

Quantifying heterogeneity in human behavior: An empirical analysis of forklift operations through multilevel modeling

D. Loske^{1,2} · M. Klumpp^{1,2,3}

Received: 25 May 2021 / Accepted: 6 October 2021 / Published online: 6 April 2022
© The Author(s) 2022 This article is published with Open Access at www.bvl.de/lore

ABSTRACT

Human operators will remain to play an essential role in picker-to-parts order picking systems despite increasing digitalization and automation of warehouse processes. While manual order picking is a laborious and cost-intensive task in warehousing, it is extensively examined in the logistics and supply chain management literature. However, the operational and individual performance of forklift operators in warehouse picking operations has received little attention yet. We aspire to close this gap by drawing on sociotechnical systems theory and formulate a multilevel approach to evaluate heterogeneity in human behavior towards differences in picking performance. We use batch execution times as the dependent variable and source level of operation, target level of operation, filling level of the palette, the necessity to correct replenishment quantities, as well as travel distance as independent variables on the first level. For the second level, we utilize forklift operators to quantify whether heterogeneity in human behavior is impacting the performance of forklift operators. We find that 15.1% of the variance among batch execution times results from heterogeneity in human behavior. In a further simulation, we show that this method can be used to assess the performance of order pickers through a multi-dimensional parametric production frontier analysis. Our findings are highly relevant for logistics management when aspiring to forecast the necessary capacity of forklift operators in a warehouse or building bonus systems that are based on more than the existing two-dimensional measures such as process time per operation.

KEYWORDS: Human factor · behavioral operations management · forklift operators · multilevel modeling



Dominic Loske^{1,2} (Corresponding author)
dominic.loske@wiwi.uni-goettingen.de,
+49 6222 574 530

Matthias Klumpp^{1,2,3}

¹ Georg-August-University of Göttingen, Platz der Göttinger Sieben 3, 37073 Göttingen, Germany

² FOM University of Applied Sciences Essen, Leimkugelstr. 6, 45141 Essen, Germany

³ Fraunhofer IML, J.-v.-Fraunhofer-Str. 2-4, 44227 Dortmund, Germany

1. INTRODUCTION

Intralogistics processes are characterized by ongoing discussions regarding further automation and digitalization [1–4], which include new robotics technologies [5], e.g., robotic sorting systems [6], autonomous mobile robots [7], or automated guided vehicles [8]. However, it is becoming increasingly obvious that the human factor is a highly relevant input factor for the performance of logistics systems, for example, in warehousing or order picking [9–11]. Despite this increasing relevance, analytical approaches and models, as well as empirical data on the performance impact of heterogeneity in human behavior in intralogistics operations, are very scarce. Three explanatory approaches might explain this existing gap: First, digital performance documentation has only recently become available to track individual intralogistics movements of forklifts with individual employees. Second, union and personal influence, as well as data protection regulations, might have prohibited a thorough investigation of personal performance inputs and developments. Third, management and leadership, as well as logistics research attention, might have been occupied with other topical fields such as technological development and large-scale logistics operations optimization

from transport to location planning. Nevertheless, the identified research gap is highly relevant because great effort is implemented to attract suitable warehouse and intralogistics personnel, but detailed knowledge about which individual competencies are actually in high demand and highly efficient for intralogistics processes is lacking. It is an interesting objective for logistics research how the heterogeneity in human behavior can be measured for their contribution to overall efficiency and performance.

Therefore, we aspire to answer the following research questions to address this identified gap: (1) What are the relevant factors that significantly impact the batch execution times of forklift operators? (2) How and to what extent are differences among the batch execution times of forklift operators quantifiable and traceable to the heterogeneity in human behavior? The relevance of these questions becomes obvious when considering the role of human factors in manual picker-to-parts order picking systems [12, 13]. From a theoretical viewpoint, we draw on sociotechnical systems theory [14], a sub-field of general systems theory [15], and are especially interested in the behavioral aspects of human-technology collaboration [16].

This paper is structured as follows: The literature framework in Section 2 positions this study in the area of manual picker-to-parts order picking systems, highlights the theoretical background of this study, and presents existing research in behavioral operations and management (BeOM) and human-centric analyses on forklift operation. Finally, parametric models that measure the efficiency of human operators in warehouses are presented. In Section 3, we describe the data and sample used in this study and introduce the multilevel model together with our strategy for validity and reliability. In the empirical results in Section 4, we discuss the impact of heterogeneity in human behavior and develop the final model in hierarchical steps. Based on the β -values of our final model, we formulate a multi-dimensional production frontier to evaluate operators' performances in Section 5. Finally, managerial learning, limitations, future research direction, and a summary are presented in Section 6.

2. THEORETICAL FRAMEWORK

2.1. Forklift operations in manual picker-to-parts order picking systems

Warehouses play an essential role in most of the existing retail supply chains [14], and they impact the retailer's business success to a large extent because they control product flows to meet customer demands [17]. Based on current figures on the global share of retail sales, e-retail sales accounted for 18% in 2020 [18]. Despite impressive growth rates in e-commerce, stationary retailing still accounts for the majority of global retail sales. However, as recent warehousing

literature focuses on e-commerce, this proportion is not mirrored in recent scientific discussions [19]. Moreover, the underlying circumstances between warehousing in e-retailing and brick-and-mortar retailing are different. E-commerce warehouses mostly face low-volume-high-mix orders, as private households order few pieces from a large assortment provided by the online retailer [19]. An example is the average order demand at Amazon warehouses in Germany, which amounts to 1.6 pieces per order [20]. In contrast, the distribution centers of brick-and-mortar retailers process high-volume-low-mix orders. Thus, we position our paper in the area of warehousing for traditional brick-and-mortar retailing. Because a detailed review of existing warehousing literature is beyond the scope of this paper, the reader is referred to [21–25] for systematic reviews with a general view on warehousing, [26–28] for reviews on performance measurement in warehousing, [29–31] for robotization and automation in warehousing, [19] for e-commerce warehousing and [32] for e-fulfillment in multi-channel distribution.

Main operations in brick-and-mortar retail logistics include receiving, transferring, and storing incoming items from suppliers for order picking and shipping to grocery stores [33]. Regarding the existing warehouse system, different order characteristics, including order size, assortment, workload variation, load stability, store-specific build-up, product expiry, and lead time mainly determine what is applicable for a given retail setting [20, 34]. Thus, suitable order picking systems may range from fully-automated case picking systems to fully manual picker-to-parts order picking systems. However, most of the existing warehouses in western Europe still adopt traditional picker-to-parts setups in their operations [35] due to the following facts: First, manual picker-to-parts order picking systems are highly flexible and scalable, as further order picking capacity can be generated by integrating additional order pickers and forklifts into the work system. Second, large assortments in retailing lead to high variability of products and special requirements for grasping and stacking the items (for example, fragility), which makes the picking task difficult to automate. In summary, our paper belongs to the research stream of manual picker-to-parts order picking systems where pickers travel with vehicle support.

Within the warehouses of traditional brick-and-mortar retailing, order picking is a necessary task, as incoming items from suppliers are usually received and stored in the form of large unit loads, for example, pallets. Customer orders originating from grocery stores are organized in fixed delivery cycles [36], and they comprise different items that mirror the assortments offered in the store while the volume is limited to the stores' shelf capacity [37]. Meanwhile, a process that retrieves items from the large unit loads in response to a specific customer request is required, which is defined as order picking [38]. While picking operations are mostly done on ground level, the upper

shelf levels are the reserve space used for unit loads, and they represent the inventory of a picking place. Forklifts are utilized to store unit loads in the reserve levels after receiving them from the supplier, and they retrieve the unit loads from the reserve levels to the ground level to replenish the picking places [20]. In summary, we are especially interested in the operators steering forklifts in manual picker-to-parts order picking systems with vehicle support in the context of traditional brick-and-mortar retailing. After specifying the object of research that we aim to empirically investigate in this study, we now establish the theoretical background.

2.2. Systems theory, sociotechnical systems and human-system interaction

Triggered by the introduction of computational technology in the mid-20th century, scholars and practitioners have been increasingly confronted with large-scale problems. The goal of integrating similarities within science, promoting communication across disciplines, and establishing a common theoretical basis for general scientific discussions and education have led to system thinking grounded in system science [15]. General systems theory enables the abstraction of large-scale problems by simplifying them on a common theoretical basis while simultaneously capturing the multi-dimensionality of the underlying problem [39]. The theoretical systems view has been applied to various circumstances where humans interact with other systems existing in reality, including socio-ecological systems interlinking human (social) and nature (ecological) [40] or socio-political systems [41]. As a first step, we position our research in systems theory.

In the 1950s, the transformation of working conditions led to hardly predictable work environments where interactions between humans and technology had to be considered [42]. The first set of principles for a sociotechnical systems design was proposed by Cherns in 1976 [43], which interlinked humans (social) and technology (machines). Since then, sociotechnical systems thinking has often been applied to human-machine work systems [14]. It has also been used to explain more modern phenomena such as human-computer systems [44] and has been extended to cyber-sociotechnical systems interlinking intelligent software (cyber) and humans (social) as well as machines (technical) [45]. In summary, sociotechnical systems theory is the theoretical basis that deals with the interaction of humans with any kind of engineered system [46]. However, an interaction such as a reciprocal action or influence is not trivial and may differ depending on several dimensions, such as workspace, work time, or the overall aim of the work system. The taxonomy applied by [16] systemizes these interactions as follows: human-system coexistence sharing work time and workspace; human-system cooperation sharing work time, workspace and aims; and human-system collaboration sharing work time,

workspace, aims, and constantly in contact. In our case of forklift operation in manual picker-to-parts order picking systems, the human operator and machine share work time and workspace, and they both aim to fulfill retrieval operations to supply picking places at the ground level and fulfill storage operations from the ground level to the reserve level.

Additionally, they are constantly in contact because the operator sits in the forklift and utilizes the steering wheel to maneuver and control the machine parts to lift and lower forks. Hence, we are especially interested in the human-technology interaction of forklift operators.

2.3. Human-technology collaboration in forklift operation through the lens of human factors

As a scientific discipline, the human factor (synonymous with the term “ergonomics”) is concerned with the investigation of the interaction between humans and systems, or elements of a system, to optimize human well-being and overall system performance [47]. The theoretical basis is systems thinking and systems theory explained in the previous section, together with the claim that humans are part of operations and production systems and that they need to be integrated.

Calls for more human-centered research in these systems by [48], [49], [50], and more recently [51], point to the importance of human factors. The studies argued that deterministic models neglect the human factor by treating humans and technology, either implicitly or explicitly, as independent from each other. In recent reviews [9, 12, 13, 35], the human factor was divided into perceptual aspects (information processing, reading), mental aspects (learning, forgetting, behavior), physical aspects (workload, lifting, carrying), and psychosocial aspects (motivation, stress, feedback, monotony). However, the impact of these aspects on the forklift operator’s task is difficult to assess on a general level as mental and physical reactions to a perception depend on the respective tasks. Meanwhile, [46] argued that the application of the perception-cognition-motor-action cycle framework is relevant for all human-system interactions.

The cycle includes (1) gathering information through the human sensory system, (2) proceeding with the sensory input cognitively, (3) planning the action of the musculoskeletal system, and (4) responding with an action [52]. This action leads to a system output that is again perceived by the human operator. An existing schema in the human brain is then reinforced, extended, or newly built. This perception-action cycle is grounded on the circular cybernetic flow of cognitive information, a basic biological principle that links humans to their environment [53]. Table 1 systemizes the process steps of forklift operation and the relevant aspects of human factors for a detailed view of human-technology collaboration in forklift operation.

Table 1.: Forklift sub-processes through the lens of human factors

| Operator tasks | Perceptual | Mental | Physical | Psychosocial |
|-----------------------|---|--|---|---|
| Setup | Perceiving the setup of the operation, e.g., batch details on a touch display. | Receive and understand the setup, e.g., batch details on a touch display. | Movements when setting up the workstation, e.g., bending or neck flexion extension during a departure check. | |
| Travel | Perceive warehouse layout, including the visual record and orientation in aisles, shelves, and places. | Understand and remember a route when starting at a source and traveling to a target location. | Steer the forklift through the warehouse, including, e.g., arm movement while sitting or standing in front of a steering wheel. | Aspects regarding motivation, stress, workload, boredom, and interaction with co-workers within all operator tasks, including the setup, travel, search, as well as store/retrieve tasks. |
| Search | Perceive arrangement of places with visual record and orientation with the place logic of the warehouse | Search and identify locations and items, e.g., load unit in target level. | Movements during the search process, including neck flexion extension when, e.g., searching the environment. | |
| Retrieve/store | Perceive height of racks with visual record and orientation when operating within several rack levels. | Decide on how to handle sub-processes, e.g., how to position the forklift and how high to lift forks during a retrieval operation. | Movements during the search process, including neck flexion extension when, e.g., bending for optimal view on forks. | |

Note: Examples in brackets. Initially introduced by [35] for order picking.

2.4. Research on human behavior and the human factor in forklift operation

Research on human behavior in operations systems aims to contribute to the field of BeOM, which has become an accepted sub-field of operations management [54, 55] and is relevant for supporting managers in decision making, designing and optimizing employees' working conditions, as well as improving processes for customers [56]. While BeOM starts at the micro-level to facilitate a better understanding of the heterogeneity in human behavior, the research approaches associated with this stream analyze decisions, skills, and capabilities of individuals or small groups [57, 58]. This may include a cognitive perspective, for example, individual risk aversion in the context of the newsvendor problem [59], individual cognitive reflection in the same context [60], individual perception [61], social preferences [62], or

trust in technology [63–65]. The physical perspective is rarely addressed by BeOM, but it is extensively addressed by the more engineering-driven human factor stream. Although forklift operators are less studied than order pickers, there is still a considerable body of research dealing with them in the context of human-factor analysis. [66] examined cold-related symptoms among forklift drivers operating in a cool warehouse in Thailand. Further approaches that deal with the physical factor explore neck pain [67], neck and backloads [68], energy consumption [69], the impact of whole-body vibration [70], or low back pain [71]. Additionally, a well-examined sub-process of forklift operation is the driving process itself, which is the reason many studies investigate collision warning systems [72, 73], safety systems [74], and sensor-based assistant systems [75]. Regarding mental processes, fatigue and learning are

well examined. [76, 77] examined fatigue associated with the sub-process of driving. An approach that contributes to this direction is exploring the situation awareness of forklift operators for preventing collisions [78]. Hence, we identify a research gap regarding the interplay of perceptual, mental, and physical aspects when humans operate forklifts within a human–technology setting. This may include the fact that some individuals are more efficient in combining task sets than others. As humans are an essential part of these work systems, they significantly impact the outcomes of warehouse systems, including performance and quality. In the next section, we will focus on methodologies to operationalise, evaluate and benchmark human-operator performance in warehousing.

2.5. Parametric models for performance measurement of human operators in warehousing

When multilevel modeling is applied to warehousing, the basic idea is to measure the performance of order picks and decompose large datasets that contain order batches. However, it is often difficult to determine how and to what extent human capabilities and skills cause performance differences. In one approach, [10] applied multilevel modeling to order picking logs and proved that almost 10.3% of the total time variance in batch execution time could be traced to individual differences. Another approach was presented by [79] using pick duration as the dependent variable. They analyzed the impact of the independent variables through a survival model, more precisely a parametric accelerated-failure-time. More parametric models have been applied to measure fatigue and learn order pickers [80, 81], using Wright’s, Sandford-B, or DeJong’s learning curve. Most of these approaches use cumulated picks and investigate the development of the time per pick while concluding that minimizing the pick time is equal to a learning effect where a human operator improves picking skills.

3. DATA AND METHODOLOGY

3.1. Warehouse and dataset description

We analyzed the performance of forklift operators employed in the distribution center of a large German grocery retail chain. Within this distribution center, refrigerated and non-refrigerated foods are stored, picked, packed, and delivered to supermarkets. This

study focuses on one warehouse storing perishable and non-perishable non-cooled items. In our sample warehouse, orders are picked through a manual picker-to-parts order picking system, and storage locations used for picking are replenished by forklift operators. The warehouse has two cross-aisles, 28 aisles on the ground level, one depot storing empty roll cages, and drop-off points for the pickers at every aisle. Since the racking system comprises several gripping levels (minimum, ground level; maximum, 1.05 m) and five reserve levels (level 1, 2.10 m; level 2, 4.20 m; level 3, 6.30 m; level 4, 8.40 m; level 5, 10.5 m), the storage of pallets in the higher levels is particularly a challenge. The maximum height for the delivered pallets (including pallet timber) is 1.05 m (CCG1) or 1.95 m (CCG2). Altogether, storing or retrieving a 1.95-m CCG2 pallet in a storage place with a height of 1.95 m on the highest rack level (10.5 m) is a challenging task. This is why we expect individual heterogeneity and behavioral factor to be relevant for the batch execution time. Many warehouse management systems (WMS) store replenishment and order picking logs, which quantify operations on a detailed level. The batch execution times for individual forklift operators may depend on warehouse-specific circumstances, for example, a varying amount of order pickers operating in the same warehouse, which depends on work time and workgroup decisions. However, as human capabilities, skills, and heterogeneity in human behavior have been identified as relevant for intralogistics [12, 80], behavioral factors may also influence batch execution times.

We included these behavioral factors implicitly through the available data by considering and evaluating the past performance of the forklift operators. One batch is equal to a storage or retrieval operation of one single full pallet. The warehouse management data obtained for this study on batch execution times for replenishment operations of forklift operators included: (1) batch ID, (2) forklift operator ID, (3) starting date of the batch, (4) ending date of the batch, (5) starting time of batch, (6) ending time of batch, (7) distance traveled by a forklift as a difference between the source and target locations, (8) height of the pallet, (9) level of the source location, (10) level of the target location, (11) degree of filling for each pallet as a percentage of the CCG1 or CCG2 type, (12) necessity of correcting the amount on a picking location as errors occur in upstream processes (as a dichotomy variable, 0 = no and 1 = yes).

Table 2.: Descriptive statistics of the observed forklift operators

| Operator ID | Observations | Mean execution time | Standard deviation | Variance |
|-------------|--------------|---------------------|--------------------|----------|
| 1 | 1,198 | 155.63 | 79.74 | 0.51 |
| 2 | 500 | 161.39 | 67.14 | 0.42 |
| 3 | 735 | 114.65 | 67.35 | 0.59 |
| 4 | 908 | 190.81 | 91.19 | 0.48 |
| 5 | 626 | 199.17 | 91.16 | 0.46 |
| 6 | 469 | 176.10 | 99.66 | 0.57 |
| 7 | 920 | 112.53 | 70.55 | 0.63 |
| 8 | 673 | 171.85 | 102.11 | 0.59 |
| 9 | 961 | 118.21 | 67.84 | 0.57 |
| 10 | 446 | 177.35 | 68.66 | 0.39 |
| 11 | 1,022 | 179.73 | 87.82 | 0.49 |
| 12 | 1,357 | 147.11 | 75.45 | 0.51 |
| 13 | 932 | 107.78 | 43.05 | 0.40 |
| 14 | 981 | 207.23 | 89.98 | 0.43 |
| 15 | 1,261 | 193.05 | 111.66 | 0.58 |
| 16 | 1,154 | 111.07 | 48.71 | 0.44 |
| 17 | 467 | 250.46 | 122.42 | 0.49 |
| 18 | 423 | 195.05 | 104.30 | 0.54 |
| 19 | 945 | 157.13 | 66.67 | 0.42 |
| 20 | 537 | 194.59 | 98.72 | 0.51 |
| 21 | 1,167 | 183.64 | 74.00 | 0.40 |
| 22 | 778 | 153.20 | 98.58 | 0.64 |
| 23 | 789 | 153.40 | 87.72 | 0.57 |
| 24 | 431 | 222.00 | 102.23 | 0.46 |
| 25 | 633 | 139.48 | 91.98 | 0.66 |
| 26 | 635 | 210.72 | 96.94 | 0.46 |
| 27 | 706 | 99.92 | 38.57 | 0.39 |
| 28 | 1,129 | 204.17 | 110.12 | 0.54 |
| 29 | 1,275 | 165.98 | 102.15 | 0.62 |
| 30 | 525 | 185.00 | 96.88 | 0.52 |
| 31 | 980 | 169.92 | 75.97 | 0.45 |
| 32 | 707 | 125.07 | 72.03 | 0.58 |
| 33 | 709 | 193.94 | 85.01 | 0.44 |

The whole dataset contained 27,985 batches (equal to full loads on pallets) for replenishment operations performed by 45 forklift operators during a 1.5-month period. All operations were performed between 7 a.m. and 6 p.m., and all the operators used identical technology. As we use real-life data, some datasets were polluted by personal breaks during batches. Therefore, we filtered out all batches lasting longer than 15 mins and all batches without operator ID. Additionally, we excluded zero values and checked for unreasonable traveling speed within the warehouse. After the data cleaning, the dataset containing 26,979 out of 27,985 batches and 33 out of 45 forklift operators were used for further analysis. Table 2 summarizes the descriptive statistics of the operators.

3.2. Multilevel modeling

Linear regression models are used to explain the statistical relationship between a dependent variable y_i and an independent variable x_i . They can be extended to multiple linear regression with more than one independent variable. Therefore, one possible approach is to explain the batch execution time of a forklift operator for batch i as the dependent variable and the travel distance and filling level of a pallet, both on the individual batch level I , as independent variables. The standard ordinary least squares (OLS) regression, which includes the error term ϵ , is given by Eq. (1):

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i \quad (1)$$

Multilevel modeling depicts a series of nested OLS regression analyses in which the coefficients at level 1 become the dependent variable at another level [82]. In

the first level of our model, we tried to explain the batch execution time with the five independent variables used to quantify heterogeneity in human behavior. Table 3 summarizes the dependent and independent variables for the first level of our model to operationalize behavioral factors.

In our multilevel model, we represent the sub-tasks of the individual operator by linear regression model betas. Let $\beta_{i,j}$ be one sub-task within $i \in I$, where i is the number of the examined task, and I is the total set of tasks described in Table 3. Additionally, $j \in J$ defines the forklift operator with an individual operator j and a set of examined operators J . After clarifying the basic idea of our approach based on multiple linear regression, we added a second level. Multilevel modeling is used to additionally distinguish between groups [83–85], which are represented by forklift operators in our case. Herein, we added group variables that enabled us to differentiate between groups and to evaluate whether or not this differentiation is beneficial for the model. Furthermore, we concretized the dependent variable y_{ij} as the batch execution time t for a set of possible batches $r \in R$ from the set of order O that we also used for the independent variables. The hierarchical two-level model was notated as a group-dependent regression, with j reflecting the group level. We used the notation of [10] formulating a multilevel approach for order pickers to forecast the picking time in the first level and grouping according to operator ID:

$$t_{r,w} = \beta_{0,w} + \sum_{i \in I} \beta_{i,w} x_{i,r} + \epsilon_{w,r} \quad (2)$$

$$\beta_{0,w} = y_i + u_{i,w} \quad (3)$$

Table 3.: Operationalization of human behavior and variables for level 1

| Behavioral factor | Role in model | Operationalization |
|--|---------------------------|---------------------------|
| The performance of the forklift operator shall be high when the processing time for storage and replenishment operations is minimized | Dependent | Batch execution time |
| The heterogeneity in the speed of work execution when performing the sub-task of retrieving pallets from the racking system with different levels | Independent (β_1) | Source level of operation |
| The heterogeneity in the speed of work execution when performing the sub-task of storing pallets on the grasping level for the order picker | Independent (β_2) | Target level of operation |
| The heterogeneity heterogeneity in the speed of work execution when performing the sub-task of handling different heights of pallets where the difficulty increases with high filling levels (less space for maneuveringmanoeuvring) | Independent (β_3) | Filling level of a pallet |
| The heterogeneity in the speed of work execution when performing the sub-task dealing with errors and correct them through his mobile computer connected to the WMS | Independent (β_4) | Necessity to correct |
| The heterogeneity in the speed of work execution when performing the sub-task of driving a forklift fast and saving within the assigned warehouse area | Independent (β_5) | Travel distance |

where

- $t_{r,w}$: batch execution time
- r : individual batch (indicated as a level-one unit)
- w : operator ID (indicated as a level-two unit)
- β : regression weight for the independent variable
- w operator ID (indicated as a level-two unit)
- i individual task
- x independent variable
- ε within-group error term
- Y slope effect dependent variable
- u group-level error term

The models were formulated using IBM SPSS Statistics and its multilevel modeling options. As we associate a low batch execution time $t_{r,w}$ with a highly efficient forklift operator, lower β_i are beneficial for low execution times. Therefore, lower β_i values of the individual operators lead to decreasing batch execution times. Consequently, high variations in β_i values are linked to higher heterogeneity in human behavior, and we set the relationship as follows: (1) high variations in β_1 as high heterogeneity in human behavior when retrieving pallets, (2) high variations in β_2 as high heterogeneity in human behavior when storing pallets on the grasping level, (3) high variations in β_3 as high heterogeneity in human behavior when handling different heights of pallets, (4) high variations in β_4 as high heterogeneity in human behavior when dealing with errors, and (5) high variations in β_5 as high heterogeneity in human behavior when driving a forklift within the assigned warehouse area.

To assess the model fit for our multilevel approach, we followed the approach by [86] and calculated R_2 values for levels 1 and 2, as well as for the total model by opposing the results of our initial null-model (baseline) with the models integrating the independent variables.

$$total R^2 = \frac{(\hat{\sigma}_\varepsilon^2 + \hat{\sigma}_{\mu_0}^2)_{baseline} - (\hat{\sigma}_\varepsilon^2 + \hat{\sigma}_{\mu_0}^2)_{comparison}}{(\hat{\sigma}_\varepsilon^2 + \hat{\sigma}_{\mu_0}^2)_{baseline}} \quad (4)$$

$$R_1^2 = \frac{\hat{\sigma}_{\varepsilon,baseline}^2 - \hat{\sigma}_{\varepsilon,comparison}^2}{\hat{\sigma}_{\varepsilon,baseline}^2} \quad (5)$$

$$R_2^2 = \frac{\hat{\sigma}_{\mu,baseline}^2 - \hat{\sigma}_{\mu,comparison}^2}{\hat{\sigma}_{\mu,baseline}^2} \quad (6)$$

After calculating the null model as a basis for further evaluation in Section 4.1, we used the model fits in Eqs. (4)–(6) to assess whether or not the integration of additional dependent variables is beneficial for the total R^2 . This approach is similar to hierarchical regression analysis and enables the evaluation of integrated level one predictors.

4. EMPIRICAL RESULTS

4.1. Differences between operators

In the first step, we calculated a variance components model without predictors $x_{i,r}$, which is well-known as a null model in multilevel modeling [82]. Herein, one fixed effect of the grand mean, in our case, the mean of the batch execution times, was estimated for all batch IDs and operator IDs. The null model is given as follows:

$$t_{r,w} = \beta_{0,w} + \varepsilon_{w,r} \quad (7)$$

$$\beta_{0,w} = \gamma_i + u_{i,w} \quad (8)$$

Based on this, we computed the intraclass correlation coefficient (ICC) to evaluate between-group variations [87–89]. In our case, this measure can explain whether or not variations in batch execution times occur between forklift operators and if it is suitable to apply multilevel modeling. For the mathematical formulation, the ICC is given by [82]:

$$ICC = \frac{\sigma_{\mu_{o,j}}^2}{\sigma_{\varepsilon_{i,j}}^2 + \sigma_{\mu_{o,j}}^2} \quad (9)$$

where

$\sigma_{\mu_{o,j}}$: variance of batch execution time in the second level (operator ID)

$\sigma_{\varepsilon_{i,j}}$: variance of batch execution time in the first level (batch ID)

The ICC for our case is 0.1510, indicating that 15.1% of the variance among the batch execution time originates from the forklift operators on the second level of our multilevel model, and 84.9% can be traced back to the batches themselves. [10] conducted the same analysis for order pickers, and their model showed that 10.3% of the variance in batch execution time was due to the individual differences of order pickers. The results are presented in Table 4, and they confirm that our results are significant at the alpha level of 0.05. As the Wald-Z value of 4.03 is highly significant ($p < .001$), we can reject the null hypothesis that there are no significant differences in batch execution time among the examined forklift operators.

Table 4.: Summary of null-model for intraclass correlation coefficient test

| Parameter | Estimate | Standard error | Wald-Z | p | Lower bound | Upper bound |
|------------------------------|----------|----------------|--------|--------|-------------|-------------|
| $\sigma_{\varepsilon_{i,j}}$ | 7293.55 | 62.84 | 116.07 | .000 | 7171.43 | 7417.75 |
| $\sigma_{\mu_{o,j}}$ | 1297.91 | 322.144 | 4.03 | < .001 | 797.95 | 2111.13 |

Note: The lower and upper bounds are for an alpha level of 0.05.

4.2. A two-level model with fixed predictors and random intercepts

In a second analysis stage, we built on the null model and added travel distance per batch as the first independent variable on the first level. Additionally, we added the operator IDs in the second level. For the initial bi-variate linear regression model with execution time as the dependent variable, we find a high statistical linear relationship with travel distance as the independent variable. This is also illustrated in Figure 1 with linear regression lines per forklift operator.

The individual and different slopes per operator indicate the high heterogeneity in human behavior

observable for the examined forklift operators. Furthermore, we integrated this independent variable, together with the other independent variables introduced in Table 2, into our multilevel approach. For the integration of travel distance, we formulated the level 1 in Equ. (10) and the level 2 in Equ. (11, 12) as follows:

$$t_{r,w} = \beta_{0,w} + \beta_{1,w} \text{travel.distance}_{i,r} + \varepsilon_{w,r} \quad (10)$$

$$\beta_{0,w} = y_i + u_{i,w} \quad (11)$$

$$\beta_{1,w} = \gamma_{10} \quad (12)$$

Fig. 1.: Scatterplot for batch execution time and travel distance per operator

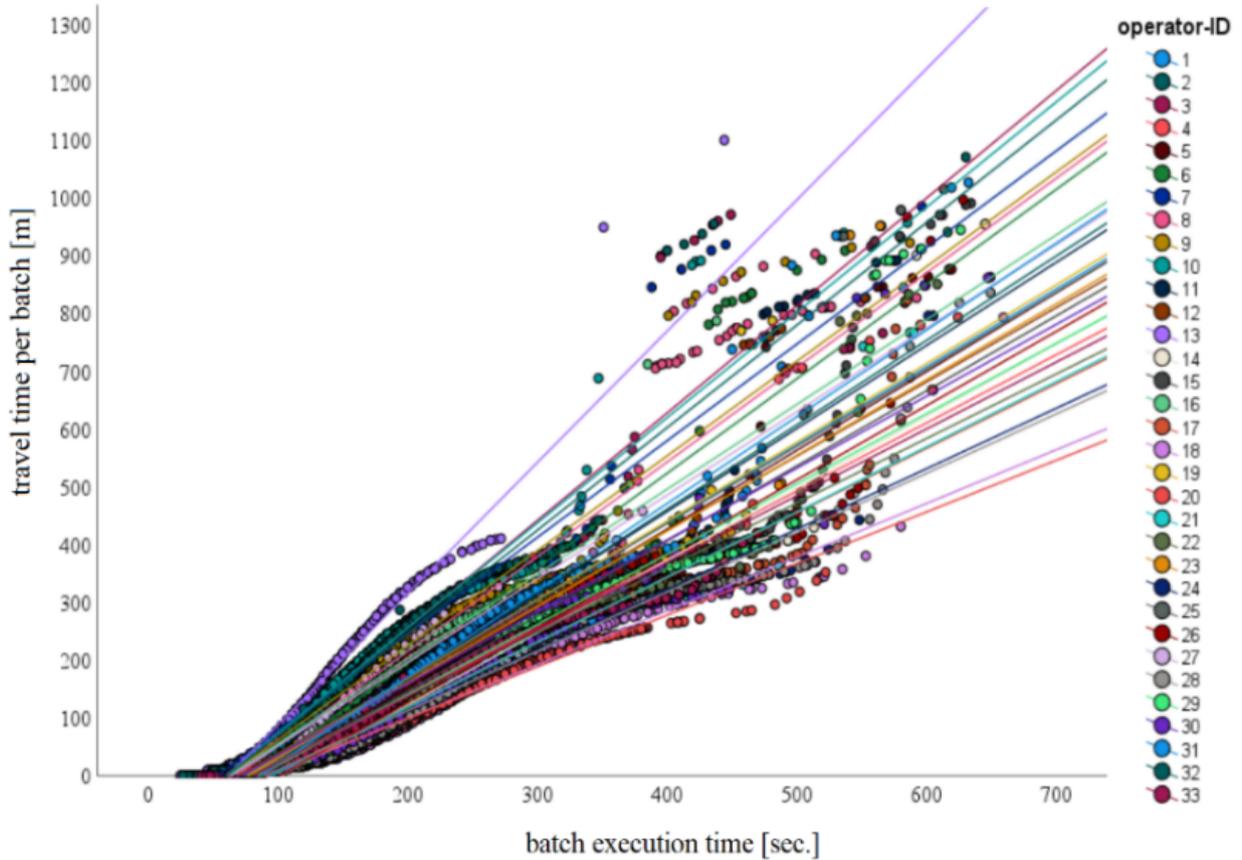


Table 5.: Summary of a multilevel model with fixed predictors and random intercepts

| Model | R^2 | R_1^2 | R_2^2 | Parameter | Estimate | AIC | Wald-Z | p |
|------------|-------|---------|---------|------------------------------|----------|--------|--------|--------|
| Null-model | | | | $\sigma_{\varepsilon_{i,j}}$ | 7293.65 | 316703 | 116.07 | .000 |
| | | | | $\sigma_{\mu_{o,j}}$ | 1297.91 | | 4.03 | < .001 |
| Model 2 | .84 | .92 | .38 | $\sigma_{\varepsilon_{i,j}}$ | 609.42 | 249802 | 116.07 | .000 |
| | | | | $\sigma_{\mu_{o,j}}$ | 800.29 | | 4.06 | .000 |
| Model 3 | .86 | .94 | .41 | $\sigma_{\varepsilon_{i,j}}$ | 402.69 | 238638 | 116.06 | .000 |
| | | | | $\sigma_{\mu_{o,j}}$ | 769.74 | | 4.06 | .000 |
| Model 4 | .86 | .94 | .41 | $\sigma_{\varepsilon_{i,j}}$ | 402.59 | 238634 | 116.07 | .000 |
| | | | | $\sigma_{\mu_{o,j}}$ | 769.73 | | 4.06 | .000 |
| Model 5 | .86 | .94 | .41 | $\sigma_{\varepsilon_{i,j}}$ | 400.78 | 238513 | 116.07 | .000 |
| | | | | $\sigma_{\mu_{o,j}}$ | 764.68 | | 4.06 | .000 |
| Model 6 | .86 | .94 | .41 | $\sigma_{\varepsilon_{i,j}}$ | 400.63 | 238505 | 116.07 | .000 |
| | | | | $\sigma_{\mu_{o,j}}$ | 764.25 | | 4.06 | .000 |

Note: AIC is used as an estimator to predict error rates in the model.

Compared to the initial null-model, our model 2 (null-model + travel time as dependent variable) showed an improved Akaike information criterion (AIC), indicating a reduction of the error rate in our model. We added the additional independent variables stepwise until we reached the final model described in Eqs. (2) and (3) with model 3 (model 2 + the filling level of the pallet), model 4 (model 3 + source level of operation), model 5 (model 4 + target level of operation) and model 6 (model 5 + necessity to correct). Table 5 summarizes the results.

The results indicate that the level 1 prediction power of the final model (model 6) increases by 94% when integrating the variables described above. This finding on predicting the batch execution time is not surprising, and it still proves that it can be forecasted with the given

independent variables. Additionally, the integration of operator IDs as a second level increases the null model up to 41%. This again confirms that heterogeneity in human behavior is a critical factor. In total, our final model increased the predicting power up to 86% compared to the initial null model. Considering the model betas, the results for the source level of operation (β_1) are not significant and must be excluded from our analysis. Therefore, our final model comprises target level of operation (β_2), filling level of the pallet (β_3), the necessity to correct (β_4), and travel distance (β_5). In summary, our findings show that heterogeneity in human behavior plays a major role when forecasting batch execution times with these variables. Table 6 summarizes all model betas for a multilevel model with fixed predictors and random intercepts.

Table 6.: Model betas for a multilevel model with fixed predictors and random intercepts

| Parameter | β | Standard error | df | p | Lower bound | Upper bound |
|-----------------------------|---------|----------------|-------|------|-------------|-------------|
| Constant | 50.93 | 4.835 | 33 | .000 | 41.10 | 60.76 |
| Source level of operation | -.01 | .0110 | 26949 | .484 | .029 | .014 |
| Target level of operation | 1.32 | .155 | 26948 | .000 | 1.016 | 1.62 |
| Filling level of the pallet | 119.80 | 1.031 | 26951 | .000 | 117.79 | 121.83 |
| Necessity to correct | 1.22 | .383 | 26947 | .001 | .47 | 1.97 |
| Travel distance | .48 | .002 | 26949 | .000 | .48 | .49 |

Note: Degrees of freedom (df) is the number of values in the final model that are free to vary.

5. SIMULATION OF BATCH EXECUTION TIME FOR FORKLIFT OPERATORS

After formulating the multilevel approach and applying it to the real data of 33 forklift operators, we are now able to formulate a multiple regression model with the estimates calculated in the previous section. At this point, it is important to note that the source level of operation (β_i) is excluded in the simulation approach. From a managerial perspective, there are two possible tracks for simulation. First, by integrating level 1 and level 2 formulations in Eqs. (25) and (26), it is possible to forecast the batch execution times for forklift

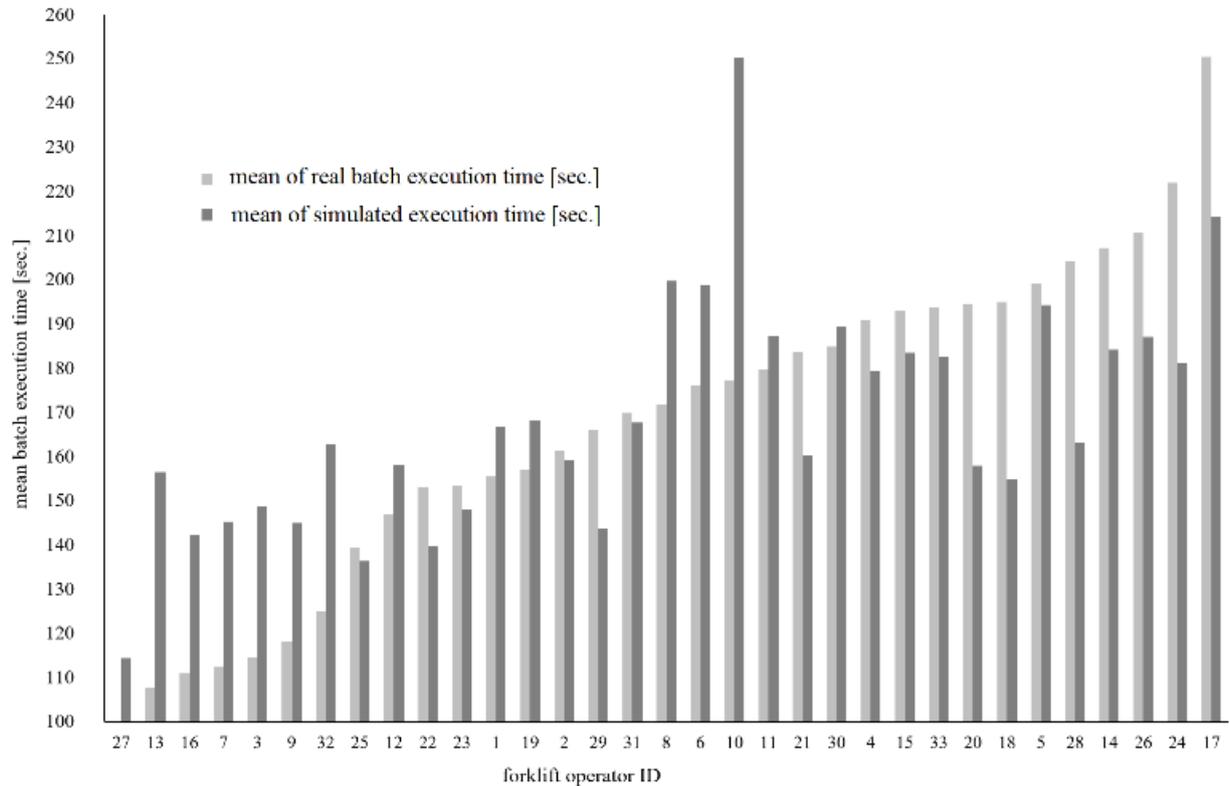
operators and the total working time for all batches. This may be beneficial when logistics managers seek to find the optimal capacity of forklifts and forklift operators for their warehouse. Second, after being aware that individual heterogeneity in human behavior has an impact on the performance of forklift operators, we adopt the level 1 regression model to forecast the batches. Hence, we can formulate a multi-dimensional and parametric production function that is beneficial to evaluate a forklift driver's performance compared to the group performance level. The formulation of our simulation approach is as follows:

$$t_{r,w} = 50.93 + 1.32 x_{2,r} + 119.8 x_{3,r} + 1.44 x_{4,r} + 0.48 x_{5,r} + \varepsilon_{w,r} \quad (13)$$

$$\beta_{0,w} = \gamma_i + u_{i,w} \quad (14)$$

$$\beta_{1,w} = \gamma_{10} \quad (15)$$

Fig. 2.: Comparison of simulation results and real batch execution times for forklift operators



After running the simulation, we found significant differences among the individual forklift operators, which is as expected from our previous findings and can be used for cross-validation purposes. As bonus schemes in intralogistics mostly rely on process time, for example, the average time per operation, a multi-dimensional parametric approach may show a promising way to integrate further performance-relevant factors. Additionally, our approach integrates time for correcting errors that result from upstream processes and are, therefore, not caused by forklift operators. However, these situations are often ignored when benchmarking blue-collar workers in instore logistics.

6. CONCLUSION

6.1. Management learnings

This paper has outlined an innovative method and empirical evaluation approach regarding the analysis of efficiency contributions from human operators with forklift operators in retail warehousing operations. A relevant explanation share of 15.1% regarding the individual batch operating time data was identified, warranting further research and also management attention towards the analysis of individual heterogeneity. Our model is transferable to all warehouses where forklift operations are performed by humans. This may include picker-to-parts or parts-to-picker order picking systems. The main requirement for transferring our formulated model to practice is the availability of historical data, e.g., documented by a WMS. Additionally, our method requires data on the aggregation level of operations. Because most modern warehouses utilize digitalized or intelligent picking systems, e.g., pick-by-voice or pick-by-vision, we expect that the data used in this paper is also available to a majority of warehouse managers. To further optimize warehouse processes, our findings may help managers to rethink batch assignment methods in human-centered work systems and reorganize work system design strategies.

First, taking the aspect of batch assignment methods, most companies use simple rules, such as the first free method. We expect a high application potential for batch assignment respecting the individual heterogeneity of the human workforce. After formulating the model, operationalizing the constructs, and inserting the data into SPSS or R, the software output provides model betas per operator. Through a comparative analysis of all operators, we are able to cluster humans according to their performance, e.g., operators that are better in driving the forklift than others. At the same time, we can cluster all batches in the WMS according to, e.g., distance. Bringing both together, we may utilize the model betas from our multilevel approach to implement human-centered batch assignment for

forklift operations. This would also allow improving training processes for new and unknown employees: Let us assume that operator A has an experience level of 10,000 cumulative operations after a time period of three months and is seen as an experienced employee. Our model and the historic WMS data can now help to estimate the model betas in the very first week quantifying a low level of experience. These model betas can then be assigned to new and unknown operator B. In this exemplary case, the human-centered batch assignment may be suitable to prevent overload and resulting stress when incorporating new employees.

Second, data-driven multilevel analyses might be beneficial to develop human-centered work system design strategies. The presented parametric multilevel models have the potential to shed light on operations with human involvement at a very detailed level. Finding that 15.1% of the total variance in time is related to differences between human operators might, e.g., lead to incentive-based sociotechnical work systems taking behavioral factors more into account. It underlines the importance of integrating ergonomic aspects for work system design. Additionally, our method allows the evaluation of operators' performance on sub-tasks, e.g., driving the forklift. This may be helpful to derive whether and to which extent assistive technologies may be necessary for certain tasks, sub-tasks, operators, or batches. For an extensive review of assistive devices in manual material handling, the reader is referred to [90]. In summary, quantifications of individual heterogeneity in operations systems might trigger managers to rethink work system design strategies.

6.2. Limitations and further research

The approach presented in this paper has specific limitations. First, although the obtained data basis of 26,979 batches remaining after the data cleaning is extensive, the quantifications are limited to the examined warehouse. Therefore, our model and the approach would benefit from further applications in alternative warehouse scenarios. Additionally, we investigate humans operating in a warehouse storing perishable and non-perishable non-cooled items. It may, therefore, be interesting to research the impact of temperature on human performance and heterogeneity in cold supply chains, as well as the performance in night shifts compared to those in day shifts.

From a methodological viewpoint, further research avenues can be directed at multilevel models for panel data. This could enable the quantification of the learning effects and the development of human skills at a certain time. Furthermore, non-parametric methodologies need to be tested to build production frontiers without setting regression weights a priori. Data envelopment analysis and free disposable hull methodologies are suitable as they apply a model-endogenous weighting of input and output factors. The importance of research analyses and concepts regarding the human-technology interaction parts of operations and logistics systems will further

increase due to future digitalization and automation developments.

From a content-oriented viewpoint, logistics managers may be interested in the aspect of how and to which extent technological improvement through digitalization and automation can accelerate operations. Further research utilizing the same historic WMS data as in our approach may, therefore, be directed in the area of parametric survival models. The event-history analysis is concerned with duration times where an event is defined as the transition from the origin to the destination state [91]. This time frame is quantified as a survival object and integrated as the dependent variable in the survival model. Thus, we propose to set the process time of interest as the survival object and the dependent variable within an accelerated failure time model. Based on this, independent variables can then be defined, allowing to quantify their accelerating or decelerating impact.

6.3. Summary and Outlook

The contribution of this paper is threefold: (1) it elaborates, justifies, and applies a parametric multilevel approach as a novel method for evaluating forklift operators, (2) it decomposes batch execution time variations in two levels to reveal the impact of heterogeneity in human behavior, (3) it elaborates and applies a simulation approach based on a parametric multi-dimensional production frontier to evaluate operators' performance with empirical data. Therefore, this paper provides the first step and is a cornerstone for future analyses regarding the efficiency impact of the human factor. This is connected to ongoing research regarding the human factor in logistics operations [19, 30, 46, 81, 92–96]. The common ground of this research stream is the empirical evidence pointing at a highly relevant efficiency and effectiveness impact of human factors towards logistics operations, as also shown empirically in this contribution. This warrants further research and analysis of the relevant impact of logistics research.

CONFLICT OF INTEREST

The authors declare that they have no conflict of interest.

REFERENCES

1. Kong XT, Zhong RY, Zhao Z, Shao S, Li M, Lin P, Chen Y, Wu W, Shen L, Yu Y, Huang GQ (2020) Cyber physical ecommerce logistics system: An implementation case in Hong Kong. An implementation case in Hong Kong. *Computers & Industrial Engineering* 139:106170. doi:10.1016/j.cie.2019.106170
2. Mörth O, Emmanouilidis C, Hafner N, Schadler M (2020) Cyber-physical systems for performance monitoring in production intralogistics. *Computers & Industrial Engineering* 142:106333. doi:10.1016/j.cie.2020.106333
3. Pan S, Trentesaux D, McFarlane D, Montreuil B, Ballot E, Huang GQ (2021) Digital interoperability and transformation in logistics and supply chain management. Editorial. *Computers in Industry* 129:103462. doi:10.1016/j.compind.2021.103462
4. Coelho F, Relvas S, Barbosa-Póvoa AP (2021) Simulation-based decision support tool for in-house logistics. The basis for a digital twin. *Computers & Industrial Engineering* 153:107094. doi:10.1016/j.cie.2020.107094
5. Ghelichi Z, Kilaru S (2021) Analytical models for collaborative autonomous mobile robot solutions in fulfillment centers. *Applied Mathematical Modelling* 91:438–457. doi:10.1016/j.apm.2020.09.059
6. Zou B, Koster R de, Gong Y, Xu X, Shen G (2021) Robotic Sorting Systems. Performance Estimation and Operating Policies Analysis. *Transportation Science*. doi:10.1287/trsc.2021.1053
7. Žulj I, Salewski H, Goeke D, Schneider M (2021) Order batching and batch sequencing in an AMR-assisted picker-to-parts system. *European Journal of Operational Research*. doi:10.1016/j.ejor.2021.05.033
8. Löffler M, Boysen N, Schneider M (2021) Picker Routing in AGV-Assisted Order Picking Systems. *INFORMS Journal on Computing*. doi:10.1287/ijoc.2021.1060
9. Grosse EH, Glock CH, Neumann WP (2017) Human factors in order picking. A content analysis of the literature. *International Journal of Production Research* 55/5(5):1260–1276. doi:10.1080/00207543.2016.1186296
10. Matusiak M, de Koster RBM, Saarinen J (2017) Utilizing individual picker skills to improve order batching in a warehouse. *European Journal of Operational Research* 263/3(3):888–899. doi:10.1016/j.ejor.2017.05.002
11. Klumpp M, Ruiner C, eds. (2021) *Digital Supply Chains and the Human Factor*. Springer, Berlin, Heidelberg
12. Sgarbossa F, Grosse EH, Neumann WP, Battini D, Glock CH (2020) Human factors in production

- and logistics systems of the future. *Annual Reviews in Control* 49:295–305. doi:10.1016/j.arcontrol.2020.04.007
13. Grosse EH, Calzavara M, Glock CH, Sgarbossa F (2017) Incorporating human factors into decision support models for production and logistics. Current state of research. *IFAC-PapersOnLine* 50/1(1):6900–6905. doi:10.1016/j.ifacol.2017.08.1214
 14. Clegg CW (2000) Sociotechnical principles for system design. *Applied ergonomics* 31/5(5):463–477. doi:10.1016/S0003-6870(00)00009-0
 15. Skyttner L (2001) *General systems theory. Ideas & applications*. World Scientific, Singapore
 16. Schmidtler J, Knott V, Hölzel C, Bengler K (2015) Human Centered Assistance Applications for the working environment of the future. *Occupational Ergonomics* 12/3(3):83–95. doi:10.3233/OER-150226
 17. Tompkins J, White JA, Bozer YA (2010) *Facilities planning*, 4th edn. Wiley, Hoboken, NJ
 18. Statista (2021) E-commerce share of total global retail sales from 2015 to 2023. Accessed 4 May 2021
 19. Boysen N, de Koster RBM, Weidinger F (2019) Warehousing in the e-commerce era. A survey. *European Journal of Operational Research* 277/2(2):396–411. doi:10.1016/j.ejor.2018.08.023
 20. Boysen N, de Koster RBM, Fübler D (2021) The forgotten sons. Warehousing systems for brick-and-mortar retail chains. *European Journal of Operational Research* 288/2(2):361–381. doi:10.1016/j.ejor.2020.04.058
 21. Baker P, Canessa M (2009) Warehouse design. A structured approach. *European Journal of Operational Research* 193/2(2):425–436. doi:10.1016/j.ejor.2007.11.045
 22. Gu J, Goetschalckx M, McGinnis LF (2010) Research on warehouse design and performance evaluation. A comprehensive review. *European Journal of Operational Research* 203/3(3):539–549. doi:10.1016/j.ejor.2009.07.031
 23. Davarzani H, Norrman A (2015) Toward a relevant agenda for warehousing research. Literature review and practitioners' input. *Logist. Res.* 8/1(1). doi:10.1007/s12159-014-0120-1
 24. Gu J, Goetschalckx M, McGinnis LF (2007) Research on warehouse operation. A comprehensive review. *European Journal of Operational Research* 177/1(1):1–21. doi:10.1016/j.ejor.2006.02.025
 25. Rouwenhorst B, Reuter B, Stockrahm V, van Houtum GJ, Mantel RJ, Zijm W (2000) Warehouse design and control. Framework and literature review. *European Journal of Operational Research* 122/3(3):515–533. doi:10.1016/S0377-2217(99)00020-X
 26. van Gils T, Ramaekers K, Caris A, de Koster RBM (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):1–15. doi:10.1016/j.ejor.2017.09.002
 27. Staudt FH, Alpan G, Di Mascolo M, Rodriguez CMT (2015) Warehouse performance measurement. A literature review. *International Journal of Production Research* 53/18(18):5524–5544. doi:10.1080/00207543.2015.1030466
 28. Gong Y, de Koster RBM (2011) A review on stochastic models and analysis of warehouse operations. *Logistics Research* 3/4(4):191–205. doi:10.1007/s12159-011-0057-6
 29. Fragapane G, de Koster RBM, Sgarbossa F, Strandhagen JO (2021) Planning and control of autonomous mobile robots for intralogistics. Literature review and research agenda. *European Journal of Operational Research*. doi:10.1016/j.ejor.2021.01.019
 30. Azadeh K, de Koster RBM, Roy D (2019) Robotized and automated warehouse systems. Review and recent developments. *Transportation Science* 53/4(4):917–945. doi:10.1287/trsc.2018.0873
 31. Lerher T, Ekren YB, Sari Z (2015) Simulation Analysis of Shuttle Based Storage and Retrieval Systems. *International Journal of Simulation Modelling*:48–59. doi:10.2507/IJSIMM14(1)5.281
 32. Agatz NA, Fleischmann M, van Nunen JA (2008) E-fulfillment and multi-channel distribution – A review. *European Journal of Operational Research* 187/2(2):339–356. doi:10.1016/j.ejor.2007.04.024
 33. Bartholdi J, Hackman ST (2019) *Warehouse and distribution science*. Release 0.98.1. Supply Chain and Logistics Institute
 34. Richards G (2021) *Warehouse management. A complete guide to improving efficiency and minimizing costs in the modern warehouse*. KoganPage, London
 35. Grosse EH, Glock CH, Jaber MY, Neumann WP (2015) Incorporating human factors in order picking planning models: framework and research opportunities. *International Journal of Production Research* 53/3(3):695–717. doi:10.1080/00207543.2014.919424
 36. Hübner A, Holzapfel A, Kuhn H (2015) Operations management in multi-channel retailing. An exploratory study. *Operations Management Research* 8/3(3):84–100. doi:10.1007/s12063-015-0101-9
 37. Hübner A, Kuhn Heinrich, Sternbeck Michael G. (2013) Demand and supply chain planning in grocery retail. An operations planning framework. *International Journal of Retail & Distribution Management* 41/7(7):512–530. doi:10.1108/IJRDM-05-2013-0104
 38. de Koster RBM, Le-Duc T, Roodbergen KJ (2007) Design and control of warehouse order

- picking. A literature review. *European Journal of Operational Research* 182/2(2):481–501. doi:10.1016/j.ejor.2006.07.009
39. Rice AK (1958) Productivity and social organization. The ahmedabad experiment: technical innovation, work organization and management, 7th edn. Routledge, London
 40. Colding J, Barthel S (2019) Exploring the social-ecological systems discourse 20 years later. *Ecology and Society* 24/1(1). doi:10.5751/ES-10598-240102
 41. van de Sand K (1995) The socio-political system and development. *Intereconomics* 30/1(1):3–8. doi:10.1007/BF02926355
 42. Trist EL, Bamforth KW (1951) Some Social and Psychological Consequences of the Longwall Method of Coal-Getting. *Human Relations* 4/1(1):3–38. doi:10.1177/001872675100400101
 43. Cherns A (1976) The Principles of Sociotechnical Design. *Human Relations* 29/8(8):783–792. doi:10.1177/001872677602900806
 44. Walker G (2015) Come back sociotechnical systems theory, all is forgiven ... *Civil Engineering and Environmental Systems* 32/1-2(1-2):170–179. doi:10.1080/10286608.2015.1024112
 45. Patriarca R, Falegnami A, Costantino F, Di Gravio G, Nicola A de, Villani ML (2021) WAX: An integrated conceptual framework for the analysis of cyber-socio-technical systems. *Safety Science* 136:105142. doi:10.1016/j.ssci.2020.105142
 46. Neumann WP, Winkelhaus S, Grosse EH, Glock CH (2021) Industry 4.0 and the human factor – A systems framework and analysis methodology for successful development. *International Journal of Production Economics* 233:107992. doi:10.1016/j.ijpe.2020.107992
 47. International Ergonomics Association (2021) Human Factors/Ergonomics (HF/E). Definition and Applications
 48. Boudreau J, Hopp W, McClain JO, Thomas LJ (2003) On the Interface Between Operations and Human Resources Management. *Manufacturing & Service Operations Management* 5/3(3):179–202. doi:10.1287/msom.5.3.179.16032
 49. Bendoly E, Donohue K, Schultz KL (2006) Behavior in operations management. Assessing recent findings and revisiting old assumptions. *Journal of Operations Management* 24/6(6):737–752. doi:10.1016/j.jom.2005.10.001
 50. Gino F, Pisano GP (2008) Toward a Theory of Behavioral Operations. *Manufacturing & Service Operations Management* 10/4(4):676–691. doi:10.1287/msom.1070.0205
 51. Wieland A (2020) Dancing the Supply Chain: Toward Transformative Supply Chain Management. *Journal of Supply Chain Management*. doi:10.1111/jscm.12248
 52. Helander MG (2006) A guide to human factors and ergonomics, 2nd edn. CRC/Taylor & Francis, Boca Raton, Fla.
 53. Fuster JM (2004) Upper processing stages of the perception-action cycle. *Trends in cognitive sciences* 8/4(4):143–145. doi:10.1016/j.tics.2004.02.004
 54. Croson R, Schultz K, Siemsen E, Yeo ML (2013) Behavioral operations. The state of the field. *Journal of Operations Management* 31/1-2(1-2):1–5. doi:10.1016/j.jom.2012.12.001
 55. Loch CH, Wu Y (2005) Behavioral Operations Management. *Foundations and Trends in Technology, Information and Operations Management* 1/3(3):121–232. doi:10.1561/02000000009
 56. Bendoly E, Croson R, Goncalves P, Schultz K (2010) Bodies of knowledge for research in behavioral operations. *Production and Operations Management* 19/4(4):434–452. doi:10.1111/j.1937-5956.2009.01108.x
 57. Bendoly E, Donohue K, Schultz KL (2006) Behavior in operations management. Assessing recent findings and revisiting old assumptions. *Journal of Operations Management* 24/6(6):737–752. doi:10.1016/j.jom.2005.10.001
 58. Bendoly E (2006) Incorporating behavioral theory in OM empirical models. *Journal of Operations Management* 24/6(6):735–736. doi:10.1016/j.jom.2006.09.003
 59. Véricourt F de, Jain K, Bearden JN, Filipowicz A (2013) Sex, risk and the newsvendor. *Journal of Operations Management* 31/1-2(1-2):86–92. doi:10.1016/j.jom.2012.11.001
 60. Moritz B, Hill A, Donohue K (2013) Individual differences in the newsvendor problem. Behavior and cognitive reflection. *Journal of Operations Management* 31/1-2(1-2):72–85. doi:10.1016/j.jom.2012.11.006
 61. Cragg T, Loske D (2019) Perceived work autonomy in order picking systems. An empirical analysis. *IFAC-PapersOnLine* 52/13(13):1872–1877. doi:10.1016/j.ifacol.2019.11.475
 62. Urda J, Loch CH (2013) Social preferences and emotions as regulators of behavior in processes. *Journal of Operations Management* 31/1-2(1-2):6–23. doi:10.1016/j.jom.2012.11.007
 63. Inagaki T, Itoh M (2013) Human's Overtrust in and Overreliance on Advanced Driver Assistance Systems: A Theoretical Framework. *International Journal of Vehicular Technology* 2013:1–8. doi:10.1155/2013/951762
 64. Hoff KA, Bashir M (2015) Trust in automation. Integrating empirical evidence on factors that influence trust. *Human Factors* 57/3(3):407–434. doi:10.1177/0018720814547570
 65. Lee JD, See KA (2004) Trust in automation. Designing for appropriate reliance.

- Human Factors 46/1(1):50–80. doi:10.1518/hfes.46.1.50.30392
66. Auttanate N, Chotiphon C, Maruo SJ, Näyhä S, Jussila K, Rissanen S, Sripaiboonkij P, Ikäheimo TM, Jaakkola J, Phanprasit W (2020) Cold-related symptoms and performance degradation among Thai poultry industry workers with reference to vulnerable groups: A cross-sectional study. *BMC Public Health* 20/1(1). doi:10.1186/s12889-020-09272-6
 67. Flodin U, Rolander B, Löfgren H, Krapi B, Nyqvist F, Wåhlin C (2018) Risk factors for neck pain among forklift truck operators: A retrospective cohort study. *BMC Musculoskeletal Disorders* 19/1(1). doi:10.1186/s12891-018-1956-3
 68. Mkandawire C, Imler S, Smith J (2016) Obese forklift operator neck loads and back loads on a sit down lift truck during a sudden drop. *ASME International Mechanical Engineering Congress and Exposition, Proceedings (IMECE)* 14. doi:10.1115/IMECE201665169
 69. Al-Shaebi A, Khader N, Daoud H, Weiss J, Yoon SW (2017) The Effect of Forklift Driver Behavior on Energy Consumption and Productivity. *Procedia Manufacturing* 11:778–786. doi:10.1016/j.promfg.2017.07.179
 70. Botelho C, Luisa Matos M (2017) Whole-body vibration exposure in forklift operators – a short review. *Occupational Safety and Hygiene V – Proceedings of the International Symposium on Occupational Safety and Hygiene, SHO 2017*. doi:10.1201/9781315164809-24
 71. Joubert DM, London L (2007) A cross-sectional study of back belt use and low back pain amongst forklift drivers. *International Journal of Industrial Ergonomics* 37/6(6):505–513. doi:10.1016/j.ergon.2007.02.005
 72. Lang A, Günthner WA (2017) Evaluation of the usage of support vector machines for people detection for a collision warning system on a forklift. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 10293 LNCS:322–337. doi:10.1007/978-3-319-58481-2_25
 73. Lang A, Fottner J (2017) Intelligent collision warning for forklift operators [Intelligente Kollisionswarnung für Gabelstaplerfahrer]. *Technische Sicherheit* 7/11-12(11-12):59–61
 74. Moseley J, Howell J, Lim S (2017) A novel forklift safety system utilizing Xbees and directional Microstrip Patch Antennas. *Conference Proceedings – IEEE SOUTHEASTCON*. doi:10.1109/SECON.2017.7925280
 75. Cao L, Depner T, Borstell H, Richter K (2019) Discussions on sensor-based assistance systems for forklifts. *Smart SysTech 2019 – European Conference on Smart Objects, Systems and Technologies*
 76. Du Z, Zhang X (2020) Research on fatigue monitoring of forklift driver driving in an automobile enterprise. *Lecture Notes in Electrical Engineering* 645:251–259. doi:10.1007/978-981-15-6978-4_31
 77. Widanarko B, Modjo R (2017) Occupational risk factors for acute fatigue symptoms among Indonesian beverage industry workers. *Journal of International Dental and Medical Research* 10/3(3):1052–1054
 78. Choi M, Ahn S, Seo J (2020) VR-Based investigation of forklift operator situation awareness for preventing collision accidents. *Accident analysis and prevention* 136. doi:10.1016/j.aap.2019.105404
 79. Bell DR, Gallino S, Moreno A (2014) How to win in an omni-channel world. *MIT Sloan Management Review* 56/1(1):45–55
 80. Grosse EH, Glock CH (2015) The effect of worker learning on manual order picking processes. *International Journal of Production Economics* 170:882–890. doi:10.1016/j.ijpe.2014.12.018
 81. Winkelhaus S, Sgarbossa F, Calzavara M, Grosse EH (2018) The effects of human fatigue on learning in order picking: an explorative experimental investigation. *IFAC-PapersOnLine* 51/11(11):832–837. doi:10.1016/j.ifacol.2018.08.442
 82. Heck RH, Thomas SL, Tabata LN (2014) *Multilevel and longitudinal modeling with IBM SPSS*, 2nd edn. Routledge, New York
 83. Krull JL, MacKinnon DP (2001) Multilevel Modeling of Individual and Group Level Mediated Effects. *Multivariate behavioral research* 36/2(2):249–277. doi:10.1207/S15327906MBR3602_06
 84. Preacher KJ, Curran PJ, Bauer DJ (2006) *Computational Tools for Probing Interactions in Multiple Linear Regression, Multilevel Modeling, and Latent Curve Analysis*. *Journal of Educational and Behavioral Statistics* 31/4(4):437–448. doi:10.3102/10769986031004437
 85. Hox JJ (2010) *Multilevel analysis. Techniques and applications*, 2nd edn. Routledge Taylor & Francis, New York
 86. McCoach DB, Black AC (2012) Introduction to estimation issues in multilevel modeling. *New Directions for Institutional Research* 2012/154(154):23–39. doi:10.1002/ir.20012
 87. McGraw KO, Wong SP (1996) Forming inferences about some intraclass correlation coefficients. *Psychological Methods* 1/1(1):30–46. doi:10.1037/1082-989X.1.1.30
 88. Raudenbush SW, Bryk AS (2002) *Hierarchical linear models. Applications and data analysis methods*, 2nd edn. Sage Publ, Thousand Oaks, Calif.
 89. Koo TK, Li MY (2016) *A Guideline of Selecting and Reporting Intraclass Correlation Coefficients*

- for Reliability Research. *Journal of chiropractic medicine* 15/2(2):155–163. doi:10.1016/j.jcm.2016.02.012
90. Glock CH, Grosse EH, Neumann WP, Feldman A (2020) Assistive devices for manual materials handling in warehouses. A systematic literature review. *International Journal of Production Research*:1–24. doi:10.1080/00207543.2020.1853845
91. Mills M (2011) *Introducing survival and event history analysis*. Sage, Los Angeles
92. Battini D, Glock CH, Grosse EH, Persona A, Sgarbossa F (2016) Human energy expenditure in order picking storage assignment: A bi-objective method. *Computers & Industrial Engineering* 94:147–157. doi:10.1016/j.cie.2016.01.020
93. Grosse EH, Glock CH, Neumann WP (2017) Human factors in order picking. A content analysis of the literature. *International Journal of Production Research* 55/5(5):1260–1276. doi:10.1080/00207543.2016.1186296
94. Elbert RM, Franzke T, Glock CH, Grosse EH (2017) The effects of human behavior on the efficiency of routing policies in order picking: The case of route deviations. *Computers & Industrial Engineering* 111:537–551. doi:10.1016/j.cie.2016.11.033
95. Grosse EH, Glock CH, Jaber MY, Neumann WP (2015) Incorporating human factors in order picking planning models: framework and research opportunities. *International Journal of Production Research* 53/3(3):695–717. doi:10.1080/00207543.2014.919424
96. Pasparakis A, de Vries J, de Koster RBM (2021) In Control or under Control? Human-Robot Collaboration in Warehouse Order Picking. *SSRN Journal*. doi:10.2139/ssrn.3816533