics Research 36 (5-7), 597–617.

Gombolay, M. C., Gutierrez, R. A., Clarke, S. G., Sturla, G. F., Shah, J. A., 2015. Decision-making authority, team efficiency and human worker satisfaction in mixed human–robot teams. Autonomous Robots 39 (3), 293–312.

Groop, J., Ketokivi, M., Gupta, M., Holmström, J., 2017. Improving home care: Knowledge creation through engagement and design. Journal of Operations Management 53-56 (1), 9–22.

Grosse, E. H., Glock, C. H., Jaber, M. Y., Neumann, W. P., 2015. Incorporating human factors in order picking planning models: Framework and research opportunities. International Journal of Production Research 53 (3), 695–717.

Grosse, E. H., Glock, C. H., Neumann, W. P., 2017. Human factors in order picking: A content analysis of the literature. International Journal of Production Research 55 (5), 1260–1276.

Gu, J., Goetschalckx, M., McGinnis, L. F., 2007a. Research on warehouse design and performance evaluation: A comprehensive review. European Journal of Operational Research 203 (3), 539–549.

Gu, J., Goetschalckx, M., McGinnis, L. F., 2007b. Research on warehouse operation: A comprehensive review. European Journal of Operational Research 177 (1), 1–21.

Guest, G., Bunce, A., Johnson, L., 2006. How many interviews are enough? Field Methods 18 (1), 59–82.

Gutelius, B., Theodore, N., 2019. The future of warehouse work: Technological change in the U.S. logistics industry. Working Partnerships USA, UC Berkeley Labor Center, available at https://laborcenter.berkeley.edu/pdf/2019/Future-of-Warehouse-Work.pdf.

Hedenstierna, C. P. T., Disney, S. M., Eyers, D. R., Holmström, J., Syntetos, A. A., Wang, X., 2019. Economies of collaboration in build–to–model operations. Journal of Operations Management 65 (8), 753–773.

Higgins, E. T., 1997. Beyond pleasure and pain. American Psychologist 52 (12), 1280–1300.

Holton, J. A., 2012. The Coding Process and Its Challenges. In: Bryant, A., Charmaz, K. (Eds.), The SAGE Handbook of Grounded Theory. Sage Publications, London, UK, pp. 265–289.

Hübner, A., Holzapfel, A., Kuhn, H., 2015. Operations management in multi-channel retailing: An exploratory study. Operations Management Research 8 (3), 84–100.

Hübner, A., Holzapfel, A., Kuhn, H., Obermair, E., 2019. Distribution in omnichannel grocery retailing: An analysis of concepts realized. In: Gallino, S., Moreno, A. (Eds.), Operations in an omnichannel world. Springer series in supply chain management. Springer Nature Switzerland AG, Cham.

Ibanez, M. R., Staats, B. R., 2018. Behavioral Empirics and Field Experiments. In: Donohue, K., Katok, E., Leider, S. (Eds.), The Handbook of Behavioral Operations. Wiley online library. John Wiley & Sons, Inc, Hoboken, NJ, USA, pp. 121–147.

IEEE, 2020. Symphony of humans and machines or "Modern Times"? https://cmte.ieee.org/futuredirections/2020/01/03/symphony-of-humans-and-machines-or-modern-times [07.08.2022].

IFR, 2020. Executive summary world robotics 2020 industrial robots. World Robotics Report, IFR, available at https://ifr.org/img/worldrobotics/Executive_Summary_WR_2020_Industrial_Robots_1.pdf.

IHCI, 2020. Amazon has >200,000 kiva robots navigating through its warehouses across the globe. https://ihci.sbf.org.sg/docs/default-source/application-guides/ag28-rfw/app-guide-28_case-1.pdf?sfvrsn=8e119260_2 [06.08.2022].

Ilk, N., Shang, G., Goes, P., 2020. Improving customer routing in contact centers: An automated triage design based on text analytics. Journal of Operations Management 66 (5), 553–577.

Instawork, 2022. State of warehouse labor. https://info.instawork.com/state-of-warehouse-labor [25.07.2022].

ISO, 2012. ISO 8373:2012. Robots and robotic devices — Vocabulary. ISO/TC 299 Robotics.

Jaghbeer, Y., Hanson, R., Johansson, M. I., 2020. Automated order picking systems and the links between design and performance: A systematic literature review. International Journal of Production Research 58 (15), 4489–4505.

Johnson, M., Burgess, N., Sethi, S., 2020. Temporal pacing of outcomes for improving patient flow: Design science research in a National Health Service hospital. Journal of Operations Management 66 (1-2), 35–53.

Jung, J. H., Schneider, C., Valacich, J., 2010. Enhancing the motivational affordance of information systems: The effects of real-time performance feedback and goal setting in group collaboration environments. Management Science 56 (4), 724–742.

Kadir, B. A., Broberg, O., 2020. Human well-being and system performance in the transition to Industry 4.0. International Journal of Industrial Ergonomics 76, 102936.

Kaipia, R., Holmström, J., Småros, J., Rajala, R., 2017. Information sharing for sales and operations planning: Contextualized solutions and mechanisms. Journal of Operations Management 52, 15–29.

Kaplan, A. D., Sanders, T., Hancock, P. A., 2019. The relationship between extroversion and the tendency to anthropomorphize robots: A Bayesian analysis. Frontiers in Robotics and AI 5, 135.

Karwowski, W., 2005. Ergonomics and human factors: The paradigms for science, engineering, design, technology and management of human-compatible systems. Ergonomics 48 (5), 436–463.

KC, D., 2020. Worker productivity in operations management. Foundations and Trends® in Technology, Information and Operations Management 13 (3), 151–249.

Kellogg, K. C., Valentine, M. A., Christin, A., 2020. Algorithms at work: The new contested terrain of control. Academy of Management Annals 14 (1), 366–410.

Kihlstrom, J. F., Park, L., 2018. Cognitive psychology: Overview. Reference Module in Neuroscience and Biobehavioral Psychology, 1–14.

Kostami, V., Rajagopalan, S., 2014. Speed–Quality trade-offs in a dynamic model. Manufacturing & Service Operations Management 16 (1), 104–118.

Kretschmer, V., Plewan, T., Rinkenauer, G., Maettig, B., 2018. Smart Palletisation: Cognitive Ergonomics in Augmented Reality Based Palletising. In: Karwowski, W., Ahram, T. Z. (Eds.), Intelligent Human Systems Integration. Advances in Intelligent Systems and Computing. Springer, Cham, Switzerland, pp. 355–360.

Kudelska, I., Niedbal, R., 2020. Technological and organizational innovation in warehousing process – Research over workload of staff and efficiency of picking stations. E&M Economics and Management 23 (3), 67–81.

Kumar, S., Mookerjee, V., Shubham, A., 2018. Research in operations management and information systems interface. Production and Operations Management 27 (11), 1893–1905.

Lamballais, T., Roy, D., De Koster, R., 2020. Inventory allocation in robotic mobile fulfillment systems. IISE Transactions 52 (1), 1–17.

Land, M. J., Thürer, M., Stevenson, M., Fredendall, L. D., Scholten, K., 2021. Inventory diagnosis for flow improvement—A design science approach. Journal of Operations Management 67 (5), 560–587.

Langer, M., Landers, R. N., 2021. The future of artificial intelligence at work: A review on effects of decision automation and augmentation on workers targeted by algorithms and third-party observers. Computers in Human Behavior 123, 106878.

Langley, A., Smallman, C., Tsoukas, H., van de Ven, A. H., 2013. Process studies of change in organization and management: Unveiling temporality, activity, and flow. Academy of Management Journal 56 (1), 1–13.

Larco, J. A., De Koster, R., Roodbergen, K. J., Dul, J., 2017. Managing warehouse efficiency and worker discomfort through enhanced storage assignment decisions. International Journal of Production Research 55 (21), 6407–6422.

Larson, R. B., 2019. Controlling social desirability bias. International Journal of Market Research 61 (5), 534–547.

Latham, G. P., Locke, E. A., 2006. Enhancing the benefits and overcoming the pitfalls of goal setting. Organizational Dynamics 35 (4), 332–340.

Latham, G. P., Mitchell, T. R., Dossett, D. L., 1978. Importance of participative goal setting and anticipated rewards on goal difficulty and job performance. Journal of Applied Psychology 63 (2), 163–171.

Lewis-Beck, M. S., Bryman, A., Liao, T. F. (Eds.), 2004. The Sage Encyclopedia of Social Science Research Methods. Sage Publications, Thousand Oaks, CA.

Lincoln, Y. S., Guba, E. G., 1985. Naturalistic Inquiry. Sage Publications, Thousand Oaks, CA.

Linderman, K., Schroeder, R. G., Choo, A. S., 2006. Six Sigma: The role of goals in improvement teams. Journal of Operations Management 24 (6), 779–790.

Loch, C. H., Wu, Y., 2005. Behavioral operations management. Foundations and Trends® in Technology, Information and Operations Management 1 (3), 121–232.

Locke, E. A., Latham, G. P., 1990. Work motivation and satisfaction: Light at the end of the tunnel. Psychological Science 1 (4), 240–246.

Locke, E. A., Latham, G. P., 2002. Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. American Psychologist 57 (9), 705–717.

Locke, E. A., Latham, G. P., 2019. The development of goal setting theory: A half century retrospective. Motivation Science 5 (2), 93–105.

Logistics Management, 2018. 2018 Warehouse / Distribution center survey: Labor crunch driving automation. https://www.logisticsmgmt.com/article/2018_warehouse_distribution_center_survey_labor_crunch_driving_automation [09.06.2022].

Logistics Management, 2021. There is no way to dodge the ongoing labor shortage issues. https://www.logisticsmgmt.com/article/there_is_no_way_to_dodge_the_ongoing_labor_shortage_issues [01.08.2022].

Logistics Manager, 2020. Hermes increases returns capacity with L.A.C. https://www.logisticsmanager.com/hermes-increases-returns-capacity-with-l-a-c-conveyors-automation [08.08.2022].

Lorson, F., Fügener, A., Hübner, A., 2022. New team mates in the warehouse: Human interactions with automated and robotized systems. IISE Transactions (forthcoming).

Lots of Bots, 2022. We help you find the right robot. https://www.lotsofbots.com/en/robot-finder// [05.06.2022].

Lughofer, E., Smith, J. E., Tahir, M. A., Caleb-Solly, P., Eitzinger, C., Sannen, D., Nuttin, M., 2009. Human–Machine interaction issues in quality control based on online image classification. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans 39 (5), 960–971.

Maettig, B., Hering, F., Doeltgen, M., 2019. Development of an intuitive, visual packaging assistant. In: Nunes, I. L. (Ed.), Advances in Human Factors and Systems Interaction. Vol. 781 of Advances in Intelligent Systems and Computing. Springer, Cham, Switzerland, pp. 19–25.

Maettig, B., Kretschmer, V., 2019. Smart packaging in intralogistics: An evaluation study of human-technology interaction in applying new collaboration technologies. In: Hawaii International Conference on System Sciences. HICSS, Grand Wailea, Hawaii, pp. 739–748.

Magazino, 2021. Magazino extends robot fleet at Zalando. https://www.magazino.eu/magazino-extends-robot-fleet-at-zalando/?lang=en [25.07.2022].

Marchet, G., Melacini, M., Perotti, S., 2015. Investigating order picking system adoption: A case-study-based approach. International Journal of Logistics Research and Applications 18 (1), 82–98.

Marks, M. A., Mathieu, J. E., Zaccaro, S. J., 2001. A temporally based framework and taxonomy of team processes. Academy of Management Review 26 (3), 356–376.

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

Masood, T., Egger, J., 2019. Augmented reality in support of Industry 4.0—Implementation challenges and success factors. Robotics and Computer-Integrated Manufacturing 58, 181–195.

Matusiak, M., De Koster, R., Saarinen, J., 2017. Utilizing individual picker skills to improve order batching in a warehouse. European Journal of Operational Research 263 (3), 888–899.

McCutcheon, D. M., Meredith, J. R., 1993. Conducting case study research in operations management. Journal of Operations Management 11 (3), 239–256.

McKinsey & Company, 2021a. Automation in European grocers' supply chains has reached its tipping point. https://www.mckinsey.com/industries/retail/our-insights/automation-in-european-grocers-supply-chains-has-reached-its-tipping-point [07.08.2022] .

McKinsey & Company, 2021b. Navigating the labor mismatch in US logistics and supply chains. =https://www.mckinsey.com/business-functions/operations/our-insights/navigating-the-labor-mismatch-in-us-logistics-and-supply-chains [16.06.2022].

Meller, R. D., Klote, J. F., 2004. A throughput model for carousel/VLM pods. IIE Transactions 36 (8), 725–741.

Miles, M. B., Huberman, A. M., Saldaña, J., 2013. Qualitative Data Analysis. A Methods Sourcebook, 4th Edition. Sage Publications, Thousand Oaks, CA.

Modern Materials Handling, 2019. Top 20 worldwide materials handling system suppliers in 2019. https://www.mmh.com/article/top_20_worldwide_materials_handling_system_suppliers_in_2019 [08.08.2022].

Mohsen, M. D. H., 2010. A framework for selection of material handling equipment in manufacturing and logistics facilities. Journal of Manufacturing Technology Management 21 (2), 246–268.

Moniz, A., Krings, B.-J., 2016. Robots working with humans or humans working with robots? Searching for social dimensions in new human-robot interaction in industry. Societies 6 (3), 1–23.

Monostori, L., Kádár, B., Bauernhansl, T., Kondoh, S., Kumara, S., Reinhart, G., Sauer, O., Schuh, G., Sihn, W., Ueda, K., 2016. Cyber-physical systems in manufacturing. CIRP Annals 65 (2), 621–641.

Muter, İ., Öncan, T., 2021. Order batching and picker scheduling in warehouse order picking. IISE Transactions 54 (5), 435–447.

Newman, D. T., Fast, N. J., Harmon, D. J., 2020. When eliminating bias isn't fair: Algorithmic reductionism and procedural justice in human resource decisions. Organizational Behavior and Human Decision Processes 160, 149–167.

Öhman, M., Hiltunen, M., Virtanen, K., Holmström, J., 2021. Frontlog scheduling in aircraft line maintenance: From explorative solution design to theoretical insight into buffer management. Journal of Operations Management 67 (2), 120–151.

Oliva, R., Watson, N., 2011. Cross-functional alignment in supply chain planning: A case study of sales and operations planning. Journal of Operations Management 29 (5), 434–448.

Olivia, R., 2019. Intervention as a research strategy. Journal of Operations Management 65 (7), 710–724.

Olsen, T. L., Tomlin, B., 2020. Industry 4.0: Opportunities and challenges for operations management. Manufacturing & Service Operations Management 22 (1), 113–122.

Otto, A., Battaïa, O., 2017. Reducing physical ergonomic risks at assembly lines by line balancing and job rotation: A survey. Computers & Industrial Engineering 111, 467–480.

Papadopoulos, A.-A., Kordonis, I., Dessouky, M., Ioannou, P., 2019. Coordinated freight routing with individual incentives for participation. IEEE Transactions on Intelligent Transportation Systems 20 (9), 3397–3408.

Parasuraman, R., Manzey, D. H., 2010. Complacency and bias in human use of automation: an attentional integration. Human Factors 52 (3), 381–410.

Parasuraman, R., Riley, V., 1997. Humans and automation: Use, misuse, disuse, abuse. Human Factors 39 (2), 230–253.

Parasuraman, R., Sheridan, T. B., Wickens, C. D., 2000. A model for types and levels of human interaction with automation. IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans 30 (3), 286–297.

Parasuraman, R., Sheridan, T. B., Wickens, C. D., 2008. Situation awareness, mental workload, and trust in automation: Viable, empirically supported cognitive engineering constructs. Journal of Cognitive Engineering and Decision Making 2, 140–160.

Pasparakis, A., de Vries, J., de Koster, M., 2021. In control or under control? Human-Robot collaboration in warehouse order picking. Available at SSRN: `https://ssrn.com/abstract=3816533`.

Patel, P. C., Baldauf, C., Karlsson, S., Oghazi, P., 2021. The impact of free returns on online purchase behavior: Evidence from an intervention at an online retailer. Journal of Operations Management 67 (4), 511–555.

Pazour, J.A., Meller, R. D., 2014. A framework and analysis to inform the selection of piece-level order-fulfillment technologies. Progress in Material Handling Research.

Pazour, J. A., Meller, R. D., 2011. An analytical model for A-frame system design. IIE Transactions 43 (10), 739–752.

Perera, H. N., Hurley, J., Fahimnia, B., Reisi, M., 2019. The human factor in supply chain forecasting: A systematic review. European Journal of Operational Research 274 (2), 574–600.

Petersen, C. G., Siu, C., Heiser, D. R., 2005. Improving order picking performance utilizing slotting and golden zone storage. International Journal of Operations & Production Management 25 (10), 997–1012.

Plewan, T., Mättig, B., Kretschmer, V., Rinkenauer, G., 2021. Exploring the benefits and limitations of augmented reality for palletization. Applied ergonomics 90, 103250.

Proctor, R. W., van Zandt, T., 2018. Human factors in simple and complex systems, 3rd Edition. CRC Press, Boca Raton, FL.

Prologis, 2022. Logistics real estate: Highest demand, fastest rent growth in history. https://www.prologisgermany.de/en/logistics-industry-research/logistics-real-estate-highest-demand-fastest-rent-growth-history [05.05.2022].

Qu, S. Q., Dumay, J., 2011. The qualitative research interview. Qualitative Research in Accounting & Management 8 (3), 238–264.

Randall, W. S., Mello, J. E., 2012. Grounded theory: An inductive method for supply chain research. International Journal of Physical Distribution & Logistics Management 42 (8/9), 863–880.

Robert, L. P., Pierce, C., Marquis, L., Kim, S., Alahmad, R., 2020. Designing fair AI for managing employees in organizations: A review, critique, and design agenda. Human–Computer Interaction 35 (5-6), 545–575.

Rodrigue, J.-P., 2020. The Geography of Transport Systems, 5th Edition. Taylor & Francis Group, Milton.

Roels, G., Staats, B. R., 2021. OM Forum—People-centric operations: Achievements and future research directions. Manufacturing & Service Operations Management 23 (4), 745–757.

Roodbergen, K. J., Vis, I. F., 2009. A survey of literature on automated storage and retrieval systems. European Journal of Operational Research 194 (2), 343–362.

Roser, C., 2021. Amazon Kiva storage strategies. https://www.allaboutlean.com/amazon-kiva-storage-strategies/ [09.02.2022].

Rouwenhorst, B., Reuter, B., Stockrahm, V., van Houtum, G. J., Mantel, R. J., Zijm, W., 2000. Warehouse design and control: Framework and literature review. European Journal of Operational Research 122 (3), 515–533.

Roy, D., Nigam, S., De Koster, R., Adan, I., Resing, J., 2019. Robot-storage zone assignment strategies in mobile fulfillment systems. Transportation Research Part E: Logistics and Transportation Review 122, 119–142.

Roy, S., Edan, Y., 2018. Investigating joint-action in short-cycle repetitive handover tasks: The role of giver versus receiver and its implications for human-robot collaborative system design. International Journal of Social Robotics 12, 973–988.

Russell, M. L., Meller, R. D., 2003. Cost and throughput modeling of manual and automated order fulfillment systems. IIE Transactions 35 (7), 589–603.

Salvendy, G. (Ed.), 2012. Handbook of Human Factors and Ergonomics, 4th Edition. Wiley, Hoboken, NJ.

Sanders, M. S., McCormick, E. J., 1993. Human Factors In Engineering and Design, 7th Edition. McGraw-Hill, New York, NY.

Sanders, T., Kaplan, A., Koch, R., Schwartz, M., Hancock, P. A., 2019. The relationship between trust and use choice in human-robot interaction. Human Factors 61 (4), 614–626.

Sauppé, A., Mutlu, B., 2015. The social impact of a robot co-worker in industrial settings. Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, 3613–3622.

Schäfer, F., Lorson, F., Huebner, A., 2022. Finding the right one: Decision model to select cost-efficient order picking solutions. Working paper TU Munich.

Schiffer, M., Boysen, N., Klein, P. S., Laporte, G., Pavone, M., 2022. Optimal picking policies in e-commerce warehouses. Management Science (forthcoming).

Schmidtler, J., Knott, V., Hölzel, C., Bengler, K., 2015. Human centered assistance applications for the working environment of the future. Occupational Ergonomics 12 (3), 83–95.

Schultz, K. L., McClain, J. O., Thomas, L., 2003. Overcoming the dark side of worker flexibility. Journal of Operations Management 21 (1), 81–92.

Schultz, K. L., Schoenherr, T., Nembhard, D., 2010. An example and a proposal concerning the correlation of worker processing times in parallel tasks. Management Science 56 (1), 176–191.

Schulz, R., Kratzer, P., Toussaint, M., 2018. Preferred interaction styles for human-robot collaboration vary over tasks with different action types. Frontiers in Neurorobotics 12 (36).

See, J. E., Drury, C. G., Speed, A., Williams, A., Khalandi, N., 2017. The role of visual inspection in the 21st Century. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 61 (1), 262–266.

Sgarbossa, F., Grosse, E. H., Neumann, W. P., Battini, D., Glock, C. H., 2020. Human factors in production and logistics systems of the future. Annual Reviews in Control 49, 295–305.

Shen, C., Wu, Y., Zhang, D., 2010. A selection method of manual and semi-automated order picking systems based on filling curve and time model. 2010 IEEE International Conference on Automation and Logistics, 169–176.

Smith, A. D., Plowman, D. A., Duchon, D., Quinn, A. M., 2009. A qualitative study of high-reputation plant managers: Political skill and successful outcomes. Journal of Operations Management 27 (6), 428–443.

Snyder, H., 2019. Literature review as a research methodology: An overview and guidelines. Journal of Business Research 104, 333–339.

Solow, D., Ning, J., Zhu, J., Cai, Y., 2020. Improved heuristics for finding balanced teams. IISE Transactions 52 (12), 1312–1323.

Song, H., Tucker, A. L., Murrell, K. L., Vinson, D. R., 2018. Closing the productivity gap: Improving worker productivity through public relative performance feedback and validation of best practices. Management Science 64 (6), 2628–2649.

Staats, B. R., Gino, F., 2012. Specialization and variety in repetitive tasks: Evidence from a Japanese bank. Management Science 58 (6), 1141–1159.

Statista, 2020. Global warehouse automation market size. `https://www.statista.com/statistics/1094202/global-warehouse-automation-market-size/` [03.02.2022].

Statista, 2022. Global warehouse automation market size. `https://www.statista.com/statistics/1094202/global-warehouse-automation-market-size/` [09.08.2022].

Staudt, F. H., Alpan, G., Di Mascolo, M., Rodriguez, C. M. T., 2015. Warehouse performance measurement: A literature review. International Journal of Production Research 53 (18), 5524–5544.

Stein, V., Scholz, T. M., 2019. Manufacturing revolution boosts people issues: The evolutionary need for 'Human–Automation Resource Management' in smart factories. European Management Review 17 (2), 391–406.

Stoltz, M.-H., Giannikas, V., McFarlane, D., Strachan, J., Um, J., Srinivasan, R., 2017. Augmented reality in warehouse operations: Opportunities and barriers. IFAC-PapersOnLine 50 (1), 12979–12984.

Sun, J., Zhang, D. J., Hu, H., van Mieghem, J. A., 2021. Predicting human discretion to adjust algorithmic prescription: A large-scale field experiment in warehouse operations. Management Science 68 (2), 846–865.

Sætre, A. S., Van den Ven, A., 2021. Generating theory by abduction. Academy of Management Review 46 (4), 684—-701.

Tan, T. F., Netessine, S., 2014. When does the devil make work? An empirical study of the impact of workload on worker productivity. Management Science 60 (6), 1574–1593.

Tan, T. F., Netessine, S., 2019. When you work with a superman, will you also fly? An empirical study of the impact of coworkers on performance. Management Science 65 (8), 3495–3517.

Tappia, E., Roy, D., Melacini, M., De Koster, R., 2019. Integrated storage-order picking systems: Technology, performance models, and design insights. European Journal of Operational Research 274 (3), 947–965.

TechCrunch, 2021. Attack of the $200M robotic raises. https://techcrunch.com/2021/09/23/attack-of-the-200m-robotic-raises/ [09.08.2022].

Teigen, K. H., 1994. Yerkes-Dodson: A law for all seasons. Theory & Psychology 4 (4), 525–547.

TGW, 2016. Latest TGW technologies for Zalando. https://www.tgw-group.com/us/news/press-releases/2016/latest-tgw-technologies-for-zalando [25.07.2022].

The Logistics iQ, 2020. Warehouse automation market. https://www.thelogisticsiq.com/research/warehouse-automation-market/ [24.02.2022].

The Logistics iQ, 2022. Warehouse Automation Market. https://www.thelogisticsiq.com/research/warehouse-automation-market/ [03.08.2022].

Tompkins, J. A., White, J., Bozer, Y., Tanchoco, J., 2010. Facilities Planning. John Wiley & Sons, Hoboken, NJ.

Trautrims, A., Grant, D. B., Cunliffe, A. L., Wong, C., 2012. Using the "documentary method" to analyse qualitative data in logistics research. International Journal of Physical Distribution & Logistics Management 42 (8/9), 828–842.

Trebilcock, B., 2020. Robotic software takes control. https://www.mmh.com/article/robotic_software_takes_control [09.08.2022].

Udenio, M., Vatamidou, E., Fransoo, J. C., Dellaert, N., 2017. Behavioral causes of the bullwhip effect: An analysis using linear control theory. IISE Transactions 49 (10), 980–1000.

Urda, J., Loch, C. H., 2013. Social preferences and emotions as regulators of behavior in processes. Journal of Operations Management 31 (1-2), 6–23.

Usher, J., Kamal, A., Kim, S., 2001. A decision support system for economic justification of material handling investments. Computers & Industrial Engineering 39 (1-2), 35–47.

Van Aken, J., Chandrasekaran, A., Halman, J., 2016. Conducting and publishing design science research. Journal of Operations Management 47-48 (1), 1–8.

Van den Ven, A., 2007. Engaged scholarship: A guide for organizational and social research. Press on demand. Oxford University, Oxford, England.

Van der Gaast, J. P., Weidinger, F., 2022. A deep learning approach for the selection of an order picking system. European Journal of Operational Research 302 (2), 530–543.

Van Gils, T., Caris, Ramaekers, K., Braekers, K., De Koster, R. B., 2019. Designing efficient order picking systems: The effect of real-life features on the relationship among planning problems. Transportation Research Part E: Logistics and Transportation Review 125, 47–73.

Van Gils, T., Ramaekers, K., Caris, A., De Koster, R., 2018. Designing efficient order picking systems by combining planning problems: State-of-the-art classification and review. European Journal of Operational Research 267 (1), 1–15.

Van Lent, M., Souverijn, M., 2020. Goal setting and raising the bar: A field experiment. Journal of Behavioral and Experimental Economics 87, 101570.

Van Mieghem, J. A., 2013. OM Forum—Three Rs of Operations Management: Research, relevance, and rewards. Manufacturing & Service Operations Management 15 (1), 2–5.

Vanheusden, S., van Gils, T., Ramaekers, K., Cornelissens, T., Caris, 2022. Practical factors in order picking planning: State-of-the-art classification and review. International Journal of Production Research, 1–25.

Vischer, J. C., 2007. The effects of the physical environment on job performance: Towards a theoretical model of workspace stress. Stress and Health 23 (3), 175–184.

Wang, J., Zhou, Y.-P., 2018. Impact of queue configuration on service time: Evidence from a supermarket. Management Science 64 (7), 3055–3075.

Wang, Z., Sheu, J.-B., Teo, C.-P., Xue, G., 2021. Robot scheduling for mobile–rack warehouses: Human–Robot coordinated order picking systems. Production and Operations Management 31 (1), 95–116.

Webster, J., Watson, R. T., 2002. Analyzing the past to prepare for the future: Writing a literature review. MIS Quarterly 26 (2), xiii–xxiii.

Weidinger, F., Boysen, N., 2018a. Scattered storage: How to distribute stock keeping units all around a mixed-shelves warehouse. Transportation Science 52 (6), 1412–1427.

Weidinger, F., Boysen, N., 2018b. Storage assignment with rack-moving mobile robots in KIVA warehouses. Transportation Science 52 (6), 1479–1495.

WEKA, 2020. Materialfluss Martkübersichten. `https://www.materialfluss.de/materialfluss-marktuebersichten.htm` [01.08.2022].

Wu, Z., Choi, T. Y., 2005. Supplier-supplier relationships in the buyer-supplier triad: Building theories from eight case studies. Journal of Operations Management 24 (1), 27–52.

Wu, Z., Pullman, M. E., 2015. Cultural embeddedness in supply networks. Journal of Operations Management 37 (1), 45–58.

Xie, L., Thieme, N., Krenzler, R., Li, H., 2021. Introducing split orders and optimizing operational policies in robotic mobile fulfillment systems. European Journal of Operational Research 288 (1), 80–97.

Yang, H., Kumara, S., Bukkapatnam, S. T., Tsung, F., 2019. The internet of things for smart manufacturing: A review. IISE Transactions 51 (11), 1190–1216.

Yoon, C. S., Sharp, G. P., 1996. A structured procedure for analysis and design of order pick systems. IIE Transactions 28 (5), 379–389.

Yuan, R., Graves, S. C., Cezik, T., 2019. Velocity-based storage assignment in semi-automated storage systems. Production and Operations Management 28 (2), 354–373.

Zaerpour, N., Yu, Y., de Koster, R. B., 2017. Optimal two-class-based storage in a live-cube compact storage system. IISE Transactions 49 (7), 653–668.

Zhang, P., 2008. Motivational affordances: Reasons for ICT design and use. Communications of the ACM 51 (11), 145–147.