In pursuit of humanised order picking planning: methodological review, literature classification and input from practice

Peer-reviewed author version

DE LOMBAERT, Thomas; BRAEKERS, Kris; De Koster, Rene & RAMAEKERS, Katrien (2023) In pursuit of humanised order picking planning: methodological review, literature classification and input from practice. In: International journal of production research (Print), 61 (10), p. 3300-3330.

DOI: 10.1080/00207543.2022.2079437
Handle: http://hdl.handle.net/1942/37627

In pursuit of humanised order picking planning: methodological review, literature classification and input from practice

Thomas De Lombaert[a]*, Kris Braekers[a], René De Koster[b] and Katrien Ramaekers[a]

[a]Research Group Logistics, UHasselt - Hasselt University, Hasselt, Belgium;

[b]Rotterdam School of Management, Erasmus University, Rotterdam, The Netherlands

* Corresponding author. E-mail address: Thomas.delombaert@uhasselt.be (T. De Lombaert)

Abstract: At the core of every high-performing warehouse is an efficient order picking (OP) system. To attain such a system, policy choices should be carefully aligned with subjects responsible for the actual picking within the established system. Despite recent advancements in automating the picking process due to Industry 4.0, human operators will continue to play a crucial role in the future of warehousing. However, unlike robots, human operators have specific skills, conduct, and perceptions, which are only partly accounted for in current planning models. This review adopts a multimethod approach to identify and analyse how these phenomena are currently integrated into OP planning problems. In addition, we assess the relevance and adequacy of human factors modelling in academic literature with practice-based insights gathered via semi-structured interviews. This leads to five major human factors integration constructs and dedicated recommendations on how to refine them. We then take the analysis one step further and make suggestions on how to integrate these constructs with leading research methodologies in the context of Industry 5.0. The results highlight the prevalent need to increasingly account for psychosocial phenomena and their impact on operational performance. Future research opportunities provide a substantiated foundation to assist in human-centric work design.

Keywords: logistics, warehouse operations management, human factors, literature review, multimethod approach, Industry 4.0, Industry 5.0

1. Introduction

Rigorous operational planning enables warehouses to remain resilient and competitive when coping with current and upcoming logistic challenges. These challenges, such as globalisation and changes in consumer purchasing behaviour, lead to an intensified competitive business environment in which warehouses process many orders within stringent time windows (Marchet, Melacini, and Perotti 2015). Several prominent activities, such as receiving, storing, picking, and shipping stock keeping units (SKUs), are performed in a warehouse. Among all the activities, order picking (OP), defined as the process of retrieving products from their storage location in a warehouse, is found to be the costliest (Tompkins et al. 2010). Consequently, previous research efforts have focused on the most efficient way of organising the order picking task. Warehouse managers often focus on short-term normative optimality, such as taking measures for attaining daily targets without considering their long-term repercussion, to minimise the costs related to OP (Vanheusden et al. 2022). However, a driving force behind these costs is the degree of congruence between the system's configuration, stipulated by the OP planning, and the subjects operating in this system (Glock et al. 2017; Gajšek et al. 2020). In other words, a paradigmatic shift towards a holistic harmonisation of the subjects who pick orders and their planning is vital to fulfil customers' requirements efficiently.

On the one hand, OP planning entails the ensemble of strategic, tactical and operational decisions that determine how an order picking system is organised. Van Gils et al. (2018) provide a comprehensive review of these planning problems. The resources that operate within the configured OP system, such as humans or robots, have inherent capabilities, conduct, and interdependencies, which also need to be considered (Grosse et al. 2015; Pasparakis, De Vries, and De Koster 2021). Michel (2016) estimates that no more than 10% of all warehouses can be categorised as parts-to-picker despite the growing interest in automated warehouses. Moreover, he found that less than 3% of all warehouses are entirely automated. In fact, according to Azadeh, De Koster, and Roy (2019), about 40 warehouses in Western Europe were fully automated and operational at that time. Since the vast majority of modern-day warehouses still (partly) rely on human participation for their day-to-day operations (M14 Intelligence 2021), human operators are currently omnipresent in warehouse environments.

Despite their presence, human operators and their mutually distinctive characteristics are often not or unrealistically reflected in planning models (Boudreau et al. 2003). To date, the apprehensions of Boudreau et al. (2003) are still relevant and frequently resonate within more recent literature (e.g. Neumann and Village 2012; Ostermeier 2020; Sgarbossa et al. 2020; Gräßler, Roesmann, and Pottebaum 2021). In addition, the advent of Industry 4.0 (I4.0) has radically altered the role of human operators in operating systems, while appropriate interactive human-technology considerations still lag behind (Neumann et al. 2021). I4.0 symbolises the technology-oriented development of advanced digitalised and automated work environments (Frank, Dalenogare, and Ayala 2019). Winkelhaus, Grosse, and Morana (2021) consolidate findings from the I4.0 research stream and apply them in an OP environment, establishing key characteristics of OP 4.0. The authors explicitly refer to human factors in their definition of OP 4.0, thereby alluding to the recent emergence of Industry 5.0 (I5.0). The latter is a complementary stream which aims at overcoming the shortcomings of I4.0, for example by paying more attention at social outcomes and human-centric work design (Breque, De Nul, and Petridis 2021). In an OP-context, the transition from I4.0 to I5.0 is thus one that shifts attention away from a sole focus on technology as such towards a holistic integration of new technologies with sustainable worker and societal welfare. For example, where I4.0 focused on the introduction of collaborative robots to increase the efficiency of the picking process, I5.0 goes beyond and additionally optimises the human-robot collaboration such that increments in productivity do not come at the expense of picker well-being. Studies have already shown that an organisation that pays little attention to the characteristics of individual employees is more likely to induce worker dissatisfaction, chronic stress, employee turnover, and burn-outs (Devi, Munuswamy, and Jasim 2019; Min 2007). Since the (un)availability of qualified labour (also known as the 'war for talents') directly affects warehouse productivity, companies should be incentivised to prevent employee shortages, as these may negatively affect their competitiveness (Min 2004). Although the sophisticated technologies in I4.0 entail some challenges, they most certainly provide major opportunities in light of I5.0 as well (Breque, De Nul, and Petridis 2021). Neumann et al. (2021) show the existent gap between human considerations in I4.0 systems and the resulting impact on employees and society. Although earlier works studied worker-specific factors in a generic industrial context (Larco 2010), Grosse et al. (2015) were the first to systematically investigate the same premise in the specific context of order picking and detected a prevalent disconnection between OP planning and the rigorous consideration of human operators. Recently, some OP researchers have invoked insights from the Human Factors (HFs) (or ergonomics) discipline, which aims to improve worker well-being and operational performance (Bridger 2017). However, studies that integrate human factors into OP planning are still scarce (Grosse et al. 2015). In order to reap the benefits from the I4.0-to-I5.0 transition, it is crucial to understand *how* HFs are currently integrated in OP planning, such that tailored research directions with regards to human-technology interaction can be proposed. Prior review articles already mapped out which HFs are significant in OP (Gajšek et al. 2017; Grosse et al. 2015; Grosse, Glock, and Neumann 2015; 2017; Grosse et al. 2017; Vanheusden et al. 2022). However, very little attention has been paid to the methodological manner in which these HFs are concretely integrated in OP planning models. Recent review papers, such as Vanheusden et al. (2022), succeed in grouping relevant HFs, but lack a more in-depth discussion on how these HFs should be concretely integrated into planning problems. Moreover, they often resort to a purely literature-based research approach, thereby missing out on the enriching potential of additional research methods. Such complementary methods (e.g. qualitative research) may shed a unique and original light on the topic at hand, uncovering several new research avenues in light of the I4.0-to-I5.0 transition. In addition, review articles which specifically focus on I4.0-related topics are either not explicitly tailored to the OP process (Sony and Naik 2020; Reiman et al. 2021; Kadir, Broberg, and Conceição 2019), or are to a lesser extent focussed on the modelling part (Neumann et al. 2021; Winkelhaus, Grosse, and Morana 2021). Sun et al. (2019) conclude that future research should not only focus on which, but also on how relevant human factors can be included in operations planning and which methodologies should be used.

The realisations of the paper at hand are threefold. First, it bridges the gap between practice-based user requirements and their conversion into academic modelling features. Unlike prior review articles that predominantly examined *which* HFs should be integrated (e.g. Grosse et al. 2017 and Vanheusden et al. 2022), we focus on *how* HFs can be integrated to attain models that improve either operational performance, worker well-being, or both simultaneously. We identify methodological practices from the literature and classify them to present a comprehensive overview of approaches to integrating HFs into OP planning. Hence, we approach the consideration of HFs in OP from a very unique angle compared to recent reviews (e.g. Vanheusden et al. (2022)), resulting in tailored and foremost original insight and research opportunities. Second, since this paper corroborates human-centric planning design, we confront academic modelling practices with practice-based insights. In line with the I4.0-to-I5.0 transition, we consciously account for the opinion of pickers during semi-structured interviews. Being the ones that actually work in the configured system, order pickers can shed a very unique light on the respective matter. This multimethodological approach allows us to identify the quality, adequacy and gaps of existing modelling constructs, a unique approach compared to other review articles in this research domain. Third, we conceptualise HFs in OP with regard to human-technology interaction by associating theory-practice (dis)similarities with insights from the I4.0 research stream. In order to avoid incompleteness and preconceptions in the assessment of OP systems, our analysis starts with a dominant focus on HFs in OP. Thereafter, findings are tailored to the I4.0-to-I5.0 transition. According to Kadir, Broberg, and Conceição (2019), a human factors and ergonomics approach is a highly valuable research perspective for analysing and understanding human-centric work design in Industry 4.0. Our main contributions are as follows. First, we conduct a systematic review of the literature and include an up-to-date and broad spectrum of research methodologies in OP planning. Hereby, our focus is on quantitative methods, i.e., the statistical, mathematical, or numerical analysis of (collected) data. Studies that solely focus on qualitative methods are not considered, despite their value for advancing research on OP, albeit from a different perspective. Second, we cross-check academic research outputs in this field with the concerns and desires of pickers and their supervisors, which is a fairly innovative approach in this field. Finally, we propose a holistic starting point to support future research with guidelines that originate from the literature and on-the-job experiences. In this way, managers and academics find sufficient opportunities to facilitate the I4.0-to-I5.0 transition.

The remainder of this paper is structured as follows. Section 2 elaborates on the multimethod approach and its constituting elements. In Section 3, we conduct an in-depth analysis and cross-check academic practices and interview findings and raise implications for the I4.0-to-I5.0 transition. Section 4 synthesises the results and provides generic guidelines for future research and Section 5 concludes this paper and briefly touches on its limitations and managerial implications.

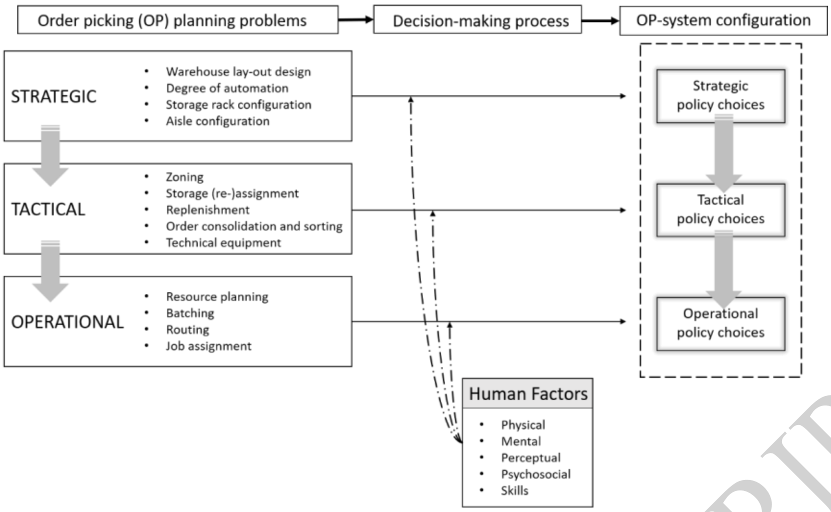
2. Methodology

We first provide a holistic understanding of human-centric modelling by assessing current methods to integrate human factors in order picking planning problems. This induces an approach that captures the interface between academic research and practitioners' interests. Multimethod approaches are suitable for coping with these types of investigations (Singhal et al. 2008). Such an approach involves applying more than one research method to investigate a certain research question with the aim of fusing each method's strengths and providing insights from different angles (Lewis-Beck, Bryman, and Liao 2004). Therefore, we first conduct a systematic review of the literature on human factors in order picking planning. In general, literature reviews provide a comprehensive and up-to-date overview of the state-of-the-art in one specific research domain (Choi, Cheng, and Zhao 2016; Snyder 2019). They are regularly enriched with potentially fruitful research avenues for academics and practitioners, which is also the aim of this paper. To increase the potential for practical applications, researchers can combine literature studies with expert focus groups or interviews (Davarzani and Norrman 2015; Setayesh et al. 2021). Studies by Boudreau et al. (2003) and Ryan et al. (2011) highlight the importance of embracing qualitative research in pursuit of high-quality Operations Management models. Trends within academic research may be opposed to preoccupations of practitioners, allowing to scrutinise the plausibly pending theory-practice gap. For this reason, we complement our literature review with semi-structured interviews with both order pickers and supervisors. This type of interview is highly valuable for data collection on behavioural issues in an order picking context (Grosse et al. 2016). Additionally, the involvement of actors with different hierarchical positions and related responsibilities allow to uncover possibly conflicting viewpoints on the same matter, thereby enriching our analysis.

2.1 Literature review

Figure 1 summarises the scope of this review. In general, order picking planning problems yield a set of policy choices at three organisational levels. Altogether, these policy choices constitute the OP-system which is the cornerstone of every company's internal warehousing operations. However, the consistent consideration of human factors during decision-making has been out of scope in many research papers. The dotted arrows in Figure 1 are currently only scarcely embedded in this process. The scope of this review is aimed at these dotted arrows, and more precisely their interference in the decision-making process. The considered set of OP planning problems is primarily based on the classification of Van Gils et al. (2018), whereas the generic human factors categories (i.e. physical, mental, perceptual, and psychosocial) are adopted from Grosse et al. (2015). The latter framework is extended with "skills", as this HF-category is frequently employed in related industrial domains, such as manufacturing and schedule modelling (e.g. Norman et al. 2002; Techawiboonwong, Yenradee, and Das 2006; Attia, Duquenne, and Le-Lann 2014; Ogbeyemi et al. 2021), and appears more and more in OP literature (e.g. Matusiak, De Koster, and Saarinen 2017; Vanheusden et al. 2022). In this way, an all-embracing view on HFs in an industrial context is adopted.

Figure 1. Scope of the review.



We used Denyer and Tranfield's (2009) review protocols as guidance to cover all relevant literature and ensure a transparent scientific approach. The first step involves identifying relevant papers for the scope of this research. We, therefore, conducted a keyword-based search on three different databases, i.e. Scopus, Web of Science and Google Scholar (see Appendix A). In the next step, we formulate selection criteria (SC) to further align the literature sample with this study's scope. We then apply the following two criteria for the legitimate and transparent inclusion or exclusion of papers from the sample while reviewing their title, keywords, and abstract.

- SC1: No duplicates, written in English, only research papers published in peer-reviewed journals or conference proceedings, i.e., excluding review papers
- SC2: A clear focus on human factors and the order picking process; The paper should look at, analyse, or model human-related phenomena in an order picking (or warehouse) context.

Finally, selection criteria 3 was applied when reading the full papers to further refine the sample.

- SC3: Only relevant papers should be included.

A paper is deemed relevant if its principal focus is on the integration of HFs in OP planning problems, which we define as '*the rigorous consideration and incorporation of human conduct, capabilities and/or interdependencies, whose impact on operation outcomes or worker well-being is explicitly examined and substantiated when planning the order picking process*'. This implies that we only consider papers that quantitatively examine multiple policies for at least one planning problem and account for at least one human factor, as visualised in Figure 1.

Papers that did not comply with these selection criteria were excluded from the final literature sample. In addition, as this respective research area is relatively nascent and the usage of uniform terminology is still in progress, we decided to use a snowball search approach. In particular, a combined forward and backward reference search was conducted. The resulting papers were reviewed under the very same scrutiny as the initial paper sample to assess their eligibility for inclusion in the final sample. This approach is highly suited to overcome the burden of incompleteness as a result of inconsistent terminology (Lecy and Beatty 2012) and is frequently used in other related literature reviews, e.g. Boysen, De Koster, and Weidinger (2019) and Grosse et al. (2015). This led to a sample of 57 papers. Figure 2 depicts a general overview of the sampling process.

Figure 2. Flowchart of study search and inclusion process.

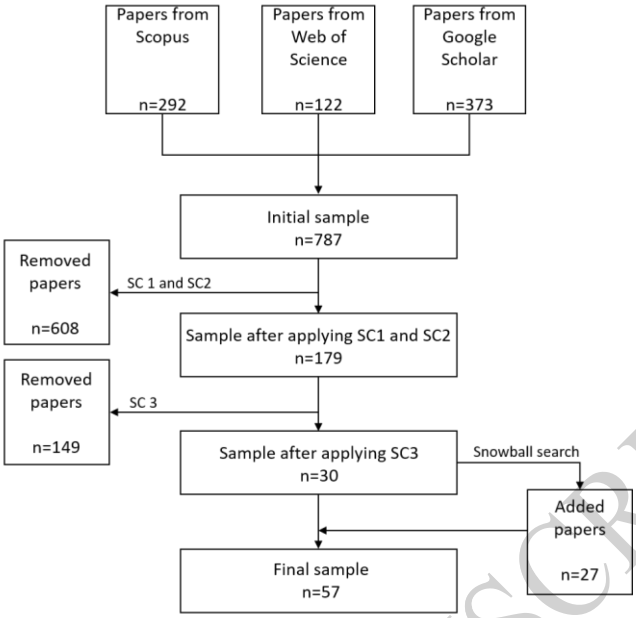


Figure 3 shows how the cumulative number of papers in our sample has evolved throughout the years. The convex shape of the trend line indicates a growing interest in the topic over the past few years. Figure 4 lists the journals in which the sample articles were published.

Figure 3. Cumulative time distribution of reviewed articles.

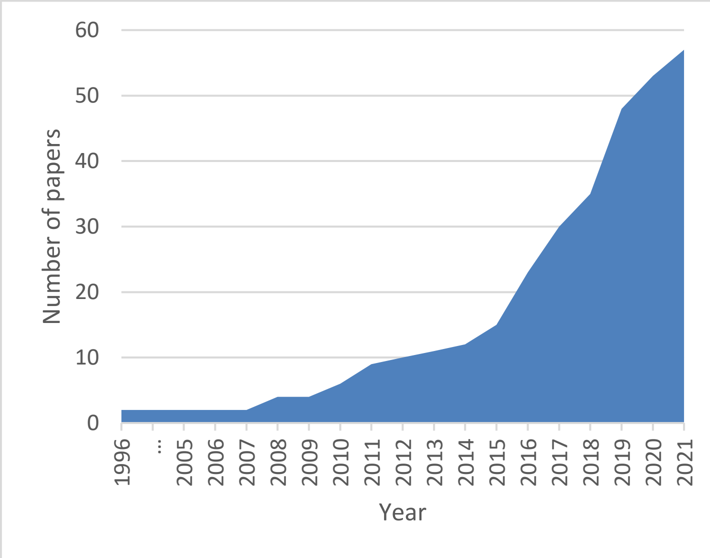
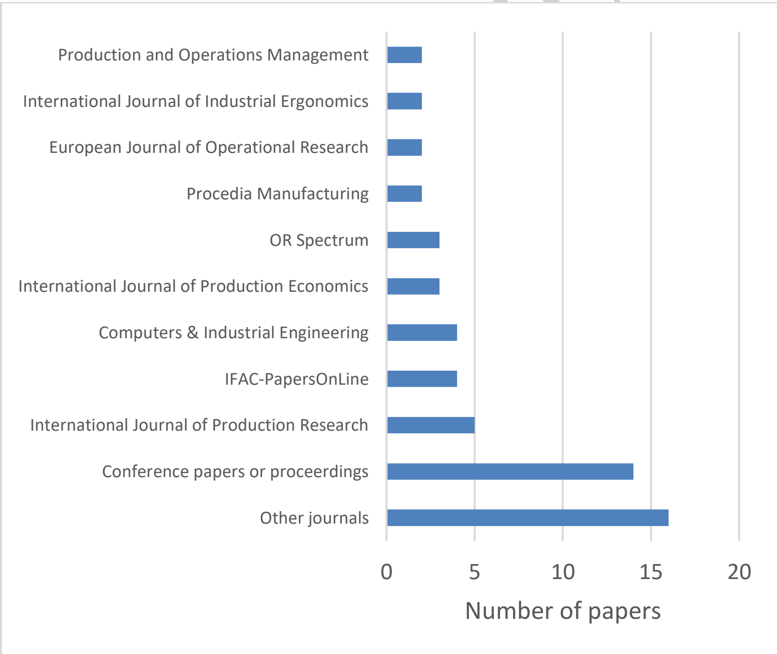


Figure 4. Journals of reviewed articles.



2.2 Expert interviews

We decided to include experts' input for external validity and future practicality (Carter 2008). Hereby, we aimed to elicit and exploit the tacit knowledge, concerns, and experiences embedded in employees who pick orders on a daily basis and their supervisors (Ferrari, Spoletini, and Gnesi 2016). Participants were selected on the basis of a purposive sampling strategy to ensure a broad spectrum of expert perspectives (Lewis-Beck, Bryman, and Liao 2004; Robinson 2014). For example, pickers were selected to include diverse ethnicities, personalities, physical characteristics, etc. To ensure a sufficient degree of representativeness, we purposefully identified well-established companies with different working conditions and environments. Furthermore, only companies with a mature operations strategy and a sufficient share of manual activities were considered. As a result, four major warehouses from various industries in the Flemish area were carefully selected and contacted. Since this research could bear mutual benefits, all four decided to participate in our study and delegated order pickers and supervisors. The warehouses differ in terms of rack type (pallet or shelf), ambient working temperature (refrigerated or not), product properties such as weight and scale, assistive technical equipment (e.g. barcode scanner or pick-by-voice), and job rotation possibilities. A direct consequence of this plurality in working conditions is the rigorous coverage of many order picking system facets and related worker preoccupations. In close consultation with the involved companies and to avoid an unevenly distributed company representation, we aimed to interview 6 to 8 pickers per company. This setting led to a sample of 29 participants (summarised in Table 1) with whom we each conducted a semi-structured interview that lasted for an average of 50 minutes. This interview type elicits information from an employee's viewpoint and keeps the degree of preconception to a minimum, which is a direct result of open-ended questions (Grosse et al. 2016). In addition, our sample follows recommendations to assure sufficient academic rigor concerning sample size (Guest, Bunce, and Johnson 2006; Hennink, Kaiser, and Marconi 2017). Before the interviews, we shared our questions with practitioners and engaged in a focus group discussion to finetune the final interview guide, which encompassed aspects related to picker well-being, perceptions regarding working conditions, and the impact of those aspects on picker performance. These respective aspects were inspired by the current state-of-the-art in this research domain.

Most interviews were held online due to the COVID-19 pandemic and related precautionary measures that impeded physical company visits. However, person-to-person interviews do not grant significantly better data per se compared to other interview settings (Knox and Burkard 2009). Since most interviews could be conducted via online videocall, both verbal and non-verbal communication could be observed. The interviews were conducted in Dutch with Dutch-speaking employees (avg. age: 37.48 (std: 12.13); avg. experience: 9.69 (std: 6.96); 9 females, 20 males) and both video and audio were recorded. After the interview, these data were rigorously transcribed and coded to identify emerging themes and uncover new insights through thematic analysis (Grosse et al. 2016; Wildemuth 2016).

Table 1. Summary of the interview participants.

Interviewee	Company	Gender	Age (in years)	Job role	Experience (in years)
1	A	Female	36	Picker	> 16
2	A	Male	52	Picker	< 4
3	A	Female	24	Picker	< 4
4	A	Male	28	Picker	4 - 8
5	A	Female	49	Supervisor	> 16
6	A	Female	37	Supervisor	8 - 16
7	B	Female	23	Picker	< 4
8	B	Male	33	Picker	4 - 8
9	B	Male	40	Picker	4 - 8
10	B	Male	23	Picker	< 4
11	B	Male	54	Picker	< 4
12	B	Male	19	Picker	< 4
13	B	Male	29	Supervisor	4 - 8
14	B	Male	34	Supervisor	8 - 16
15	C	Male	36	Picker	8 - 16
16	C	Male	55	Picker	8 - 16
17	C	Male	22	Picker	< 4
18	C	Male	52	Picker	> 16
19	C	Female	29	Picker	4 - 8
20	C	Male	37	Picker	8 - 16
21	C	Male	37	Supervisor	> 16
22	C	Female	39	Supervisor	< 4
23	D	Male	29	Picker	4 - 8
24	D	Female	55	Picker	4 - 8
25	D	Male	63	Picker	4 - 8
26	D	Male	25	Picker	< 4
27	D	Female	49	Picker	> 16
28	D	Male	49	Picker	8 - 16
29	D	Male	29	Supervisor	8 - 16

3. Analysis

Reviewed articles can be classified in numerous ways. Unlike previous literature reviews that resorted to the five generic HF categories (physical, mental, psychosocial, perceptual, and skills), or that classified papers based on the examined OP planning problem, this review rather focuses on the employed research methodology and modelling constructs. The underlying thought for this perspective relates to the aspiration of providing a hands-on toolset whose inherent intuitiveness promotes and facilitates human factor integration. Against this backdrop, we present a detailed discussion of modelling constructs (Section 3.1), their relation to the I4.0-to-I5.0 transition (Section 3.2), research methodologies (Section 3.3), and how they fit together (Section 4). Appendix B provides a comprehensive overview of the discussed papers.

3.1 Modelling constructs

Order picking planning problems can only lead to accurate and adequate policy choices if human factors are considered during the decision-making process (Grosse et al. 2017). Normative models should therefore be enriched with additional modelling functionalities. Boudreau et al. (2003) highlight seven assumptions (see Appendix C) that are frequently employed by researchers to simplify their models, thereby disregarding the consequences that these simplifications may have on the decision-making outcomes. The authors advocate a paradigmatic shift towards human-centred models to capture the inherent interaction between operations and human resources. Based on these seven assumptions, insights from interviews with practitioners, and a thorough analysis of our literature sample, we distinguish five main HF integration constructs (hereafter referred to as *modelling constructs*) that each converts relevant human-centred phenomena into modelling functionalities:

- Varying work rates
- Quantitative physical state indicators
- Stochastic worker behaviour and work execution
- Subjective worker experience and judgment
- Socio-demographic worker differentiations

Modelling constructs refer to explicit variables or mathematical expressions that researchers can integrate into their models, rather than the generic HF categories. Accordingly, they should be considered as a way to integrate HFs into OP planning problems to support and enhance policy choices when configuring the OP system. These five modelling constructs are collectively not exhaustive but rather principal constructs that cover best practices to integrate human factors in order picking planning problems. They are also not mutually exclusive, which means that occasional overlap between the modelling constructs could occur. In subsections 3.1.1 to 3.1.5, we discuss each of the constructs, how they address concerns raised by Boudreau et al. (2003), how they are used in current academic research papers, and their relevance in industrial environments.

3.1.1 Varying work rates

The time it takes an industrial worker to complete a task is characterised by a certain degree of variability (Dudley 1968), of which the worker himself is claimed to be the most significant source (Doerr and Arreola-Risa 2000). In an order picking context, this implies that pickers exhibit distinctive task execution times such as pick time, search time, and walking time. The same claim is made by Boudreau et al. (2003), who argue that modellers often portray workers as identical and stationary. In reality, however, we see that pickers differ in task completion times and that individual task completion times are prone to a certain degree of stochasticity and may even change over time. We thus distinguish between *static*, *dynamic*, and *stochastic* varying work rates (VWRs).

Static varying work rates in OP literature

The among-picker differences in work rate that remain steady over time are called static VWRs. It is reasonable to assume that job performance differences among pickers will not change significantly in the absence of phenomena such as training, learning, forgetting, boredom, etc. Several papers in our literature sample use this premise. Feng and Hu (2021) differentiate in upper and lower bound unit pick time among pickers due to intrinsic differences in working habits when configuring the order batching and sequencing in a vegetable OP environment. In the context of an apparel retailer, Batt and Gallino (2019) investigate the task assignment and routing problem, subject to the effect of different levels of experience among a cohort of pickers. Although their model can differentiate pickers in terms of pick performance, it omits the impact of an increased experience level after every single pick for each picker. Matusiak, De Koster, and Saarinen (2017) also examine the job assignment problem. Batch execution times are forecasted based on historic pick data and serve as input for a joint batching and generalised assignment problem. The forecasts reflect the different skill levels in a set of heterogeneous pickers, leading to differences in task execution times between pickers. Skill level disparities are also often used to investigate bucket brigade picking, a system where pickers progressively pick orders and hand over their work-in-progress to a downstream picker (Bartholdi and Eisenstein 1996). Studies by Fibrianto and Hong (2019), Hong (2018), Hong, Johnson, and Peters (2015), Koo (2009), and Webster, Ruben, and Yang (2012) employ empirical observations or qualitative interviews to account for pick speed differences among pickers. Most of these studies do not only incorporate variations in pick speed, but also walk time differences. In this way, real-life behaviour is simulated, as operators' walk time can be affected by work environment and biomechanical dissimilarities (Mohler et al. 2007). Finally, Al-Araidah et al. (2020) illustrate that varying walk times play a prominent role when orders are assigned to pickers.

Although these papers deviate from unrealistic normative models, they raise some limitations that should be addressed in future research to refine their models. For example, Matusiak, De Koster, and Saarinen (2017) argue that among-picker varying work rates could also be examined with a psychosocial perspective and refer to studies by Juran and Schruben (2004) and Larco (2010) to justify this premise. Apart from psychosocial aspects, many papers acknowledge that phenomena such as learning and forgetting or physical workload also strongly influence task execution times and should be accounted for to mimic real-life work systems (Al-Araidah et al. 2020).

Dynamic varying work rates in OP literature

Dynamic VWRs imply that individual task completion times may either improve or deteriorate over time. These job performance alterations stem from phenomena that are inherent to human operators. One of them is the prevalence of learning due to the repetitive nature of tasks (Vanheusden et al. 2022). Grosse and Glock (2013) empirically study the prevalence of learning in an industrial order picking setting and recommend that the importance of learning in the decision-making process should be acknowledged. However, very few research papers have heeded these early calls (Winkelhaus et al. 2018). Apart from learning, other phenomena include amongst others training, fatigue, and boredom (e.g. Bendoly and Prietula (2008) and Calzavara et al. (2019)).

Several papers in our literature sample integrate dynamic VWRs in their analyses. Grosse and Glock (2015) include a pick frequency-dependent search time function in their model to investigate the allocation of pickers to different zones in a warehouse. Their study shows that allocating pickers with the lowest learning rate to a zone with fast moving products is beneficial for the company's operational performance. However, the total time per pick is not only determined by search time but also by setup, travel, and pick time (Tompkins et al. 2010). On this ground, Grosse, Glock, and Jaber (2013) also incorporate experience-dependent pick times to analyse storage (re-)assignment decisions. The authors highlight the indispensable need to consider learning and forgetting in the decision-making process, rather than a sole dogmatic focus on travel distance. Despite the rigorousness of the study by Grosse, Glock, and Jaber (2013), some of their assumptions were challenged by Ogasawara, Ishigaki, and Yasui (2019). Especially, they claim it is impractical to presume that all picker learning is reset after a reassignment in the storage area. It would therefore be more accurate to infer that learning is only reset for the storage shelves where products are re-assigned. The study's results and recommendations are, nonetheless, in line with those of Grosse, Glock, and Jaber (2013). Zhang et al. (2019) also conclude that considering picker learning would promote the accuracy and predictability of the order fulfilment operations. Although they assume a constant travel (or walking) speed, the time to pick an item depends on an initial picking time amended by the number of times this item has already been picked. This routine is common practice in all of the aforementioned papers and inspired by Wright (1936), which is, in turn, founded on empirical observations.

Unlike learning-related phenomena, the physical repercussions of OP, i.e. fatigue, have been linked to a reduction in workers' work rate and operational performance (Winkelhaus et al. 2018). Granotto et al. (2019) make use of a fatigue constant μ to model the slowdown of pickers' speed in a bucket brigade order picking system. This fatigue constant captures the different workload levels when workers pick orders and ranges between ‘zero effort’ and ‘hard work’, resulting in a progressive decay of the worker’s maximal work pace. Similarly, Zhao et al. (2019) use a fatigue factor α to reflect the processing time’s dependence on previous efforts. The concrete constructs used by Granotto et al. (2019) and Zhao et al. (2019) are not based on empirical evidence, but rather inspired by common practices from the literature on generic industrial environments (e.g. Jaber, Givi, and Neumann (2013) and Öztürkoğlu and Bulfin (2012)). Alternatively, Feng and Hu (2021) find inspiration in biological process modelling and adopt a generalised logistic function to adjust job processing times with a work fatiguing effect. In addition to the delimitation of upper and lower working efficiency, the authors suggest that order processing times are represented in a more concrete and natural way.

The previously discussed papers identify two main streams of future research opportunities. First, researchers should address break scheduling, i.e. its frequency, length, and effects (Granotto et al. 2019; Zhao et al. 2019). Current models make some basic assumptions, but these impede a realistic integration of dynamic VWRs. Second, researchers should examine the impacts of boredom as a consequence of systematically assigning the same pickers to the same task (Grosse and Glock 2015). Neglecting these impacts could put efficiency at stake in the long run (Neumann and Dul 2010). Although specific research is yet to be done, integrating dynamic VWRs in future analyses would greatly benefit the accuracy and usability of decision support models.

Stochastic varying work rates in OP literature

In many cases, both static and dynamic VWRs view task execution times as a concept that is not exposed to stochastic variations. However, in reality, pick times are not deterministic, implying that no two picks are the same. Therefore, some papers include stochastic VWRs in their model. Batt and Gallino (2019) use a stochastic time per pick realization ($time_i \sim \mathrm{Exp}(x_i\beta + \varepsilon_i)$) that depends on two factors. The deterministic part is a linear predictor that depends on bin density, distance, and pick experience. The stochasticity is introduced by ε_i, drawn from a logistic distribution. Integrating stochastic pick times is also a common practice in papers that study bucket brigade systems. Popular statistic distributions to model the time to retrieve an item from its location are the uniform, triangular, and exponential distribution (Fibrianto and Hong 2019; Hong 2018; Hong, Johnson, and Peters 2015; Koo 2009; Webster, Ruben, and Yang 2012).

Insights from practice

Our interviewees acknowledged the prevalence of varying work rates, although they considered static and stochastic VWRs less important and impactful than their dynamic counterparts. Especially, order pickers were only to a limited extent convinced that among-worker differences play a significant role in a warehouse environment. This contrasts sharply with the importance that pickers (N=15/22) attribute to dynamic VWRs. The most frequently raised arguments to underpin this conviction were the accumulation of physical efforts during the day and their morningness versus eveningness nature (Horne and Östberg 1976). The time of the day apparently significantly influences work rates, arguably due to workers' circadian rhythm. As a result of technological advancements in Industry 4.0, companies have the opportunity to moderate this impact to some extent, e.g. through smart lightning systems (Füchtenhans, Grosse, and Glock 2021). Apart from optimising the working condition, the work itself can also be reorganised to respond to varying work rates. Although some companies consider shift preferences when planning their operations, many strictly adhere to fixed shift rotation schedules. Some interviewees (N=11/22) felt that such regimes negatively influenced their work rate and mention they wanted to participate more in these decisions. The same aspiration holds for break scheduling. Our respondents mentioned that breaks were important to avoid decrements in employees' performance (Li, Xu, and Fu 2020). However, they are only scarcely researched in an order picking context. Rijal et al. (2021), for example, touch on the benefits of integrating breaks in order picker scheduling decisions, but do not account for human factors. Future research could therefore incorporate individual preferences to boost motivation and operational performance.

Furthermore, the interviewees (N=18/22) indicated several other factors that might affect their individual work rates, such as the prevailing target regime and informal inter-colleague relationships. The former originates from a lack of motivation once the predetermined target has been reached. This has already been extensively researched (e.g. Bouwens and Kroos (2011)), although none of these studies are tailored to an order picking environment. The latter relates to the context in which the work rate of one picker is affected by other pickers (e.g. Zhang et al. (2021)), their mutual relationship, or their respective work rates. Interviewees confirmed that, for example, informal talks with colleagues can significantly interfere with their concentration levels and work rate. Hence, errors might occur and can lead to detrimental financial and operational consequences (Vidovič and Gajšek 2020). However, pickers (N=12/22) also emphasised how favourable these informal talks were for their motivation, a conviction shared with many interviewed supervisors. This seemingly paradoxical trade-off can be interesting to investigate with a research perspective similar to the one used by Glock et al. (2017).

Another concern relates to the fact that although warehouse task assignments based on skills and learning rates are favourable in terms of work rates (Grosse and Glock 2015; Matusiak, De Koster, and Saarinen 2017), pickers are not the strongest proponents of this premise. Specially, they fear that repeatedly being assigned to identical jobs would hamper long-term motivation and deteriorate work rates. They (N=17/22) mentioned that companies should include job rotation in their personnel scheduling planning. Many supervisors agreed with this premise and claimed that the introduction of well-thought-through job rotation scheduling could be a cornerstone of improved worker motivation. Despite its potential to improve worker well-being, job rotation is prone to skill considerations in different warehouse zones (e.g. one's ability to drive a forklift, specific lifting capabilities, etc.) which could complicate implementation in practice. The concept of job rotation in a warehouse context is therefore deemed to be an interesting topic of study in future research.

In conclusion, we can state that academic research papers primarily succeed in capturing the basic phenomena related to varying work rates. Nevertheless, many of the interviewees' concerns are still understudied (i.e. picker participation, target- and feedback-regimes, and job rotation schedules) and are therefore considered to be fruitful research avenues.

3.1.2 Quantitative physical state indicators

The repetitive nature of OP, occasional awkward body postures and continuous engagement in physical activities make order picking a physically demanding and exhausting job (Lee, Kim, and Chang 2016). Logically, this severely affects the physical well-being of order pickers during a working day and can lead to high absenteeism rates and work-related musculoskeletal diseases (MSDs), causing significant long-term costs (Punnett and Wegman 2004). Therefore, Boudreau et al. (2003) criticise the utopic assumption of perfect worker availability (i.e. no absenteeism) and the disregard of workers' tiredness in operational models. Some of the papers in our sample explicitly integrate pickers' physical well-being in their analyses.

Quantitative physical state indicators in academic OP literature

Rather than perceived workload or strain levels, this modelling construct category focuses on objectively measurable indicators that assess the physical state of an order picker. By integrating such indicators, the physical state of order pickers can be monitored, and the work environment and task assignments can be organised in an efficient, yet ergonomic matter. Various indicators can be employed to assess the physical state of order pickers (Stanton et al. 2004).

The vast majority of the papers in our literature sample uses the energy expenditure (EE) concept, which expresses the amount of energy that a person needs to carry out a given task (Levine 2005). Garg, Chaffin, and Herrin (1978) tailored the calculations of human energy expenditure to manual material handling environments, thereby providing an accurate foundation for further analysis. The EE indicator allows researchers to keep pickers' efforts within a range to work productively, without being exposed to excessive fatigue levels. For example, extreme EE levels can cause deteriorated concentration and related error rates (Zhao et al. 2019). To avoid this from occurring, one can either minimise the total EE associated with policy choices (Battini et al. 2016; Calzavara et al. 2017; Calzavara et al. 2019; Diefenbach and Glock 2019; Elbert and Müller 2017; Gajšek et al. 2021), or allow the EE to fluctuate between predetermined limits (Al-Araidah et al. 2020; Zangaro et al. 2019). However, many of the papers rely on standard or average EE calculations and recommend that future research should thoroughly investigate the impact of, for example, pickers' anthropometry (Al-Araidah et al. 2020; Calzavara et al. 2019; Elbert and Müller 2017). A fine addition to the EE concept is the measurement of a picker's heartbeat, pupil diameter, or muscular strains. Despite the possible necessity for additional measuring tools, this approach is frequently used to account for the physiological workload (Braam, van Dormolen, and Frings-

Dresen 1996; Chen et al. 2020; Lee, Chang, and Karwowski 2020; Passalacqua et al. 2020; Steinebach, Wakula, and Mehmedovic 2021; Stockinger et al. 2020).

Apart from integrating energy expenditure or heartbeat, some papers resort to indicators that focus on pickers' posture. Hanson et al. (2018) use the REBA (Rapid Entire Body Analysis) score to analyse the impact of tilting pallets on pickers' ergonomics. Calzavara et al. (2019) employ the OWAS (Ovako Working Posture Analysing System) index to assess the combined storage assignment and rack configuration problem, thereby preventing pickers from being overly exposed to awkward body postures. Al-Araidah et al. (2017) pursue the same objective with their heuristic, which first clusters items and then assigns each cluster to a location in the storage area. Interestingly, Gajšek et al. (2021) have the same intention, but their storage assignment model also optimises picking time and energy. The authors show that companies could significantly reduce health risks and strenuous working conditions without jeopardising decent picking times. A final study that considers ergonomic monitoring from a postural point of view is the one of Braam, Van Dormolen, and Frings-Dresen (1996), where different body postures are counted and compared among different automation levels.

Furthermore, some papers adopt indicators for quantifying ergonomic risks, for example, lower back injuries. A frequently used metric is the peak L4/L5 spinal compression. It embodies the compressive force at the L4/L5 vertebral joint. Glock et al. (2019) and Hanson et al. (2016) employ this measure to assess the effect of pallet rotations on both economic outcomes and worker well-being. They show that a deliberate pallet configuration improves time efficiency, without compromising the picker's physical integrity. The compressive forces in the intervertebral disc L5-S1 is another analogous metric to track spinal load exertions and related ergonomic risks (Steinebach, Wakula, and Mehmedovic 2021). Another indicator for ergonomic risk is the NIOSH lifting equation (Waters, Putz-Anderson, and Garg 1994). Using this equation, Otto et al. (2017) develop a model to determine products' placements in shelves and partition the work area into different zones. The authors show that a sole focus on economic performance and neglecting ergonomic impact lead to significantly higher health risks.

Three principal themes recur when papers discuss their limitations and future research opportunities. First, papers sometimes limit themselves to the abstraction of pickers, thereby ignoring the inherent heterogeneity among pickers. Accounting for picker-specific characteristics such as height and health condition might prove beneficial in terms of representing reality (Al-Araidah et al. 2020; Hanson et al. 2018). Second, few studies link physical state indicators to operational performance. More research is needed to investigate the direct interdependency between load level and performance (Glock et al. 2019). Finally, extant research is unclear about which metric should be applied in which context (Otto et al. 2017). It is noticeable that most of the papers in our sample make use of the EE indicator. Although this metric is suited to gain general insights into the physical strains acting on operators when performing order picking tasks, it does not capture different types of injury risks to which the worker is exposed. In order to develop a more comprehensive evaluation of a worker's physical strains and risks, future research will benefit from a stronger focus on other physical state indicators. Numerous tools exist to assess biomechanical exposure (e.g. OWAS, NIOSH lifting equation, LUBA, Strain index, etc.). Their usage, however, should be contemplated and thoroughly substantiated, as research seems to agree that no method holds an unconditional advantage over the others (Takala et al. 2010). Additionally, a monetary quantification for ergonomic risks will possibly enrich and motivate human-centric systems even more (Otto and Scholl 2011). Section 4 discusses this apprehension in more detail.

Insights from practice

The pickers we interviewed were convinced of the physically demanding nature of their job and its implications on their well-being. This justifies the keen interest that academics show in this respective topic. Literature has regularly shown that physical well-being and economic performance are not conflicting objectives per se (e.g. Calzavara et al. (2019)). However, only approximately half of the pickers (N=12/22) explicitly acknowledged the ergonomic efforts of their employer. In today's competitive environment where WHs experience an immense pressure on costs, economic objectives are still perceived to be dominant. A fine example relates to ergonomic training sessions. Although many companies wish to arrange sessions to boost ergonomics on the work floor, many operators mention that the prevalent economic target regime was a restraining factor for the actual adoption of session instructions. Targets that are too stringent or insufficiently account for working conditions might induce additional stress and pressure. This can, in turn, nudge order pickers to deviate from what is ergonomically optimal for their physical well-being, e.g. ergonomic body postures. In sum, the main reason for their non-compliance lies in the conviction that economic targets will not be met properly if they strictly adhere to the ergonomic guidelines, a phenomenon also studied by Garg and Saxena (1985). Re-assessing target regimes could prevent this from happening. Apart from this, pickers also seem to some extent resistant to change, which could be tempered if companies offered some kind of incentive.

The interviews also show the prevailing impact of physical fatigue on concentration and pick speed levels (N=17/29). This confirms findings from the literature and more precisely the expedient need to link physical indicators to work rates. Otto et al. (2017) alluded, for example, on the opportunity to add additional breaks to reduce the ergonomic load and increase pickers' performance. Fatigue-recovery models with a bifold focus on financial and ergonomic objectives are deemed favourable to develop such interdisciplinary decision support models in OM (Glock et al. 2019). In conclusion, current academic OP literature covers several concerns that were raised by pickers themselves. Nonetheless, some real-life practices, such as feedback links between physical state and work rates, ergonomic guideline compliance, and fatigue diminishing interventions require further research.

3.1.3 Stochastic worker behaviour and work execution

Workers and their actions are not always entirely predictable, resulting in behaviour that can be different from what is normatively assumed. The anomalies in worker behaviour occur probabilistically but can be steered in the 'right direction' (Sunstein 2018). However, it seems highly unethical and morally wrong to entirely marginalise the autonomy of human operators (Brey 1999). Therefore, these possible deviations from normativity should be embedded ex ante in operational planning models. Boudreau et al. (2003) support this mindset and criticise the utopic portrayal of deterministic and predictable workers. In addition, specific working conditions may occasionally manifest during the operational process such that workers have to deviate from the theoretical planning. Although limited in number, some papers include stochastic worker behaviour and work execution in their analysis.

Stochastic worker behaviour and work execution in OP literature

Contrary to what is normatively assumed, in reality, order pickers may deviate from prescribed procedures, defined as maverick picking (Glock et al. 2017). The unpredictability in worker behaviour puts the initially envisioned optimality at stake. Research papers address this phenomenon by including behavioural probabilities in their models. For example, Ogasawara, Ishigaki, and Yasui (2019) use parameters that represent the probability that a picker accesses a wrong shelf area, whereas Elbert et al. (2017) include route deviation probabilities. Both studies show the importance of integrating stochastic worker behaviour. Finally, studies that deliberately give pickers a certain degree of freedom in carrying out their task highlight the need to account for operators' peculiar but sometimes sub-optimal decisions (Hanson et al. 2018; Hilmola and Tolli 2016).

Insights from practice

Cross-checking academic practices with practitioners' experience shows that order pickers are less concerned with the implications of integrating stochastic behaviour in models. Many pickers (N=14/22) indicated they were seldom triggered to non-comply with what is normatively expected from them, apart from the occasional non-adoption of ergonomic guidelines (see section 3.1.2). Nonetheless, supervisors (N=6/7) partly endorsed this claim, exemplified by unforeseen incidents such as picker congestion, malfunctioning equipment, and stock shortages, causing pickers to deviate from what is normatively assumed. These events underline the importance of supervisory presence on the work floor to guarantee a smooth operating system. Although pickers are not always causing them, these adverse incidents prohibit a consistent and invariable execution of the work planning, thereby contributing to stochasticity in the work execution. It should be noted though that pickers who come across as assertive seem to have more incentives to deviate from predetermined norms. A plausible explanation may be found in people's regulatory focus (Higgins 1998), thereby corroborating the substantial, yet understudied role that psychological effects may have on the order picking process. Moreover, pickers justified the prevalence of maverick picking with occasional stock shortages of products.

It is noteworthy that no research papers have integrated human factors in the replenishment order picking problem. Future research could examine this gap, as well as other events that may induce non-compliant worker behaviour. Finally, we advocate the use of data-driven research approaches such as process mining (Rozinat and van der Aalst 2008). Since people might be subjected to the Hawthorne effect (Boudreau et al. 2003), these methods could be beneficial to investigate any other stochastic incidents that were not raised by interviewees.

3.1.4 Subjective worker experience and judgment

This modelling construct covers the personal perception of workers towards their working environment. Unlike the physical state indicators (section 3.1.2), which focus on objectively measurable metrics, this category comprises subjective and personal experiences, both in physical and mental terms. In this way, concerns raised by Boudreau et al. (2003) are offset, particularly the emotionlessness of human operators and their identical responses to stimuli. Generally, evidence suggests that accounting for workers' well-being promotes productivity and should thus be considered (García-Buades et al. 2020).

Subjective worker experience and judgment in OP literature

Most papers in our sample that used this modelling construct examine the technical equipment planning problem. Therefore, we refer the interested reader to Glock et al. (2021) for an in-depth discussion of these respective papers. The papers use various metrics, either combined or individually, to assess perceived workload, namely the NASA-TLX score, the USQ score, the WCF-scale, the SSQ score, the RSME, user preferences, etc. (Baechler et al. 2016; Baumann et al. 2011; Braam, van Dormolen, and

Frings-Dresen 1996; Chen et al. 2020; Funk et al. 2015; Guo et al. 2014; Kim, Nussbaum, and Gabbard 2019; Kretschmer et al. 2018; Kreutzfeldt, Renker, and Rinkenauer 2019a, 2019b; Lin et al. 2021; Murauer et al. 2018; Reif et al. 2010; Reif and Walch 2008; Scheuermann et al. 2016; Schwerdtfeger et al. 2011; Stockinger et al. 2020; Weaver et al. 2010; Yeow and Goomas 2014). However, the NASA-TLX score is considerably more popular than other metrics. As a matter of fact, the scientific publication that elaborates on its development and usage (Hart and Staveland 1988) has already been cited in over 13,000 studies, which indicates its influence in HF-related research. The reason for the NASA-TLX's popularity among researchers is twofold. First, it is easy to implement and measure. Second, it is multidimensional, as the overarching score is constituted by different subscales, i.e. mental demand, physical demand, temporal demand, performance, effort, and frustration (Cao et al. 2009). This makes the NASA-TLX score a highly detailed, but foremost sophisticated instrument to assess the perceived workload on human operators when performing a given task.

Other papers resort to more advanced psychological concepts. Bräuer and Mazarakis (2019), for example, invoke the three basic psychological needs, autonomy, competency, and relatedness, to determine the level of intrinsic motivation in a gamified environment. Gamification involves using game design elements, e.g. badges and leaderboards, in non-game contexts (Deterding et al. 2011), and is a currently up-and-coming initiative to boost worker motivation (Donohue and Schultz 2019). The authors mention the potentially beneficial effects of gamification, but also warn for the possible negative effects on pickers' motivation in the long run. Passalacqua et al. (2020) use emotional and cognitive engagement measures in their gamification analysis and also raise the concern of (understudied) long-term effects. Another and related measure is work autonomy. Cragg and Loske (2019) demonstrate that pick-by-terminal systems induce higher autonomy levels compared to pick-by-voice systems. The possibility of comfortable inter-colleague communication and lower concentration levels are the main drivers for their results. In order to diminish the high mental strain on order pickers, companies can also aim to reduce the amount of information that they have to process throughout a day. Brynzér and Johansson (1996) focus on storage assignment while keeping picker information as low as possible. Apart from perceptions about their mental workload, pickers also experience subjective workload in physical terms. Steinebach, Wakula, and Mehmedovic (2021) use the Borg RPE scale to assess the subjectively perceived exertion to evaluate the effect of an ergonomic storage location assignment (SLA). Pickers rated the proposed SLA significantly less straining than an SLA that only accounts for travel distance, while cycle times did not differ significantly between the two environments. Larco et al. (2016) use the Borg CR-10, a Category-Ratio (CR) scale, as a predictor of perceived physical workload when a picker picks a product from a specific location. The authors adopt a direct feedback method to obtain input values for the ordinary least squares model that generates discomfort estimates. Results show that a sole focus on minimising discomfort might be detrimental if short cycle times are pursued. Minimising cycle times, on the other hand, is shown to have positive repercussions on pickers' perceived working comfort as well.

In sum, a large number of papers discuss subjective workload. However, several papers raise some critical concerns for future research. First, most studies were carried out in a lab context, which could hamper generalizability. More research is needed to validate findings in an industrial environment, and future studies should also investigate the long-term implications of the current conclusions (Lin et al. 2021; Smith et al. 2021). Second, several papers recognise the plausible impact and enrichment of socio-demographic variables (see section 3.1.5). The quantity of papers that account for interpersonal differences is rather limited. Exceptions include Bräuer and Mazarakis (2019) and Cragg and Loske (2019). Finally, many papers mention that future research should shift its scope towards studies that account for subjective measures as well as objective indicators to provide a more holistic analysis.

Insights from practice

It was clear that order pickers essentially acknowledge the practical relevance of this modelling construct. Human-centric work design is a necessary precondition to boost worker motivation and retention rates. In this light, some pickers (N=12/22) complained about the lack of possibilities to have a say in operational decisions, an aspiration to which many supervisors (N=5/7) are not adverse. Management currently determines break and rotation schedules, order assignments, and employed technical equipment. Consequently, order pickers feel micromanaged and restricted in their autonomy, resulting in poor engagement and low motivation. The feeling of being robotised can cause employees to perform worse than anticipated, most certainly in the long run. Future research should therefore continue to include subjective measures in their models and account for possible worker participation. Utility-driven models (see Larco (2010)) can be a good stepping stone for such an approach. Nevertheless, some pickers (N=7/22) indicated to feel the complete opposite and prefer to engage in activities that require a low degree of participation. This observation supports the idea of a modelling construct that covers discrepancies in pickers' aspirations (section 3.1.5). In addition, we argue that OP planning models might benefit from insights from the occupational psychology research stream (e.g. Bakker and Leiter (2010)). So far, both research streams have been treated independently, which presents numerous future research opportunities. For example, future research could investigate the driving forces behind phenomena such as job satisfaction and motivation and their repercussions on individual performance, specifically tailored to an order picking context.

Finally, despite being beyond the scope of this paper, pickers and supervisors (N=20/29) highlighted the paramount importance of ex post integration of mental well-being. This could include small rewards for consistently meeting a target or the possibility to consult a company psychologist. The compatibility of these initiatives with ex ante design incorporations is, however, currently understudied.

3.1.5 Socio-demographic worker differentiations

This modelling construct is a response to the utopic assumption that all pickers are the same and that calculations are based on an 'average picker' principle. In reality, people differ in terms of physical and demographic characteristics such as length, weight, gender, education level, cultural background, health status, physical abilities, and more. In addition, also social aspects such as personality, values, and attitudes can differ among people (Ashton 2013). In response to upcoming demographic changes (Cohrssen 2021) and evidence of among-worker differences in personality traits (Revelle and Condon 2015), researchers advocate employing models that account for lifelike characteristics (e.g. Juran and Schruben (2004)). The same claim is made by Boudreau et al. (2003), who denounce the practice of portraying workers as identical and emotionless.

Socio-demographic worker differentiations in OP literature

Integrating socio-demographic worker differentiations in a research analysis allows to assess the effects and influences of a particular socio-demographic variable on the researched results. Some of the papers from our literature sample adopt physical characteristics such as body mass and length and highlight the significant impact of these variables on productivity and ergonomics (Al-Araidah et al. 2020; Bräuer and Mazarakis 2019). Hanson et al. (2018) also suggest assigning picking tasks to pickers purely based on height, although they express some ad hoc caution since this practice might be considered discriminatory. Age is another frequently used demographic variable. In light of the prevalent ageing work population (Ilmarinen 2001), considering people's age seems like an absolute must to guarantee accurate decision models. Bräuer and Mazarakis (2019) find evidence that leaderboards in a gamified environment have a stronger impact on relatively young individuals compared to their older co-workers. De Vries, De Koster, and Stam (2016c) also look at the impact of age on the effectiveness of assistive technological equipment and find that older people experience more difficulties when using pick-by-voice systems than their younger colleagues. Other demographic variables that can enrich models and analyses include years of employment (Cragg and Loske 2019) and education level (Hilmola and Tolli 2016).

Apart from typical demographic variables, one can also include social differentiators in research analyses. De Vries, De Koster, and Stam (2016a) invoke the regulatory focus theory and investigate the alignment of order picking methods and incentive systems. According to regulatory focus theory, people are distinguishable between promotion-focused (oriented on achieving one's personal goals) and prevention-focused (oriented on fulfilling one's duties) (Higgins 1998). The authors show that substantial productivity improvements can be achieved if psychological differences among pickers are considered. De Vries, De Koster, and Stam (2016c) reach the same conclusion. They use the Big Five model (Digman 1990) to distinguish individuals based on their personality and find that personality characteristics like neuroticism and extraversion play a significant role in predicting picker behaviour.

Papers that fall under this modelling construct category highlight the capability of socio-demographic variables to enhance the accuracy of operational models. Furthermore, we can infer that socio-demographic variables present a significant overlap or interference with previously discussed modelling constructs and should therefore be considered together.

Insights from practice

Both pickers and their supervisors confirmed the indisputable relevance of socio-demographic-oriented work design. For example, pickers (N=12/22) indulgently agreed with job assignments favouring relatively older people or workers with a certain body type, e.g. lightweight below-chest picks for older and shorter order pickers (Steinebach et al. 2021). Nonetheless, this principle of fairness should not be pushed to its limits. The OP literature has not yet examined the trade-off between favourable assignments and fairness, although it could seriously affect long-term motivation. Another example relates to cultural differences, as one of the interviewees mentioned a Muslim order picker who preferred not to pick products that contained alcohol. Given the cultural diversity in modern society, these kinds of differences should be taken into account.

Finally, the actual differences in people's mindsets stood out during the interviews. Some pickers continuously reflect on possible improvements and want to engage in participatory group discussions. In contract, others confine themselves to executing their jobs and prefer not to be included in additional commitments. These aspirations are a practical manifestation of regulatory focus theory and motivate a work (environment) design tailored to individual pickers' needs. Although such an approach might induce a compatibility challenge (e.g., the harmonizing different technical equipment types), it could boost worker well-being and operational performance.

3.2 Relation to I4.0-to-I5.0 transition

On some topics, interviewees clearly had an opposite view compared to academic literature, while other topics induced consensus. The most prominent points of (dis)agreement are summarised in Table 2, whose last column non-exhaustively relates each row's concepts to the I4.0-to-I5.0 transition. These relations can be grouped into two major application domains of I4.0 within the OP process, i.e. hardware (H) and software (S).

Table 2. Most prominent (dis)agreements among academic literature and interviews.

Modelling construct category	Literature review findings	Interview findings	Conflicting (↔) or affirmative (✓) findings?	Relation to I4.0-to-I5.0 transition
Varying work rates	Distinguishing between people in terms of productivity to assign tasks efficiently.	Conviction that people are more or less the same and fear for monotonous jobs.	↔	The emergence of advanced planning models that aim for operational excellence whilst considering psychosocial considerations. (S)
Varying work rates	Dynamic varying work rates prevail in an order picking environment.	The most dominant type of varying work rates are the dynamic ones.	✓	The possibility for accurate collection of pick data as input for (big) data analytics. (S)
Varying work rates	A narrow set of factors that influence pickers' work rates	Pickers' work pace is affected by several factors, amongst others the prevailing target regime, colleagues, and the accumulation of physical efforts.	↔	The development of smart wearables allows to trace pickers' work pace and physical strain as well as to gamify the working environment. (H)
Quantitative physical state indicators	Extensive scientific research on physical state of employees.	Pronounced consciousness regarding the physically demanding nature of order picking.	✓	Research on and introduction of monitoring devices to assess workers' physical state as well as assistive devices to support physically demanding activities. (H)
Stochastic worker behaviour and work execution	Very little research efforts on picker behaviour which deviates from normativity. In consequence, few papers include such phenomena.	Supervisors emphasise the frequent occurrence of non-compliant worker behaviour.	↔	The availability of advanced information systems which generate accurate time stamps as input for the detection of deviations from normativity. (S)
Subjective worker experience and judgment	Decision-making is assumed to be top down and centralised.	Feeling of being neglected in operational decisions can result in poor motivation and decrements in operational performance.	↔	The configuration of, for example, technical equipment in consultation with pickers to account for their well-being. (H)
Socio-demographic worker differentiations	Assignments and decisions without considering plausible repercussions on pickers' attitude.	Indulgent attitude towards job assignments that favour certain socio-demographic characteristics.	↔	The availability of collaborative robots makes it possible to redistribute tasks in response to the workforce's characteristics. (H)

I4.0 hardware encompasses the physical tools and devices that either support or substitute human labour during the OP process (Winkelhaus, Grosse, and Morana 2021). Examples include smart wearables, monitoring devices, exoskeletons, and cobots. Although these devices might increase the trackability of and reduce the physical impact on humans, their deployment may result in motivational issues (Lager, Virgillito, and Buchberger 2021). Acceptance is a fine example of the "subjective worker experience and judgement" construct category which ought to be considered when deciding upon the acquisition or configuration of new devices. The usage of a technology acceptance model, like TAM, can provide decision makers additional insights in why and how I4.0 hardware will (not) lead to proper user adoption (e.g. Manis and Choi (2019)). Shifting from I4.0 towards I5.0 necessitates research that pays equal attention to physical and psychological human aspects, as to operational outcomes when judging I4.0 hardware.

On the other hand, I4.0 software involves the intangible collection of procedures and instructions which underlie the configuration of an OP system, for example information systems and planning tools. In comparison with I4.0 hardware, I4.0 software is not focused on technological devices as such, but rather on how technology can be used for planning and analytical purposes. With the advent of I4.0, automation and digitalisation have found their way to many warehouses and increase the connectivity and trackability of entities. This might induce another challenge for many warehouse managers, as the data multitude increases significantly. However, with a proper data management system, this is not necessarily a burden but rather a plausible competitive advantage. Given that a system's output is only as qualitative as its inputs, a plurality in data might generate more sophisticated planning models. For example, due to increased connectivity, dynamic varying work rates can be captured more accurately to serve as enriched input for planning models that require such information for break and shift rotation schedules. Future research could therefore increase the physical and mental well-being of workers if planning models do take advantage of this increased data provision. A holistic analysis which also accounts for investment costs and the impact of more accurate decision support models is deemed a fruitful research avenue, e.g. Mocan and Draghici (2018).

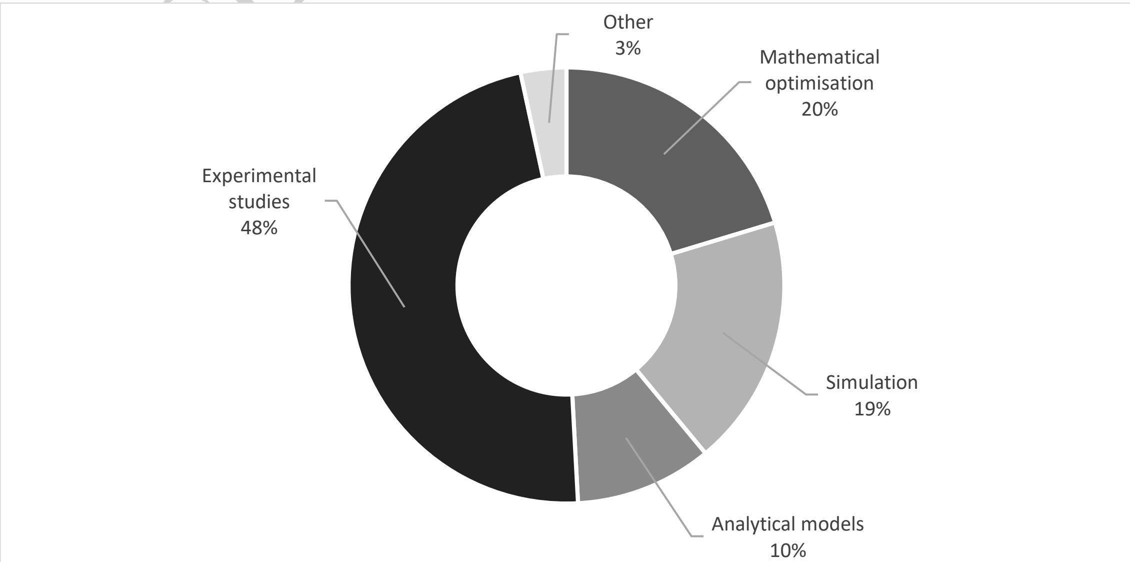
3.3 Research methodologies

This section elaborates on the research methodologies used to analyse and solve human-centric OP planning problems. The following five research approaches are frequently employed in the literature and are discussed in the remainder of this section.

- Mathematical optimisation
- Simulation
- Analytical models
- Experimental studies
- Other

Figure 5 depicts the respective shares of these methods in the literature sample. Note that a single paper can be assigned to more than one category. Experimental studies are most prominently present, followed by mathematical optimisation and simulation.

Figure 5. Research methodologies used for analysing order picking planning problems.



Mathematical optimisation refers to the ensemble of formalised mathematical expressions which aim to minimise or maximise a certain objective function subject to a set of constraints. This research methodology allows researchers to include expressions that capture problem-specific dimensions such as behavioural processes. In terms of modelling choices, authors can either optimise for one single objective, e.g. Diefenbach and Glock (2019), Glock et al. (2019) and Matusiak, De Koster, and Saarinen (2017), or decide to invoke more than one objective function, e.g. Battini et al. (2016), Larco et al. (2016) and Zhao et al. (2019). The latter, often called bi- or multi-objective optimisation, is a popular research methodology when studying possible trade-offs between seemingly conflicting objectives (Deb and Deb 2014). Moreover, it is up to the researcher to decide whether the objective function is directed towards economic (e.g. Al-Araidah et al. (2017)) or ergonomic (e.g. Otto et al. (2017)) optimisation. Despite the highly convenient flexibility in terms of modelling choices, solving these self-configured mathematical optimisation models can become increasingly time-consuming in terms of computational time if an exact solution approach is pursued. Therefore, many research papers resort to metaheuristic algorithms to generate a reasonably good feasible solution within an acceptable computing time (Glover and Sörensen 2015). The popularity of a metaheuristic approach lies in its multi-applicability and potency to rapidly generate feasible solutions that do not deviate significantly from the optimal one. In this light, metaheuristics are highly suitable for solving operational OP planning problems, as some of these have to be addressed frequently throughout the day. Which metaheuristic framework should be used is food for debate (Glover and Sörensen 2015) and may depend on the research context. Metaheuristic algorithms (MHs) that are used by papers in our literature sample include local search MHs (Al-Araidah et al. 2017; Glock, Grosse, Abedinnia, et al. 2019; Otto et al. 2017) and population-based MHs (Feng and Hu 2021;

Zhao et al. 2019). Table 3 gives an overview of the papers that apply exact and heuristic solution approaches.

Table 3. Overview of the solution approach(es) per paper that used mathematical optimisation.

	Exact	Heuristic
Battini et al. (2016)	•	
Larco et al. (2016)	•	•
Al-Araidah et al. (2017)		•
Matusiak, De Koster, and Saarinen (2017)		•
Otto et al. (2017)		•
Diefenbach and Glock (2019)	•	
Glock et al. (2019)		•
Zangaro et al. (2019)	•	
Zhang et al. (2019)		•
Zhao et al. (2019)		•
Feng and Hu (2021)		•
Gajšek et al. (2021)	•	

Simulation refers to a wide range of methods and applications to imitate the inherent characteristics and behaviour of real-life systems (Kelton 2002). The main purpose of simulation is to conduct numerical experiments to understand the actual behaviour of a system and its constituting components. For instance, an order picking system and the order pickers can be modelled such that various operational scenarios can be tested and analysed before actual implementation. However, although simulation is a powerful tool to mimic and provide insights into the true dynamics of a system, it requires a considerable modelling and scenario development time, depends heavily on the accuracy of input data, and is only generalisable to a limited extent (Feinstein and Cannon 2001; Vanbrabant et al. 2019). Different types of simulations can be used to investigate operational policies. For example, discrete-event simulation (DES) is a well-established process-oriented type of simulation where a system is modelled as a network of activities and queues. Agent-based simulation (ABS), on the other hand, is a rather up-and-coming approach focusing on self-governing individual agents (Maidstone 2012). ABS is deemed highly suitable to model the specific behaviour of and interactions between agents, whereas DES is more top-down and drawn to modelling the system in detail (Siebers et al. 2010). The ultimate choice between these two types is, however, entirely up to the researcher. Another simulation type is Monte Carlo simulation, which relies on repeated random sampling to provide insights if outcome scenarios are subjected to a degree of uncertainty (Raychaudhuri 2008). Note that system dynamics, another simulation type, is not included in the discussion since none of the papers in our sample uses this method. Table 4 summarises the research papers which use simulation.

Table 4. Overview of research papers using simulation models.

	Agent-based simulation	Discrete-event simulation	Monte Carlo simulation
Koo (2009)		•	
Webster, Ruben, and Yang (2012)		•	
Hong, Johnson, and Peters (2015)		•	
Battini et al. (2016)			•
Elbert et al. (2017)	•		
Elbert and Müller (2017)	•		
Hong (2018)		•	
Batt and Gallino (2019)			•
Fibrianto and Hong (2019)		•	
Granotto et al. (2019)		•	
Al-Araidah et al. (2020)			•

Analytical models involve a set of interconnected mathematical expressions which approximate complex systems' actual behaviour by linking the outcome variable of interest to several system parameters (Van Gils et al. 2018). This research approach can predict and evaluate the system's performance under different operational scenarios solely by determining ranges of possible parameter values. Although analytical models are less time-consuming to develop than simulation models and fairly intuitive to evaluate, they also risk being simplified representations of real operational systems. In response, a deliberate trade-off should be made regarding the degree of complexity and related representation of reality. In order picking literature, papers that integrate ergonomics and economics using analytical models are still scarce, although more researchers have recently started to employ this research methodology (Calzavara et al. 2019). Sgarbossa et al. (2020) point out that mostly physical-related human factors were considered in analytical models. However, apart from physical considerations (e.g. Calzavara et al. (2017) and Calzavara et al. (2019)), these models also include learning-related phenomena (Grosse, Glock, and Jaber 2013; Grosse and Glock 2015) and differences in skill levels (Webster, Ruben, and Yang 2012). Although the employment of analytical models is currently not very popular, it does not imply that the quality of its results are inferior to those of other research methodologies (Meller and Klote 2004).

Most research methodologies pertain to experimental studies. In this research methodology, subjects are typically assigned to treatments, either in an industrial or a lab environment, so that the researcher can make inferences about the relationship between independent and dependent variables (McClave and Sincich 2007). Experimental studies are highly suitable for investigating the impact of human behaviour on operational outcomes and are often supplemented with surveys regarding subjects' retrospective perceived experiences (Bendoly et al. 2010). There are various statistical analysis procedures (SAPs) to convert collected data into usable information. Table 5 presents an overview of these commonly used procedures and the respective papers in which they are applied. Correlation analysis is defined as an in-depth analysis and discussion of correlations between two or more variables in the data set. Frequently used metrics are the Pearson, Spearman's rank, and point-biserial correlation coefficient. Regression analysis is the statistical procedure of estimating the relationship between independent and dependent variables. Hypothesis testing is the most frequently used SAP and is primarily used for establishing (non-)significant disparities in outcome variables among treatments. Lastly, other SAP refers to procedures such as principal component analysis, effect-size evaluation, and descriptive statistics comparisons.

Table 5. Overview of statistical analysis procedures (SAPs) and papers in which they are used.

Statistical tool	Number of research papers	Papers
Correlation analysis	5	Bräuer and Mazarakis (2019); Cragg and Loske (2019); De Vries, De Koster, and Stam (2016a, 2016c); Kreutzfeldt, Renker, and Rinkenauer (2019b)
Regression analysis	3	De Vries, De Koster, and Stam (2016a, 2016c); Passalacqua et al. (2020)
Hypothesis testing	25	Baechler et al. (2016); Baumann et al. (2011); Bräuer and Mazarakis (2019); Chen et al. (2020); De Vries, De Koster, and Stam (2016a, 2016c); Funk et al. (2015); Guo et al. (2014); Hanson et al. (2016); Hanson et al. (2018); Hilmola and Tolli (2016); Kim, Nussbaum, and Gabbard (2019); Kretschmer et al. (2018); Kreutzfeldt, Renker, and Rinkenauer (2019a, 2019b); Lee, Chang, and Karwowski (2020); Lin et al. (2021); Murauer et al. (2018); Passalacqua et al. (2020); Reif et al. (2010); Reif and Walch (2008); Schwerdtfeger et al. (2011); Steinebach, Wakula, and Mehmedovic (2021); Stockinger et al. (2020); Weaver et al. (2010)
Other SAP	4	Chen et al. (2020); De Vries, De Koster, and Stam (2016a); Scheuermann et al. (2016); Yeow and Goomas (2014)

Finally, the 'other' research methodology category comprises case studies (Braam, van Dormolen, and Frings-Dresen 1996) and conceptual models (Brynzér and Johansson 1996). It is worth mentioning that some methodologies are rarely used to study human factors in order picking planning problems. These methods include queueing models, fuzzy set theory, discrete choice modelling and data or process mining.

4. Discussion

Section 3 discussed five hands-on modelling constructs to enrich the pivotal decision-making process in order picking systems. In addition, we presented some leading research methodologies in which these modelling constructs can be embedded. Throughout their review, academic practices were cross-checked with insights from practice to assess the modelling approaches from a practitioners' point of view. This has led to some recommendations for future research. On top of that, this section provides an integrated synthesis of the proceedings mentioned above with a threefold perspective. First, Section 4.1 compares our general findings with those of previously developed frameworks. Second, Section 4.2 elaborates on the degree of compatibility between modelling constructs and research methodologies. Finally, we holistically evaluate the extent to which the currently used modelling fails to comply with practitioners' concerns and aspirations and how this gap should be addressed (Section 4.3).

4.1 General findings comparison

Previous literature reviews related to the integration of human factors into order picking planning classify papers based on the generic HF categories, namely physical, perceptual, mental, psychosocial, and skills. Although we adopt another perspective to investigate the same research topic, it should be noted that our article sample's constitution is in line with recent findings by Sgarbossa et al. (2020) and Vanheusden et al. (2022). Specially, it confirms that previous research papers have mostly focused on the physical HF to mitigate the negative externalities of picking orders, e.g. the prevention of musculoskeletal disorders. Consequently, papers that focus on other HF categories lag behind in terms of quantity, thereby leaving aspects of the order picking process underexploited. However, our literature sample comprises many papers that account for the perceptual human factor (Appendix D). This is due to the prominent presence of papers that study the technical equipment planning problem, which is intrinsically connected with information processing and thus perceptual considerations. Its notable presence is mainly driven by the growing societal interest of exploiting the supportive capabilities of assistive devices in manual material handling. The recent literature review of Glock et al. (2021) about this specific topic strengthens this supposition. Furthermore, our sample is also characterised by a high degree of disparity between the number of papers that research respective OP planning problems (Appendix D). The number of papers dealing with problems such as storage (re-)assignment, assistive technical equipment (e.g. pick-by-voice or pick-by-vision), and zoning is a multiple of those that consider job assignment, batching, warehouse lay-out design, etc. While the aforementioned observations and suggestions on how to tackle them is briefly discussed in Vanheusden et al. (2022), our review succeeds in making more concrete methodological research propositions, partly originating from practitioners. In this way, the article at hand proposes concrete suggestions and guidelines to facilitate the development of human-centric OP systems.

Finally, we would also like to highlight the uniqueness of our taxonomy compared to the initial framework of Grosse et al. (2015). The classification categories that are being proposed in the paper at hand do not exhibit a one-to-one relation with the generic HF categories presented in Grosse et al. (2015) or Vanheusden et al. (2022). An illustration is the varying work rates construct, which can be linked up to three of the generic HF categories, namely physical, mental, and skills. Conversely, the perceptual category proposed by Grosse et al. (2015) is not easily convertible to one of our integrative constructs. Therefore, both research perspectives should not be seen as contesting but rather as supplementary and possibly synergic in integrating human factors in planning problems.

In addition, since we initiate the analysis with a HF-perspective, our approach varies from those adopted by prior articles about I4.0, which dominantly adopt a technology-oriented approach. Nonetheless, both approaches reach the same conclusion, namely the prevalent disconnection between the I4.0-to-I5.0 transition and rigorous human factors consideration. Therefore, the paper at hand should be seen as an enriching and complementary analysis.

4.2 Compatibility assessment

As already established, papers that adopt modelling constructs in the decision-making process use a wide range of research methodologies. However, past papers have not studied the correlation between the adopted modelling construct and the used research methodologies. Such studies could uncover prevalent patterns and assess the suitability and popularity of modelling construct-research methodology combinations. Table 6 shows how frequently a particular modelling construct has been adopted per research methodology in the reviewed papers. Note that one single research paper can have multiple methodologies and/or more than one adopted modelling construct. Details can be found in Appendix B.

Table 6. Frequency of modelling construct-research methodology combinations in our review.

	Mathematical optimisation	Simulation	Analytical models	Experimental studies	Other	Total
Varying work rates (VWRs)	4	8	4	-	-	**16**
Quantitative physical state indicators	8	3	2	5	1	**19**
Stochastic worker behaviour and work execution	-	1	1	2	-	**4**
Subjective worker experience and judgment	1	-	-	22	2	**25**
Socio-demographic worker differentiations	-	1	-	6	-	**7**
Total	**13**	**13**	**7**	**35**	**3**	**71**

Mathematical optimisation models commonly appear in combination with VWRs or quantitative physical state indicators. Most of these combinations are somehow associated with deterministic physical metrics, like energy expenditure or differences in work rates among pickers. One of the underlying reasons for this predominance can be found in the representational nature of mathematical optimisation models. Specifically, this research methodology is less capable of appropriating stochastic dynamics of a particular problem context (Chandra and Grabis 2007). In addition, only the study of Larco et al. (2016) integrates aspects related to pickers' work perception. The latter example proves that mathematical optimisation models should not be considered prohibitive for mental- and psychosocial-related phenomena. We rather consider the scarcity of works of this particular combination as a fruitful yet underexploited research avenue.

Studies that adopt simulation techniques predominantly focus on VWRs and quantitative physical state indicators and currently omit subjective worker experience and judgment. The historical development and use of simulation in OR modelling may lie at the base of this underrepresentation. Namely, discrete-event simulation has been the most prominent simulation type until recently (Siebers et al. 2010) (see Table 4). As discussed earlier, this simulation type is oriented towards systemic process modelling, rather than a detailed examination of agents within the system. With the emergence of agent-based simulation, circumstances are considerably more favourable to account for phenomena that are covered by subjective worker experience and judgment, and stochastic worker behaviour and work execution. Additionally, simulation is deemed extremely suitable to cope with stochasticity (Kelton 2002). This premise manifests itself in the observation that more than half of the research papers that use simulation include stochastic VWRs. Due to their inherent ability to deal with complex problem settings, simulation models are a particularly interesting research methodology to adopt multiple modelling constructs simultaneously.

Analogous to mathematical optimisation models, analytical models frequently resort to VWRs or quantitative physical state indicators. Since this research methodology exhibits characteristics from both mathematical optimisation models and simulation, a slight degree of stochasticity can be embedded in the model design. However, due to the limited number of research papers that use this and the ‘other’ research methodology, it seems too premature to draw definitive conclusions.

Finally, experimental studies principally invoke subjective worker experience and judgment, and socio-demographic worker differentiations. The possibility to conduct post-experimental picker surveys facilitates the integration of individualised worker perceptions. Unfortunately, empirically observed phenomena are seldom converted into expressions that can be used by other research methodologies. This finding is in line with earlier calls from Bendoly et al. (2010), who advocate converting field-based empirical observations into mathematical expressions. Despite being challenging, we are convinced that this practice will help to refine and improve future models. An excellent example is the study of Grosse et al. (2013), in which the empirical findings of Grosse and Glock (2013) are translated into mathematical expressions to serve in an analytical model. Recovery rates and mental state-dependent work pace are an illustration of relevant topics for which such approach is currently non-existent in OP literature.

Apart from methodology-specific considerations, future research should also account for the following overarching remarks. The first pertains to the practice of incautiously extrapolating short-term experimental findings. Many papers explicitly highlight the short time frame of their analysis as a limiting factor of their research (e.g. experiments of 10 minutes). In consequence, long-term implications are in many cases not accounted for, although they could severely interfere with initially obtained research outcomes. Despite their short-term planning horizon, even operational decisions should consider long-term repercussions (e.g. repetitiveness, boredom…). A strategy that sequentially adopts short-term optimality may not end up in a long-term global optimum. Additional research should therefore be tailored to these long-term implications and examine how they interfere with work performance. This partially brings us to our second remark. More precisely, although picking time might be an appropriate economic performance measures in the short run, it can be argued that other metrics are at least equally important in the long run. The performance of a warehouse can be assessed by productivity-, quality-, time-, cost- and safety-related indicators (De Koster, Stam, and Balk 2011; Staudt et al. 2015). Very few of the papers in our literature sample holistically evaluate system performance using multiple performance indicators. Data envelopment analysis (DEA) is a fine example of a method that could be invoked for such an evaluation. Its inherent ability to account for different indicators makes DEA an appropriate and widely accepted method to assess the performance of warehouses or their order picking system (De Vries, De Koster, and Stam 2016b; Klumpp and Loske 2021). In addition, the reviewed papers in our sample hardly include feedback links between human conduct and these indicators (e.g. poor working conditions induce absenteeism, which in turn causes significant operational losses (Kocakulah et al. 2016)). We, therefore, recommend well-thought-out performance measures which additionally capture the long-term impacts of short-term decisions.

4.3 Theory-practice gap

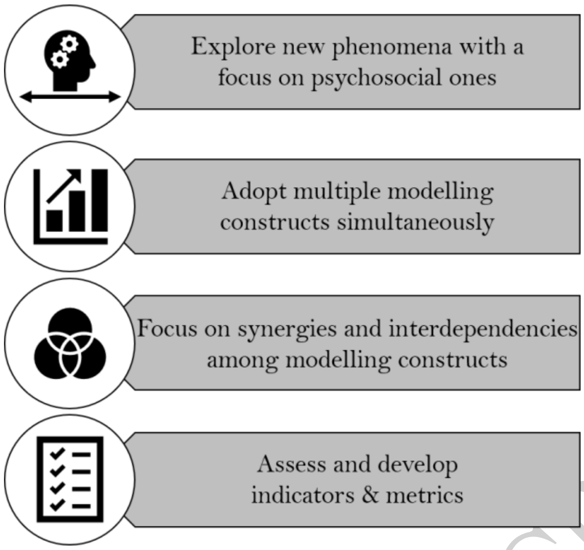
The modelling constructs presented in Section 3.1 offer researchers the opportunity to shift away from incorrect axiomatic modelling assumptions. Nonetheless, evidence from interviews suggests that they still necessitate a thorough development process in order to capture all the facets and dynamics of an OP system with the ultimate purpose of improving operational performance and/or worker well-being. Throughout the analysis of current academic literature and the conducted interviews, four fundamental focal points kept recurring and embody generic recommendations that could serve to bridge this theory-practice gap. They are discussed in the remainder of this section and are summarised in Figure 6.

By adopting the discussed modelling constructs, the veracity and consequently the usefulness of prevailing policy choices have been considerably reinforced. However, previous academic research efforts have predominantly been dedicated to the physical human factor, and to a much lesser extent to the psychosocial and mental one (Neumann and Dul 2010; Sgarbossa et al. 2020). The discrepancy between the present state-of-the-art and pickers' concerns is at the very least remarkable. More precisely, during the interviews, pickers repeatedly emphasised the significance of psychological and mental aspects on the work floor for their well-being. In addition, it has already been shown that psychosocial strains, such as high perceived work load and time-pressure, can cause musculoskeletal symptoms (Bongers et al. 1993). This finding results in a first generic recommendation, i.e. *future research should increasingly pursue to embrace and thoroughly investigate the impacts of psychosocial phenomena in an order picking environment*. As mentioned before, spanning this gap involves cross-disciplinary research with, for example, occupational and work psychology. Qualitative and experimental studies could be of great use to explore phenomena that simultaneously impact pickers' prosperity and operational performance. Even more important and enriching for this research field would be the rigorous adoption of findings from such studies in advanced mathematical optimisation, analytical or simulation models.

It should also be noted that the presented modelling constructs are currently sequentially discussed in Section 3.1. Accordingly, this might induce the conviction that they should be treated independently. However, interviews revealed that the modelling constructs should be considered interconnected as well as not mutually exclusive. A first clarification relates to their mutual non-exclusiveness. It became clear throughout the interview sessions that all of the modelling constructs entail the inherent representational capacity to reflect real-life phenomena in academic research. In order to pursue productivity and/or well-being improvements, models could benefit from a 'the more, the better' principle in terms of modelling constructs. The more modelling constructs, the better a model can capture the true dynamics in a warehouse. However, the number of research papers that adopt more than one modelling construct is rather limited. Therefore, we propose a second generic recommendation: *future research should apply multiple modelling constructs simultaneously to avoid decision-making based on incomplete information*.

The aforementioned interconnectedness of modelling constructs also deserves some attention. Particularly, feedback links between separate modelling constructs may severely impact their combined effect (e.g. Asadayoobi, Jaber, and Taghipour (2021)). Interviewees provided valuable examples to justify this premise. An example is the impact of socio-demographic worker differentiations on quantitative physical state indicators, as older order pickers have a different recovery rate compared to their younger colleagues. Another example relates to motivation-dependent work rates, which integrate varying work rates and subjective worker experience. Ample of these illustrations were raised by interviewees, which substantiates a third recommendation: *future research should focus on synergies and interdependencies among modelling constructs*. Besides, each modelling construct holds a plurality of metrics and parameters that can be applied (e.g. Takala et al. (2010)). However, the order picking research stream still lacks a comprehensive work that rigorously outlines relevant metrics and evaluates their context-dependent adequacy. Therefore, we propose a fourth recommendation: *future research should thoroughly assess and develop indicators and metrics used in decision support models*.

Figure 6. Generic recommendations for future research.



Finally, we end this section with a selection of concerns which were frequently raised by order pickers or their supervisors (see Appendix E). These could be a fine addition to some of the modelling constructs and aid in the I4.0-to-I5.0 transition, as they were claimed to influence operational performance and picker well-being. Moreover, it is noteworthy that most of these concerns have already been raised as future research opportunities by previous reviews and research articles. Despite their proven relevance and potential fruitfulness, they remain rather understudied to date, at least in an OP context.

A first concern is related to individual preferences and picker participation in the decision-making process. On several occasions, interviewees expressed the conviction that accounting for their individual preferences could indeed improve motivation. A slight deviation from normatively optimal decisions in terms of productivity may significantly increase pickers' motivation level, resulting in higher overall performance compared to the normative case. Since I4.0, planning models have the possibility to account for a wider and more accurate spectrum of input data. Consequently, detailed picker preferences can serve as additional input to support operational decisions. Converting preferences into modelling constructs could come from a utility-underpinned research method (Larco 2010), thereby contributing to our first and fourth recommendations. For example, the use of discrete choice experiments can be an excellent method to rigorously derive quantitative measures of people's aspirations, e.g. Schuir and Teuteberg (2021). Once the preferences of individual pickers are elicited, a mathematical optimisation model can be developed in which worker utility (as a proxy for work satisfaction) and/or economic performance are maximised. Frameworks for the integration of discrete choice models in mixed-integer already exist (e.g. Pacheco et al. 2021), albeit not tailored to an industrial context like order picking. Heterogeneity in picker preferences, and the resulting impact on the subjective worker well-being modelling construct(s) as a result of centralised decision-making, can be captured by for example a random parameter (or mixed) logit model. In addition, DCEs are conventionally enriched with a socio-demographic questionnaire. It would be interesting to investigate whether preference patterns can be discovered among the workforce and to what extent these patterns could contribute to WH planning. The above-discussed approach can be applied to many decision problems, such as the job assignment planning problem since it is very likely that pickers have diverse preferences for certain order attributes.

A second aspect that frequently recurred involves the timely provision of feedback on picker performance. I4.0 hardware tools, such as smart watches, are useful to provide hands-on real-time feedback. Studies have already demonstrated the potential benefits of accurate real-time feedback systems (Zhang et al. 2021). Despite the observation that not all pickers are proponents of immediate feedback provision, companies should contemplate its implantation, as it could influence work-related stress perceptions, thereby contributing to the I4.0-to-I5.0 transition. Experimental studies and use cases are suitable to configure an appropriate individualised user-interface (Schwerdtfeger et al.

2011). In this way, pickers can have tailor-made tools at their disposal to optimise for motivational and work rate-related impacts. The experimentally quantified repercussions of feedback provision can be integrated in simulation or mathematical optimisation models in the form of work rates that depend on relative performance. For instance, an agent-based simulation model would be highly convenient to incorporate dynamic work rates that depend on the frequency and content of real-life feedback reports. Based on the ad hoc circumstances, the ex-ante optimal order-to-picker assignments might underperform in practice and require adjustments. Another example pertains to mathematical optimisation models, where the processing time of an operator can be corrected for his expected remaining efforts for attaining the daily target. In studies that examine the feedback process, varying work rates and socio-demographic worker differentiations should be carefully evaluated as an addition to subjective worker experience and judgement modelling constructs. This is linked to our second and third generic recommendations.

The third concern relates to so-called 'higher-level' human factors such as teamwork and leadership and are greatly important for worker well-being. For example, studies show that the importance of a competent team leader cannot be overstated for a safe and efficiently operating warehouse (De Vries, De Koster, and Stam 2016b; Lambrechts et al. 2021). An interesting study that addresses this topic responds to earlier claims from Elbert et al. (2017). The authors infer that measures to mitigate route deviation probabilities are worthy of investigation. An experimental study could shed a new light on the respective matter. In particular, a factorial design regarding supervisory presence (e.g. limited, moderate, or continually) and style (e.g. authoritarian, coaching, or pacesetting) could be used to evaluate the appropriate supervision strategy for a given routing policy. The study's investigated modelling constructs can be manifold, for example stochastic worker behaviour (e.g. route deviation probabilities), subjective worker experience and judgement (e.g. job control and monitoring acceptance measures), and varying work rates (e.g. supervisory-dependent work rates). As a consequence of their explorative nature, case studies and qualitative studies are deemed appropriate to unveil other relevant 'higher-level' HFs, their psychological impact, and interdependencies with other modelling constructs. This directly responds to our first and fourth recommendations.

A final concern is the occurrence of boredom as a result of monotonous work days. Boredom can be associated with accidents, decreased job satisfaction and reduced performance (Azizi, Zolfaghari, and Liang 2010). Mathematical expressions for boredom have, however, not yet found their way to research on OP, despite its repetitive nature. For example, minimising pickers' boredom when assigning workers to WH zones would be an effective manner to improve worker well-being. Insights from the boredom modelling research stream can be a good starting point (Azizi, Liang, and Zolfaghari 2013). Interviewees suggested a well-designed break and job rotation schedule as one of the plausible means to mitigate work monotony and related boredom. When pursuing an accurate job rotation schedule, it is essential to account for a wide range of factors, including intervening in the learning process, preventing boredom, broadening pickers' skill set, and many more. This directly relates to our second recommendation, and is feasible as a result of advanced I4.0 technological developments. Although it could be claimed that break and job rotation schedules have to be stipulated at an aggregate level (Grosse et al. 2015), they could also be integrated with specific OP planning problems. As an illustration, the replenishment planning problem is currently not linked to considerations with regards to break and job rotation schedules. Replenishing is often restricted to moments at which no orders are being picked (De Koster, Le-Duc, and

Roodbergen 2007), implying that some workers might have a break. The timing and the duration of these breaks could be aligned with the replenishment operations so that the required amount of goods is at its respective storage location at the start of the next picking shift. Due to new I4.0 technologies and resulting information availability, accurately tracking WH inventory may contribute to a lean OP system. Hence, the ideal proportion of replenishers to pickers, as well as the break scheduling could be studied using an inventory holding cost minimisation model. Dedicatedly committing resources in a harmonised way could decrease stock shortages, picker congestion, and resulting pick time increments and frustrations. Mathematical optimisation and simulation models have in the past proven to be excellent research methodologies to cope with this type of research (Allwood and Lee 2004; Azizi, Zolfaghari, and Liang 2010). Although the list of proposed concerns is not exhaustive, it outlines some important features which are currently not sufficiently addressed in order picking literature in respect to the I4.0-to-I5.0 transition. More and similar modelling requirements could be uncovered by additional qualitative research (Grosse et al. 2016) and data-driven approaches such as process mining. This contributes to the first generic recommendation, namely *the exploration of new HF-related phenomena that significantly influence operator performance and well-being*.

5. Conclusion

Human influences on order picking systems have largely been ignored in the past. However, most recently, this research stream has gained momentum. Researchers have started to recognise the distinctive characteristics of human operators and have engaged in designing systems that are specifically tailored to them, partly driven by the expedient transition from Industry 4.0 towards Industry 5.0. Despite the indisputable link between this transition and human factors, few studies exist that holistically look at both concepts simultaneously. Prior review articles identified the most prominent human factors categories and examined their prevalence in planning problems. Our review proposes a concrete toolset of five modelling constructs (varying work rates, quantitative physical state indicators, stochastic worker behaviour and work execution, subjective worker experience and judgement, and socio-demographic worker differentiations) that can be used to humanise order picking planning problems. We adopt a multimethod approach, constituted by a profound literature review and semi-structured interviews, to establish these modelling constructs and to assess their compatibility with leading research methodologies. Throughout the discussion, we cross-check academic practices with practitioners' concerns, thereby guaranteeing external validity and providing a multitude of relevant research opportunities. In addition, we address why and how future I4.0-to-I5.0 research might benefit from the proposed modelling constructs. Results suggest that the elementary human-related phenomena which are embedded in research papers still necessitate a thorough development process, both in depth and in breadth. Academic research has insufficiently comprehended and incorporated human behaviour in operational models.

The findings of this study have to be seen in the light of some limitations. For example, our literature sample consists of 57 papers, which can be considered as limited. Nonetheless, this relatively small number should trigger awareness that this research topic is under-studied and requires more attention. In addition, we added a second layer of information, i.e. expert interviews, to detect relevant but currently missing aspects within this field. The sampling strategy for these interviews might pose a second limitation. Namely, interviewees were employed at large multinational companies with relatively advanced order picking systems. Hence, concerns that might arise in smaller enterprises could remain undetected. Despite this seemingly biased sample, we argue that our current sample scope is indeed the most adequate, since our focus is on improving the current state-of-the-art.

This study presents several valuable managerial insights. Researchers and warehouse managers are aware that persistently holding on to unrealistic assumptions about human operators biases the operational decision-making process. However, a set of holistic principles to rigorously integrate human factors in order picking planning was still missing. The present study lays the foundations to facilitate the process of designing human-centric order picking systems and frames them in the context of the I4.0-to-I5.0 transition. Although the proposed modelling constructs should not be considered as deeply exhaustive, they provide a convenient toolset to support human factors integration and a large set of practice-based research opportunities. Furthermore, the insights that emerged from our analysis can be applied to various (industrial) environments. More precisely, findings are not only restricted to a warehousing context, but could easily be translated to other research domains, e.g. manufacturing or luggage handling. This review thus helps to promote and facilitate the integration of human factors in generic planning models. Future models will be more usable in practice and enhance working conditions for employees that work within configured systems. We hope that the constructs discussed in this paper serve as inspiration for a more accurate system design in pursuit of Industry 5.0.

6. Acknowledgements

This research is supported by the Special Research Fund (BOF) of Hasselt University and the Research Foundation Flanders (FWO) (Strategic Basic Research project Data-driven logistics, FWO-S007318N).

7. Data availability statement

The data that support the findings of this study are available on request from the corresponding author, T. De Lombaert. The data are not publicly available due to research collaboration agreements with cooperating companies.

8. References

Al-Araidah, Omar, Doraid Dalalah, Ma'en Al-Abdel Azeez, and Mohammad T. Khasawneh. 2017. 'A Heuristic for Clustering and Picking Small Items Considering Safe Reach of the Order Picker'. *European Journal of Industrial Engineering* 11 (2): 256–69. https://doi.org/10.1504/EJIE.2017.083256.

Al-Araidah, Omar, Gul Okudan-Kremer, Elif Elcin Gunay, and Chih-Yuan Chu. 2020. 'A Monte Carlo Simulation to Estimate Fatigue Allowance for Female Order Pickers in High Traffic Manual Picking Systems'. *International Journal of Production Research* 0 (0): 1–12. https://doi.org/10.1080/00207543.2020.1770357.

Allwood, J. M., and W. L. Lee. 2004. 'The Impact of Job Rotation on Problem Solving Skills'. *International Journal of Production Research* 42 (5): 865–81. https://doi.org/10.1080/00207540310001631566.

Asadayoobi, N., M. Y. Jaber, and S. Taghipour. 2021. 'A New Learning Curve with Fatigue-Dependent Learning Rate'. *Applied Mathematical Modelling* 93 (May): 644–56. https://doi.org/10.1016/j.apm.2020.12.005.

Ashton, Michael C. 2013. *Individual Differences and Personality*. Academic Press.

Attia, El-Awady, Philippe Duquenne, and Jean-Marc Le-Lann. 2014. 'Considering Skills Evolutions in Multi-Skilled Workforce Allocation with Flexible Working Hours'. *International Journal of Production Research* 52 (15): 4548–73. https://doi.org/10.1080/00207543.2013.877613.

Azadeh, Kaveh, René De Koster, and Debjit Roy. 2019. 'Robotized and Automated Warehouse Systems: Review and Recent Developments'. *Transportation Science* 53 (4): 917–45. https://doi.org/10.1287/trsc.2018.0873.

Azizi, Nader, Ming Liang, and Saeed Zolfaghari. 2013. 'Modelling Human Boredom at Work: Mathematical Formulations and a Probabilistic Framework'. *Journal of Manufacturing Technology Management* 24 (5): 711–46. https://doi.org/10.1108/17410381311327981.

Azizi, Nader, Saeed Zolfaghari, and Ming Liang. 2010. 'Modeling Job Rotation in Manufacturing Systems: The Study of Employee's Boredom and Skill Variations'. *International Journal of Production Economics* 123 (1): 69–85. https://doi.org/10.1016/j.ijpe.2009.07.010.

Baechler, Andreas, Liane Baechler, Sven Autenrieth, Peter Kurtz, Thomas Hoerz, Thomas Heidenreich, and Georg Kruell. 2016. 'A Comparative Study of an Assistance System for Manual Order Picking – Called Pick-by-Projection – with the Guiding Systems Pick-by-Paper, Pick-by-Light and Pick-by-Display'. In *2016 49th Hawaii International Conference on System Sciences (HICSS)*, 523–31. https://doi.org/10.1109/HICSS.2016.72.

Bakker, Arnold B., and Michael P. Leiter. 2010. *Work Engagement: A Handbook of Essential Theory and Research*. Psychology Press.

Bartholdi, John J., and Donald D. Eisenstein. 1996. ‘A Production Line That Balances Itself’. *Operations Research* 44 (1): 21–34. https://doi.org/10.1287/opre.44.1.21.
Batt, Robert J., and Santiago Gallino. 2019. ‘Finding a Needle in a Haystack: The Effects of Searching and Learning on Pick-Worker Performance’. *Management Science* 65 (6): 2624–45. https://doi.org/10.1287/mnsc.2018.3059.
Battini, Daria, Christoph H. Glock, Eric H. Grosse, Alessandro Persona, and Fabio Sgarbossa. 2016. ‘Human Energy Expenditure in Order Picking Storage Assignment: A Bi-Objective Method’. *Computers & Industrial Engineering* 94 (April): 147–57. https://doi.org/10.1016/j.cie.2016.01.020.
Baumann, Hannes, Thad Starner, Hendrik Iben, Anna Lewandowski, and Patrick Zschaler. 2011. ‘Evaluation of Graphical User-Interfaces for Order Picking Using Head-Mounted Displays’. In *Proceedings of the 13th International Conference on Multimodal Interfaces*, 377–84. ICMI ’11. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/2070481.2070553.
Bendoly, Elliot, Rachel Croson, Paulo Goncalves, and Kenneth Schultz. 2010. ‘Bodies of Knowledge for Research in Behavioral Operations’. *Production and Operations Management* 19 (4): 434–52. https://doi.org/10.1111/j.1937-5956.2009.01108.x.
Bendoly, Elliot, and Mike Prietula. 2008. ‘In “the Zone”: The Role of Evolving Skill and Transitional Workload on Motivation and Realized Performance in Operational Tasks’. *International Journal of Operations & Production Management* 28 (12): 1130–52. https://doi.org/10.1108/01443570810919341.
Bongers, Paulien M, Cornells R de Winter, Michiel AJ Kompier, and Vincent H Hildebrandt. 1993. ‘Psychosocial Factors at Work and Musculoskeletal Disease’. *Scandinavian Journal of Work, Environment & Health* 19 (5): 297–312.
Boudreau, John, Wallace Hopp, John O. McClain, and L. Joseph Thomas. 2003. ‘On the Interface Between Operations and Human Resources Management’. *Manufacturing & Service Operations Management* 5 (3): 179–202. https://doi.org/10.1287/msom.5.3.179.16032.
Bouwens, Jan, and Peter Kroos. 2011. ‘Target Ratcheting and Effort Reduction’. *Journal of Accounting and Economics* 51 (1): 171–85. https://doi.org/10.1016/j.jacceco.2010.07.002.
Boysen, Nils, René De Koster, and Felix Weidinger. 2019. ‘Warehousing in the E-Commerce Era: A Survey’. *European Journal of Operational Research* 277 (2): 396–411. https://doi.org/10.1016/j.ejor.2018.08.023.
Braam, Ingrid Th. J., Max van Dormolen, and Monique H. W. Frings-Dresen. 1996. ‘The Work Load of Warehouse Workers in Three Different Working Systems’. *International Journal of Industrial Ergonomics* 17 (6): 469–80. https://doi.org/10.1016/0169-8141(95)00008-9.
Bräuer, Paula, and Athanasios Mazarakis. 2019. *Badges or a Leaderboard? How to Gamify an Augmented Reality Warehouse Setting*.
Breque, Maija, Lars De Nul, and Athanasios Petridis. 2021. *Industry 5.0: Towards a Sustainable, Human Centric and Resilient European Industry*. LU: Publications Office of the European Union. https://data.europa.eu/doi/10.2777/308407.
Brey, Philip. 1999. ‘Worker Autonomy and the Drama of Digital Networks in Organizations’. *Journal of Business Ethics* 22 (1): 15–25. https://doi.org/10.1023/A:1006199816737.
Bridger, Robert. 2017. *Introduction to Human Factors and Ergonomics*. CRC Press.

Brynzér, H., and M. I. Johansson. 1996. ‘Storage Location Assignment: Using the Product Structure to Reduce Order Picking Times’. *International Journal of Production Economics*, Proceedings of the 8th International Working Seminar on Production Economics, 46–47 (December): 595–603. https://doi.org/10.1016/0925-5273(94)00091-3.

Calzavara, Martina, Christoph H. Glock, Eric H. Grosse, Alessandro Persona, and Fabio Sgarbossa. 2017. ‘Analysis of Economic and Ergonomic Performance Measures of Different Rack Layouts in an Order Picking Warehouse’. *Computers & Industrial Engineering* 111 (September): 527–36. https://doi.org/10.1016/j.cie.2016.07.001.

Calzavara, Martina, Christoph H. Glock, Eric H. Grosse, and Fabio Sgarbossa. 2019. ‘An Integrated Storage Assignment Method for Manual Order Picking Warehouses Considering Cost, Workload and Posture’. *International Journal of Production Research* 57 (8): 2392–2408. https://doi.org/10.1080/00207543.2018.1518609.

Calzavara, Martina, Alessandro Persona, Fabio Sgarbossa, and Valentina Visentin. 2019. ‘A Model for Rest Allowance Estimation to Improve Tasks Assignment to Operators’. *International Journal of Production Research* 57 (3): 948–62. https://doi.org/10.1080/00207543.2018.1497816.

Cao, Alex, Keshav K. Chintamani, Abhilash K. Pandya, and R. Darin Ellis. 2009. ‘NASA TLX: Software for Assessing Subjective Mental Workload’. *Behavior Research Methods* 41 (1): 113–17. https://doi.org/10.3758/BRM.41.1.113.

Carter, Craig R. 2008. ‘Knowledge Production and Knowledge Transfer: Closing the Research-Practice Gap’. *Journal of Supply Chain Management*, 2008. Gale Academic OneFile.

Chandra, Charu, and Jānis Grabis, eds. 2007. ‘Mathematical Programming Approaches’. In *Supply Chain Configuration: Concepts, Solutions, and Applications*, 183–206. Boston, MA: Springer US. https://doi.org/10.1007/978-0-387-68155-9_8.

Chen, Yuhao, Chizhao Yang, Bowen Song, Nicholas Gonzalez, Yu Gu, and Boyi Hu. 2020. ‘Effects of Autonomous Mobile Robots on Human Mental Workload and System Productivity in Smart Warehouses: A Preliminary Study’. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 64 (1): 1691–95. https://doi.org/10.1177/1071181320641410.

Choi, Tsan-Ming, T. C. E. Cheng, and Xiande Zhao. 2016. ‘Multi-Methodological Research in Operations Management’. *Production and Operations Management* 25 (3): 379–89. https://doi.org/10.1111/poms.12534.

Cohrssen, Barbara. 2021. *Patty’s Industrial Hygiene, Hazard Recognition*. John Wiley & Sons.

Cragg, Tony, and Dominic Loske. 2019. ‘Perceived Work Autonomy in Order Picking Systems: An Empirical Analysis’. *IFAC-PapersOnLine*, 9th IFAC Conference on Manufacturing Modelling, Management and Control MIM 2019, 52 (13): 1872–77. https://doi.org/10.1016/j.ifacol.2019.11.475.

Davarzani, Hoda, and Andreas Norrman. 2015. ‘Toward a Relevant Agenda for Warehousing Research: Literature Review and Practitioners’ Input’. *Logistics Research* 8 (1): 1. https://doi.org/10.1007/s12159-014-0120-1.

De Koster, René, Tho Le-Duc, and Kees Jan Roodbergen. 2007. ‘Design and Control of Warehouse Order Picking: A Literature Review’. *European Journal of Operational Research* 182 (2): 481–501. https://doi.org/10.1016/j.ejor.2006.07.009.

De Koster, René, Daan Stam, and Bert M. Balk. 2011. ‘Accidents Happen: The Influence of Safety-Specific Transformational Leadership, Safety Consciousness, and Hazard Reducing Systems on Warehouse Accidents’. *Journal of Operations Management*, Special Issue: Product Safety and Security on the Global Supply Chain, 29 (7): 753–65. https://doi.org/10.1016/j.jom.2011.06.005.

De Vries, Jelle, René De Koster, and Daan Stam. 2016a. ‘Aligning Order Picking Methods, Incentive Systems, and Regulatory Focus to Increase Performance’. *Production and Operations Management* 25 (8): 1363–76. https://doi.org/10.1111/poms.12547.

———. 2016b. ‘Safety Does Not Happen by Accident: Antecedents To A Safer Warehouse’. *Production and Operations Management* 25 (8): 1377–90. https://doi.org/10.1111/poms.12546.

———. 2016c. ‘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): 2260–74. https://doi.org/10.1080/00207543.2015.1064184.

Deb, Kalyanmoy, and Kalyanmoy Deb. 2014. ‘Multi-Objective Optimization’. In *Search Methodologies: Introductory Tutorials in Optimization and Decision Support Techniques,* edited by Edmund K. Burke and Graham Kendall, 403–49. Boston, MA: Springer US. https://doi.org/10.1007/978-1-4614-6940-7_15.

Denyer, David, and David Tranfield. 2009. ‘Producing a Systematic Review’. In *The Sage Handbook of Organizational Research Methods*, 671–89. Thousand Oaks, CA: Sage Publications Ltd.

Deterding, Sebastian, Dan Dixon, Rilla Khaled, and Lennart Nacke. 2011. ‘From Game Design Elements to Gamefulness: Defining “Gamification”’. In *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 9–15. MindTrek ’11. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/2181037.2181040.

Devi, Takhellambam R., Shanmugam Munuswamy, and Mohamed Jasim. 2019. ‘Factors Influencing Job Burnout Level: An Empirical Investigation on Warehouse Employees’. *International Public Health Journal* 11 (2): 211–19.

Diefenbach, Heiko, and Christoph H. Glock. 2019. ‘Ergonomic and Economic Optimization of Layout and Item Assignment of a U-Shaped Order Picking Zone’. *Computers & Industrial Engineering* 138 (December): 106094. https://doi.org/10.1016/j.cie.2019.106094.

Digman, J M. 1990. ‘Personality Structure: Emergence of the Five-Factor Model’. *Annual Review of Psychology* 41 (1): 417–40. https://doi.org/10.1146/annurev.ps.41.020190.002221.

Doerr, Kenneth H., and Antonio Arreola-Risa. 2000. ‘A Worker-Based Approach for Modeling Variability in Task Completion Times’. *IIE Transactions* 32 (7): 625–36. https://doi.org/10.1023/A:1007659032655.

Donohue, Karen, and Kenneth Schultz. 2019. ‘The Future Is Bright: Recent Trends and Emerging Topics in Behavioral Operations’. In *The Handbook of Behavioral Operations*, 619–51. Wiley Series in Operations Research and Management Science. Wiley Blackwell.

Dudley, Norman Alfred. 1968. *Work Measurement: Research Studies*. Macmillan International Higher Education.

Elbert, Ralf M., Torsten Franzke, Christoph H. Glock, and Eric H. Grosse. 2017. ‘The Effects of Human Behavior on the Efficiency of Routing Policies in Order

Picking: The Case of Route Deviations'. *Computers & Industrial Engineering* 111 (September): 537–51. https://doi.org/10.1016/j.cie.2016.11.033.
Elbert, Ralf, and Philippe Jan Müller. 2017. 'Manoeuvres in Manual Person-to-Parts Order Picking Systems: Is There a Significant Influence on Travel Time and Metabolic Rates?' *IFAC-PapersOnLine*, 20th IFAC World Congress, 50 (1): 6894–99. https://doi.org/10.1016/j.ifacol.2017.08.1213.
Feinstein, Andrew Hale, and Hugh M. Cannon. 2001. 'Fidelity, Verifiability, and Validity of Simulation: Constructs for Evaluation'. *Developments in Business Simulation and Experiential Learning: Proceedings of the Annual ABSEL Conference* 28. https://absel-ojs-ttu.tdl.org/absel/index.php/absel/article/view/801.
Feng, Xiaochun, and Xiangpei Hu. 2021. 'A Heuristic Solution Approach to Order Batching and Sequencing for Manual Picking and Packing Lines Considering Fatiguing Effect'. Edited by Shah Nazir. *Scientific Programming* 2021 (April): 1–17. https://doi.org/10.1155/2021/8863391.
Ferrari, Alessio, Paola Spoletini, and Stefania Gnesi. 2016. 'Ambiguity and Tacit Knowledge in Requirements Elicitation Interviews'. *Requirements Engineering* 21 (3): 333–55. https://doi.org/10.1007/s00766-016-0249-3.
Fibrianto, Henokh Yernias, and Soondo Hong. 2019. 'Dynamic Order Batching in Bucket Brigade Order Picking Systems with Consecutive Batch Windows and Non-Identical Pickers'. *International Journal of Production Research* 57 (20): 6552–68. https://doi.org/10.1080/00207543.2019.1567948.
Frank, Alejandro Germán, Lucas Santos Dalenogare, and Néstor Fabián Ayala. 2019. 'Industry 4.0 Technologies: Implementation Patterns in Manufacturing Companies'. *International Journal of Production Economics* 210 (April): 15–26. https://doi.org/10.1016/j.ijpe.2019.01.004.
Füchtenhans, Marc, Eric H. Grosse, and Christoph H. Glock. 2021. 'Smart Lighting Systems: State-of-the-Art and Potential Applications in Warehouse Order Picking'. *International Journal of Production Research* 59 (12): 3817–39. https://doi.org/10.1080/00207543.2021.1897177.
Funk, Markus, Alireza Sahami Shirazi, Sven Mayer, Lars Lischke, and Albrecht Schmidt. 2015. 'Pick from Here! An Interactive Mobile Cart Using in-Situ Projection for Order Picking'. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 601–9. UbiComp '15. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/2750858.2804268.
Gajšek, Brigita, Goran Đukić, Marcin Butlewski, Tihomir Opetuk, Hrvoje Cajner, and Sonja Mlaker Kač. 2020. 'The Impact of the Applied Technology on Health and Productivity in Manual "Picker-to-Part" Systems'. *Work* 65 (3): 525–36. https://doi.org/10.3233/WOR-203107.
Gajšek, Brigita, Goran Đukić, Tihomir Opetuk, and Hrvoje Cajner. 2017. 'Human in Manual Order Picking Systems'. In *Conference Proceedings-Management of Technology – Step to Sustainable Production*. Dubrovnik. https://www.bib.irb.hr/877213/download/877213.Human_in_Manual_Order_Picking_Systems.pdf.
Gajšek, Brigita, Simona Šinko, Tomaž Kramberger, Marcin Butlewski, Eren Özceylan, and Goran Đukić. 2021. 'Towards Productive and Ergonomic Order Picking: Multi-Objective Modeling Approach'. *Applied Sciences* 11 (9): 4179. https://doi.org/10.3390/app11094179.

García-Buades, M. Esther, José M. Peiró, María Isabel Montañez-Juan, Malgorzata W. Kozusznik, and Silvia Ortiz-Bonnín. 2020. 'Happy-Productive Teams and Work Units: A Systematic Review of the "Happy-Productive Worker Thesis"'. *International Journal of Environmental Research and Public Health* 17 (1): 69. https://doi.org/10.3390/ijerph17010069.
Garg, Arun, Don B. Chaffin, and Gary D. Herrin. 1978. 'Prediction of Metabolic Rates for Manual Materials Handling Jobs'. *American Industrial Hygiene Association Journal* 39 (8): 661–74. https://doi.org/10.1080/0002889778507831.
Garg, Arun, and U. Saxena. 1985. 'Physiological Stresses in Warehouse Operations with Special Reference to Lifting Technique and Gender: A Case Study'. *American Industrial Hygiene Association Journal* 46 (2): 53–59. https://doi.org/10.1080/15298668591394383.
Glock, Christoph H., Eric H. Grosse, Hamid Abedinnia, and Simon Emde. 2019. 'An Integrated Model to Improve Ergonomic and Economic Performance in Order Picking by Rotating Pallets'. *European Journal of Operational Research* 273 (2): 516–34. https://doi.org/10.1016/j.ejor.2018.08.015.
Glock, Christoph H., Eric H. Grosse, Ralf M. Elbert, and Torsten Franzke. 2017. 'Maverick Picking: The Impact of Modifications in Work Schedules on Manual Order Picking Processes'. *International Journal of Production Research* 55 (21): 6344–60. https://doi.org/10.1080/00207543.2016.1252862.
Glock, Christoph H., Eric H. Grosse, T. Kim, W. P. Neumann, and A. Sobhani. 2019. 'An Integrated Cost and Worker Fatigue Evaluation Model of a Packaging Process'. *International Journal of Production Economics* 207 (January): 107–24. https://doi.org/10.1016/j.ijpe.2018.09.022.
Glock, Christoph H., Eric H. Grosse, W. Patrick Neumann, and Andrew Feldman. 2021. 'Assistive Devices for Manual Materials Handling in Warehouses: A Systematic Literature Review'. *International Journal of Production Research* 59 (11): 3446–69. https://doi.org/10.1080/00207543.2020.1853845.
Glover, Fred, and Kenneth Sörensen. 2015. 'Metaheuristics'. *Scholarpedia* 10 (4): 6532. https://doi.org/10.4249/scholarpedia.6532.
Granotto, Giovanni Fontana, Fabio Sgarbossa, Christoph H. Glock, and Eric H. Grosse. 2019. 'The Effect of Worker Fatigue on the Performance of a Bucket Brigade Order Picking System'. *IFAC-PapersOnLine*, 9th IFAC Conference on Manufacturing Modelling, Management and Control MIM 2019, 52 (13): 2195–2200. https://doi.org/10.1016/j.ifacol.2019.11.531.
Gräßler, Iris, Daniel Roesmann, and Jens Pottebaum. 2021. 'Model Based Integration of Human Characteristics in Production Systems: A Literature Survey'. *Procedia CIRP*, 14th CIRP Conference on Intelligent Computation in Manufacturing Engineering, 15-17 July 2020, 99 (January): 57–62. https://doi.org/10.1016/j.procir.2021.03.010.
Grosse, Eric H., Martina Calzavara, Christoph H. Glock, and Fabio Sgarbossa. 2017. 'Incorporating Human Factors into Decision Support Models for Production and Logistics: Current State of Research'. *IFAC-PapersOnLine*, 20th IFAC World Congress, 50 (1): 6900–6905. https://doi.org/10.1016/j.ifacol.2017.08.1214.
Grosse, Eric H., Shane M. Dixon, W. Patrick Neumann, and Christoph H. Glock. 2016. 'Using Qualitative Interviewing to Examine Human Factors in Warehouse Order Picking: Technical Note'. *International Journal of Logistics Systems and Management* 23 (4): 499–518. https://doi.org/10.1504/IJLSM.2016.075211.
Grosse, Eric H., and Christoph H. Glock. 2013. 'An Experimental Investigation of Learning Effects in Order Picking Systems'. *Journal of Manufacturing*

Technology Management 24 (6): 850–72. https://doi.org/10.1108/JMTM-03-2012-0036.
———. 2015. ‘The Effect of Worker Learning on Manual Order Picking Processes’. *International Journal of Production Economics*, Decision models for the design, optimization and management of warehousing and material handling systems, 170 (December): 882–90. https://doi.org/10.1016/j.ijpe.2014.12.018.
Grosse, Eric H., Christoph H. Glock, and Mohamad Y. Jaber. 2013. ‘The Effect of Worker Learning and Forgetting on Storage Reassignment Decisions in Order Picking Systems’. *Computers & Industrial Engineering* 66 (4): 653–62. https://doi.org/10.1016/j.cie.2013.09.013.
Grosse, Eric H., Christoph H. Glock, Mohamad Y. Jaber, and W. Patrick Neumann. 2015. ‘Incorporating Human Factors in Order Picking Planning Models: Framework and Research Opportunities’. *International Journal of Production Research* 53 (3): 695–717. https://doi.org/10.1080/00207543.2014.919424.
Grosse, Eric H., Christoph H. Glock, and W. Patrick Neumann. 2015. ‘Human Factors in Order Picking System Design: A Content Analysis’. *IFAC-PapersOnLine*, 15th IFAC Symposium onInformation Control Problems inManufacturing, 48 (3): 320–25. https://doi.org/10.1016/j.ifacol.2015.06.101.
———. 2017. ‘Human Factors in Order Picking: A Content Analysis of the Literature’. *International Journal of Production Research* 55 (5): 1260–76. https://doi.org/10.1080/00207543.2016.1186296.
Guest, Greg, Arwen Bunce, and Laura Johnson. 2006. ‘How Many Interviews Are Enough?: An Experiment with Data Saturation and Variability’. *Field Methods* 18 (1): 59–82. https://doi.org/10.1177/1525822X05279903.
Guo, Anhong, Shashank Raghu, Xuwen Xie, Saad Ismail, Xiaohui Luo, Joseph Simoneau, Scott Gilliland, Hannes Baumann, Caleb Southern, and Thad Starner. 2014. ‘A Comparison of Order Picking Assisted by Head-up Display (HUD), Cart-Mounted Display (CMD), Light, and Paper Pick List’. In *Proceedings of the 2014 ACM International Symposium on Wearable Computers*, 71–78. ISWC ’14. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/2634317.2634321.
Hanson, Robin, Lars Medbo, Cecilia Berlin, and Jonas Hansson. 2018. ‘Manual Picking from Flat and Tilted Pallet Containers’. *International Journal of Industrial Ergonomics* 64 (March): 199–212. https://doi.org/10.1016/j.ergon.2017.07.001.
Hanson, Robin, Lars Medbo, Patrik Jukic, and Majeed Assaf. 2016. ‘Manual Picking from Large Containers - Time Efficiency and Physical Workload’. *IFAC-PapersOnLine*, 8th IFAC Conference on Manufacturing Modelling, Management and Control MIM 2016, 49 (12): 1703–8. https://doi.org/10.1016/j.ifacol.2016.07.827.
Hart, Sandra G., and Lowell E. Staveland. 1988. ‘Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research’. In *Advances in Psychology*, edited by Peter A. Hancock and Najmedin Meshkati, 52:139–83. Human Mental Workload. North-Holland. https://doi.org/10.1016/S0166-4115(08)62386-9.
Hennink, Monique M., Bonnie N. Kaiser, and Vincent C. Marconi. 2017. ‘Code Saturation Versus Meaning Saturation: How Many Interviews Are Enough?’ *Qualitative Health Research* 27 (4): 591–608. https://doi.org/10.1177/1049732316665344.
Higgins, E. Tory. 1998. ‘Promotion and Prevention: Regulatory Focus as A Motivational Principle’. In *Advances in Experimental Social Psychology*, edited

by Mark P. Zanna, 30:1–46. Academic Press. https://doi.org/10.1016/S0065-2601(08)60381-0.

Hilmola, Olli-Pekka, and Andres Tolli. 2016. 'Warehouse Layout Implications on Picking Distance: Case of Human Factor'. *World Review of Intermodal Transportation Research* 6 (1): 43–58. https://doi.org/10.1504/WRITR.2016.078138.

Hong, Soondo. 2018. 'The Effects of Picker-Oriented Operational Factors on Hand-off Delay in a Bucket Brigade Order Picking System'. *OR Spectrum* 40 (3): 781–808. https://doi.org/10.1007/s00291-018-0523-5.

Hong, Soondo, Andrew L. Johnson, and Brett A. Peters. 2015. 'Quantifying Picker Blocking in a Bucket Brigade Order Picking System'. *International Journal of Production Economics*, Decision models for the design, optimization and management of warehousing and material handling systems, 170 (December): 862–73. https://doi.org/10.1016/j.ijpe.2015.04.012.

Horne, J. A., and O. Östberg. 1976. 'A Self-Assessment Questionnaire to Determine Morningness-Eveningness in Human Circadian Rhythms'. *International Journal of Chronobiology* 4: 97–110.

Ilmarinen, Juhani E. 2001. 'AGING WORKERS'. *Occupational and Environmental Medicine* 58 (8): 546–546. https://doi.org/10.1136/oem.58.8.546.

Jaber, M. Y., Z. S. Givi, and W. P. Neumann. 2013. 'Incorporating Human Fatigue and Recovery into the Learning–Forgetting Process'. *Applied Mathematical Modelling* 37 (12): 7287–99. https://doi.org/10.1016/j.apm.2013.02.028.

Juran, David C., and Lee W. Schruben. 2004. 'Using Worker Personality and Demographic Information to Improve System Performance Prediction'. *Journal of Operations Management*, Simulation Studies in Operations Management, 22 (4): 355–67. https://doi.org/10.1016/j.jom.2004.05.003.

Kadir, Bzhwen A., Ole Broberg, and Carolina Souza da Conceição. 2019. 'Current Research and Future Perspectives on Human Factors and Ergonomics in Industry 4.0'. *Computers & Industrial Engineering* 137 (November): 106004. https://doi.org/10.1016/j.cie.2019.106004.

Kelton, W. David. 2002. *Simulation with Arena*. McGraw-Hill. http://thuvienso.bvu.edu.vn/handle/TVDHBRVT/14447.

Kim, Sunwook, Maury A. Nussbaum, and Joseph L. Gabbard. 2019. 'Influences of Augmented Reality Head-Worn Display Type and User Interface Design on Performance and Usability in Simulated Warehouse Order Picking'. *Applied Ergonomics* 74 (January): 186–93. https://doi.org/10.1016/j.apergo.2018.08.026.

Klumpp, Matthias, and Dominic Loske. 2021. 'Order Picking and E-Commerce: Introducing Non-Parametric Efficiency Measurement for Sustainable Retail Logistics'. *Journal of Theoretical and Applied Electronic Commerce Research* 16 (4): 846–58. https://doi.org/10.3390/jtaer16040048.

Knox, Sarah, and Alan W. Burkard. 2009. 'Qualitative Research Interviews'. *Psychotherapy Research* 19 (4–5): 566–75. https://doi.org/10.1080/10503300802702105.

Kocakulah, Mehmet C., Ann Galligan Kelley, Krystal M. Mitchell, and Margaret P. Ruggieri. 2016. 'Absenteeism Problems And Costs: Causes, Effects And Cures'. *International Business & Economics Research Journal (IBER)* 15 (3): 89–96. https://doi.org/10.19030/iber.v15i3.9673.

Koo, Pyung-Hoi. 2009. 'The Use of Bucket Brigades in Zone Order Picking Systems'. *OR Spectrum* 31 (4): 759–74. https://doi.org/10.1007/s00291-008-0131-x.

Kretschmer, Veronika, Thorsten Plewan, Gerhard Rinkenauer, and Benedikt Maettig. 2018. 'Smart Palletisation: Cognitive Ergonomics in Augmented Reality Based Palletising'. In *Intelligent Human Systems Integration*, edited by Waldemar Karwowski and Tareq Ahram, 355–60. Advances in Intelligent Systems and Computing. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-73888-8_55.

Kreutzfeldt, Magali, Johanna Renker, and Gerhard Rinkenauer. 2019a. 'The Attentional Perspective on Smart Devices: Empirical Evidence for Device-Specific Cognitive Ergonomics'. In *Advances in Ergonomics in Design*, edited by Francisco Rebelo and Marcelo M. Soares, 3–13. Advances in Intelligent Systems and Computing. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-94706-8_1.

———. 2019b. 'The Influence of Gait on Cognitive Functions: Promising Factor for Adapting Systems to the Worker's Need in a Picking Context'. In *Adaptive Instructional Systems*, edited by Robert A. Sottilare and Jessica Schwarz, 420–31. Lecture Notes in Computer Science. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-22341-0_33.

Lager, Hendrik, Alfredo Virgillito, and Tom-Philipp Buchberger. 2021. 'Digitalization of Logistics Work: Ergonomic Improvements Versus Work Intensification'. In *Digital Supply Chains and the Human Factor*, edited by Matthias Klumpp and Caroline Ruiner, 33–53. Lecture Notes in Logistics. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-58430-6_3.

Lambrechts, Wim, Jessica S. Klaver, Lennart Koudijzer, and Janjaap Semeijn. 2021. 'Human Factors Influencing the Implementation of Cobots in High Volume Distribution Centres'. *Logistics* 5 (2): 32. https://doi.org/10.3390/logistics5020032.

Larco, José Martinelli. 2010. *Incorporating Worker-Specific Factors in Operations Management Models*. https://repub.eur.nl/pub/21527/.

Larco, José Martinelli, René de Koster, Kees Jan Roodbergen, and Jan Dul. 2016. 'Managing Warehouse Efficiency and Worker Discomfort through Enhanced Storage Assignment Decisions'. *International Journal of Production Research* 55 (21): 6407–22. https://doi.org/10.1080/00207543.2016.1165880.

Lecy, Jesse D., and Kate E. Beatty. 2012. 'Representative Literature Reviews Using Constrained Snowball Sampling and Citation Network Analysis'. SSRN Scholarly Paper ID 1992601. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.1992601.

Lee, Eun Young, Min Kyu Kim, and Yoon Seok Chang. 2016. 'Development of an Advanced Picking Station Considering Human Factors'. *Human Factors and Ergonomics in Manufacturing & Service Industries* 26 (6): 700–712. https://doi.org/10.1002/hfm.20669.

Lee, Joo Ae, Yoon Seok Chang, and Waldemar Karwowski. 2020. 'Assessment of Working Postures and Physical Loading in Advanced Order Picking Tasks: A Case Study of Human Interaction with Automated Warehouse Goods-to-Picker Systems'. *Work* 67 (4): 855–66. https://doi.org/10.3233/WOR-203337.

Levine, James A. 2005. 'Measurement of Energy Expenditure'. *Public Health Nutrition* 8 (7a): 1123–32. https://doi.org/10.1079/PHN2005800.

Lewis-Beck, Michael, Alan Bryman, and Tim Liao. 2004. *The SAGE Encyclopedia of Social Science Research Methods*. 3 vols. Thousand Oaks, California. https://doi.org/10.4135/9781412950589.

Li, Kai, Shuling Xu, and Hong Fu. 2020. ‘Work-Break Scheduling with Real-Time Fatigue Effect and Recovery’. *International Journal of Production Research* 58 (3): 689–702. https://doi.org/10.1080/00207543.2019.1598600.

Lin, Georgianna, Malcolm Haynes, Sarthak Srinivas, Pramod Kotipalli, and Thad Starner. 2021. ‘Towards Finding the Optimum Position in the Visual Field for a Head Worn Display Used for Task Guidance with Non-Registered Graphics’. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5 (1): 1–26. https://doi.org/10.1145/3448091.

Lin, Georgianna, Tanmoy Panigrahi, Jon Womack, Devansh Jatin Ponda, Pramod Kotipalli, and Thad Starner. 2021. ‘Comparing Order Picking Guidance with Microsoft Hololens, Magic Leap, Google Glass XE and Paper’. In *Proceedings of the 22nd International Workshop on Mobile Computing Systems and Applications*, 133–39. Virtual United Kingdom: ACM. https://doi.org/10.1145/3446382.3448729.

M14 Intelligence. 2021. ‘Warehouse Automation - Robots, Technologies, and Solutions Market, 2021 - 2030’. https://m14intelligence.com/public/index.php/product/15.

Maidstone, Robert. 2012. ‘Discrete Event Simulation, System Dynamics and Agent Based Simulation: Discussion and Comparison’, March, 1–6.

Manis, Kerry T., and Danny Choi. 2019. ‘The Virtual Reality Hardware Acceptance Model (VR-HAM): Extending and Individuating the Technology Acceptance Model (TAM) for Virtual Reality Hardware’. *Journal of Business Research* 100 (July): 503–13. https://doi.org/10.1016/j.jbusres.2018.10.021.

Marchet, Gino, Marco Melacini, and Sara Perotti. 2015. ‘Investigating Order Picking System Adoption: A Case-Study-Based Approach’. *International Journal of Logistics Research and Applications* 18 (1): 82–98. https://doi.org/10.1080/13675567.2014.945400.

Matusiak, Marek, René De Koster, and Jari Saarinen. 2017. ‘Utilizing Individual Picker Skills to Improve Order Batching in a Warehouse’. *European Journal of Operational Research* 263 (3): 888–99. https://doi.org/10.1016/j.ejor.2017.05.002.

McClave, James T., and Terry Sincich. 2007. *Statistiek, 9/e*. Pearson Education.

Meller, Russell D., and John F. Klote. 2004. ‘A Throughput Model for Carousel/VLM Pods’. *IIE Transactions* 36 (8): 725–41. https://doi.org/10.1080/07408170490458472.

Michel, Roberto. 2016. ‘2016 Warehouse/DC Operations Survey: Ready to Confront Complexity’. *Supply Chain Management Review*, no. November 2016. https://www.scmr.com/plus/SCMR1611_SUP_Warehouse_DCBenchmarkStudy.pdf.

Min, Hokey. 2004. ‘An Examination of Warehouse Employee Recruitment and Retention Practices in the USA’. *International Journal of Logistics Research and Applications* 7 (4): 345–59. https://doi.org/10.1080/13675560412331282948.

———. 2007. ‘Examining Sources of Warehouse Employee Turnover’. *International Journal of Physical Distribution & Logistics Management* 37 (5): 375–88. https://doi.org/10.1108/09600030710758437.

Mocan, Anca, and Anca Draghici. 2018. ‘Reducing Ergonomic Strain in Warehouse Logistics Operations by Using Wearable Computers’. *Procedia - Social and Behavioral Sciences*, Challenges and Innovation in Management and Entrepreneurship, 238 (January): 1–8. https://doi.org/10.1016/j.sbspro.2018.03.001.

Mohler, Betty J., William B. Thompson, Sarah H. Creem-Regehr, Herbert L. Pick, and William H. Warren. 2007. 'Visual Flow Influences Gait Transition Speed and Preferred Walking Speed'. *Experimental Brain Research* 181 (2): 221–28. https://doi.org/10.1007/s00221-007-0917-0.

Murauer, Nela, Florian Müller, Sebastian Günther, Dominik Schön, Nerina Pflanz, and Markus Funk. 2018. 'An Analysis of Language Impact on Augmented Reality Order Picking Training'. In *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference*, 351–57. PETRA '18. New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3197768.3201570.

Neumann, W Patrick, and Jan Dul. 2010. 'Human Factors: Spanning the Gap between OM and HRM'. *International Journal of Operations & Production Management* 30 (9): 923–50. https://doi.org/10.1108/01443571011075056.

Neumann, W. Patrick, Sven Winkelhaus, Eric H. Grosse, and Christoph H. Glock. 2021. 'Industry 4.0 and the Human Factor – A Systems Framework and Analysis Methodology for Successful Development'. *International Journal of Production Economics* 233 (March): 107992. https://doi.org/10.1016/j.ijpe.2020.107992.

Neumann, W.P., and J. Village. 2012. 'Ergonomics Action Research II: A Framework for Integrating HF into Work System Design'. *Ergonomics* 55 (10): 1140–56. https://doi.org/10.1080/00140139.2012.706714.

Norman, Bryan A., Wipawee Tharmmaphornphilas, Kim Lascola Needy, Bopaya Bidanda, and Rona Colosimo Warner. 2002. 'Worker Assignment in Cellular Manufacturing Considering Technical and Human Skills'. *International Journal of Production Research* 40 (6): 1479–92. https://doi.org/10.1080/00207540110118082.

Ogasawara, Ayumi, Aya Ishigaki, and Seiichi Yasui. 2019. 'Adaptive Storage Reassignment in Order Picking Systems to Picker Learning and Change of Demand'. *Procedia Manufacturing*, 25th International Conference on Production Research Manufacturing Innovation: Cyber Physical Manufacturing August 9-14, 2019 | Chicago, Illinois (USA), 39 (January): 1623–32. https://doi.org/10.1016/j.promfg.2020.01.279.

Ogbeyemi, Akinola, Wenjun Lin, Forrest Zhang, and Wenjun Zhang. 2021. 'Human Factors among Workers in a Small Manufacturing Enterprise: A Case Study'. *Enterprise Information Systems* 15 (6): 888–908. https://doi.org/10.1080/17517575.2020.1829076.

Ostermeier, Frederik Ferid. 2020. 'The Impact of Human Consideration, Schedule Types and Product Mix on Scheduling Objectives for Unpaced Mixed-Model Assembly Lines'. *International Journal of Production Research* 58 (14): 4386–4405. https://doi.org/10.1080/00207543.2019.1652780.

Otto, Alena, Nils Boysen, Armin Scholl, and Rico Walter. 2017. 'Ergonomic Workplace Design in the Fast Pick Area'. *OR Spectrum* 39 (4): 945–75. https://doi.org/10.1007/s00291-017-0479-x.

Otto, Alena, and Armin Scholl. 2011. 'Incorporating Ergonomic Risks into Assembly Line Balancing'. *European Journal of Operational Research* 212 (2): 277–86. https://doi.org/10.1016/j.ejor.2011.01.056.

Öztürkoğlu, Yücel Yılmaz, and Robert L. Bulfin. 2012. 'Scheduling Jobs to Consider Physiological Factors'. *Human Factors and Ergonomics in Manufacturing & Service Industries* 22 (2): 113–20. https://doi.org/10.1002/hfm.20257.

Pacheco, Meritxell Paneque, Michel Bierlaire, Bernard Gendron, and Shadi Sharif Azadeh. 2021. 'Integrating Advanced Discrete Choice Models in Mixed Integer

Linear Optimization'. *Transportation Research Part B: Methodological* 146 (April): 26–49. https://doi.org/10.1016/j.trb.2021.02.003.
Pasparakis, Alexandros, Jelle De Vries, and M. B. M. De Koster. 2021. 'In Control or under Control? Human-Robot Collaboration in Warehouse Order Picking'. SSRN Scholarly Paper ID 3816533. Rochester, NY: Social Science Research Network. https://doi.org/10.2139/ssrn.3816533.
Passalacqua, Mario, Pierre-Majorique Léger, Lennart E. Nacke, Marc Fredette, Élise Labonté-Lemoyne, Xinli Lin, Tony Caprioli, and Sylvain Sénécal. 2020. 'Playing in the Backstore: Interface Gamification Increases Warehousing Workforce Engagement'. *Industrial Management & Data Systems* 120 (7): 1309–30. https://doi.org/10.1108/IMDS-08-2019-0458.
Punnett, Laura, and David H. Wegman. 2004. 'Work-Related Musculoskeletal Disorders: The Epidemiologic Evidence and the Debate'. *Journal of Electromyography and Kinesiology*, State of the art research perspectives on muscoskeletal disorder causation and control, 14 (1): 13–23. https://doi.org/10.1016/j.jelekin.2003.09.015.
Raychaudhuri, Samik. 2008. 'Introduction to Monte Carlo Simulation'. In *2008 Winter Simulation Conference*, 91–100. https://doi.org/10.1109/WSC.2008.4736059.
Reif, R., W. A. Günthner, B. Schwerdtfeger, and G. Klinker. 2010. 'Evaluation of an Augmented Reality Supported Picking System Under Practical Conditions'. *Computer Graphics Forum* 29 (1): 2–12. https://doi.org/10.1111/j.1467-8659.2009.01538.x.
Reif, Rupert, and Dennis Walch. 2008. 'Augmented & Virtual Reality Applications in the Field of Logistics'. *The Visual Computer* 24 (11): 987–94. https://doi.org/10.1007/s00371-008-0271-7.
Reiman, Arto, Jari Kaivo-oja, Elina Parviainen, Esa-Pekka Takala, and Theresa Lauraeus. 2021. 'Human Factors and Ergonomics in Manufacturing in the Industry 4.0 Context – A Scoping Review'. *Technology in Society* 65 (May): 101572. https://doi.org/10.1016/j.techsoc.2021.101572.
Revelle, William, and David M. Condon. 2015. 'A Model for Personality at Three Levels'. *Journal of Research in Personality*, Integrative Theories of Personality, 56 (June): 70–81. https://doi.org/10.1016/j.jrp.2014.12.006.
Rijal, Arpan, Marco Bijvank, Asvin Goel, and René de Koster. 2021. 'Workforce Scheduling with Order-Picking Assignments in Distribution Facilities'. *Transportation Science* 55 (3): 725–46. https://doi.org/10.1287/trsc.2020.1029.
Robinson, Oliver C. 2014. 'Sampling in Interview-Based Qualitative Research: A Theoretical and Practical Guide'. *Qualitative Research in Psychology* 11 (1): 25–41. https://doi.org/10.1080/14780887.2013.801543.
Rozinat, A., and W. M. P. van der Aalst. 2008. 'Conformance Checking of Processes Based on Monitoring Real Behavior'. *Information Systems* 33 (1): 64–95. https://doi.org/10.1016/j.is.2007.07.001.
Ryan, Brendan, Rong Qu, Alex Schock, and Tony Parry. 2011. 'Integrating Human Factors and Operational Research in a Multidisciplinary Investigation of Road Maintenance'. *Ergonomics* 54 (5): 436–52. https://doi.org/10.1080/00140139.2011.562983.
Scheuermann, Constantin, Maximilian Strobel, Bernd Bruegge, and Stephan Verclas. 2016. 'Increasing the Support to Humans in Factory Environments Using a Smart Glove: An Evaluation'. In *2016 Intl IEEE Conferences on Ubiquitous Intelligence Computing, Advanced and Trusted Computing, Scalable Computing and Communications, Cloud and Big Data Computing, Internet of People, and*

Smart World Congress (UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld), 847–54. https://doi.org/10.1109/UIC-ATC-ScalCom-CBDCom-IoP-SmartWorld.2016.0134.

Schuir, Julian, and Frank Teuteberg. 2021. 'Understanding Augmented Reality Adoption Trade-Offs in Production Environments from the Perspective of Future Employees: A Choice-Based Conjoint Study'. *Information Systems and E-Business Management*, May. https://doi.org/10.1007/s10257-021-00529-0.

Schwerdtfeger, Björn, Rupert Reif, Willibald A. Günthner, and Gudrun Klinker. 2011. 'Pick-by-Vision: There Is Something to Pick at the End of the Augmented Tunnel'. *Virtual Reality* 15 (2): 213–23. https://doi.org/10.1007/s10055-011-0187-9.

Setayesh, Azin, Eric H. Grosse, Christoph H. Glock, and W. Patrick Neumann. 2021. 'Determining the Source of Human-System Errors in Manual Order Picking with Respect to Human Factors'. *International Journal of Production Research* 0 (0): 1–23. https://doi.org/10.1080/00207543.2021.1991022.

Sgarbossa, Fabio, Eric H. Grosse, W. Patrick Neumann, Daria Battini, and Christoph H. Glock. 2020. 'Human Factors in Production and Logistics Systems of the Future'. *Annual Reviews in Control* 49 (January): 295–305. https://doi.org/10.1016/j.arcontrol.2020.04.007.

Siebers, P O, C M Macal, J Garnett, D Buxton, and M Pidd. 2010. 'Discrete-Event Simulation Is Dead, Long Live Agent-Based Simulation!' *Journal of Simulation* 4 (3): 204–10. https://doi.org/10.1057/jos.2010.14.

Singhal, Vinod, Barbara B. Flynn, Peter T. Ward, Aleda V. Roth, and Vishal Gaur. 2008. 'Editorial: Empirical Elephants-Why Multiple Methods Are Essential to Quality Research in Operations and Supply Chain Management'. *Journal of Operations Management* 26 (3): 345–48. https://doi.org/10.1016/j.jom.2008.03.005.

Smith, Eboni, Reuben F. Burch V, Lesley Strawderman, Harish Chander, and Brian K. Smith. 2021. 'A Comfort Analysis of Using Smart Glasses during "Picking" and "Putting" Tasks'. *International Journal of Industrial Ergonomics* 83 (May): 103133. https://doi.org/10.1016/j.ergon.2021.103133.

Snyder, Hannah. 2019. 'Literature Review as a Research Methodology: An Overview and Guidelines'. *Journal of Business Research* 104 (November): 333–39. https://doi.org/10.1016/j.jbusres.2019.07.039.

Sony, Michael, and Subhash Naik. 2020. 'Industry 4.0 Integration with Socio-Technical Systems Theory: A Systematic Review and Proposed Theoretical Model'. *Technology in Society* 61 (May): 101248. https://doi.org/10.1016/j.techsoc.2020.101248.

Stanton, Neville Anthony, Alan Hedge, Karel Brookhuis, Eduardo Salas, and Hal W. Hendrick. 2004. *Handbook of Human Factors and Ergonomics Methods*. CRC Press.

Staudt, Francielly Hedler, Gülgün Alpan, Maria Di Mascolo, and Carlos M. Taboada Rodriguez. 2015. 'Warehouse Performance Measurement: A Literature Review'. *International Journal of Production Research* 53 (18): 5524–44. https://doi.org/10.1080/00207543.2015.1030466.

Steinebach, Tim, Jurij Wakula, and Asim Mehmedovic. 2021. 'The Influence of an Ergonomic Storage Location Assignment on Human Strain in Manual Order Picking'. In *Proceedings of the 21st Congress of the International Ergonomics Association (IEA 2021)*, edited by Nancy L. Black, W. Patrick Neumann, and

Ian Noy, 221:511–21. Lecture Notes in Networks and Systems. Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-74608-7_63.

Steinebach, Tim, Julia Wenzel, Jurij Wakula, and Ralf Elbert. 2021. *The Body Height as an Input Parameter for a Capability-Based Order Assignment in Manual Order Picking*.

Stockinger, Christopher, Tim Steinebach, Deborah Petrat, Robert Bruns, and Ilka Zöller. 2020. 'The Effect of Pick-by-Light-Systems on Situation Awareness in Order Picking Activities'. *Procedia Manufacturing*, Learning Factories across the value chain – from innovation to service – The 10th Conference on Learning Factories 2020, 45 (January): 96–101. https://doi.org/10.1016/j.promfg.2020.04.078.

Sun, Xiaoguang, Rémy Houssin, Jean Renaud, and Mickaël Gardoni. 2019. 'A Review of Methodologies for Integrating Human Factors and Ergonomics in Engineering Design'. *International Journal of Production Research* 57 (15–16): 4961–76. https://doi.org/10.1080/00207543.2018.1492161.

Sunstein, Cass R. 2018. *Nudging: A Very Short Guide*. *The Handbook of Privacy Studies*. Amsterdam University Press. https://www.degruyter.com/document/doi/10.1515/9789048540136-007/html.

Takala, Esa-Pekka, Irmeli Pehkonen, Mikael Forsman, Gert-Åke Hansson, Svend Erik Mathiassen, W Patrick Neumann, Gisela Sjøgaard, Kaj Bo Veiersted, Rolf H Westgaard, and Jørgen Winkel. 2010. 'Systematic Evaluation of Observational Methods Assessing Biomechanical Exposures at Work'. *Scandinavian Journal of Work, Environment & Health* 36 (1): 3–24.

Techawiboonwong, Atthawit, Pisal Yenradee, and Sanchoy K. Das. 2006. 'A Master Scheduling Model with Skilled and Unskilled Temporary Workers'. *International Journal of Production Economics* 103 (2): 798–809. https://doi.org/10.1016/j.ijpe.2005.11.009.

Tompkins, James A., John A. White, Yavuz A. Bozer, and J. M. A. Tanchoco. 2010. *Facilities Planning*. John Wiley & Sons.

Van Gils, Teun, Katrien Ramaekers, An Caris, and René B. M. De Koster. 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. https://doi.org/10.1016/j.ejor.2017.09.002.

Vanbrabant, Lien, Niels Martin, Katrien Ramaekers, and Kris Braekers. 2019. 'Quality of Input Data in Emergency Department Simulations: Framework and Assessment Techniques'. *Simulation Modelling Practice and Theory* 91 (February): 83–101. https://doi.org/10.1016/j.simpat.2018.12.002.

Vanheusden, Sarah, Teun van Gils, Katrien Ramaekers, Trijntje Cornelissens, and An Caris. 2022. 'Practical Factors in Order Picking Planning: State-of-the-Art Classification and Review'. *International Journal of Production Research* 0 (0): 1–25. https://doi.org/10.1080/00207543.2022.2053223.

Vidovič, Ela, and Brigita Gajšek. 2020. 'Analysing Picking Errors in Vision Picking Systems'. *Logistics & Sustainable Transport* 11 (February): 90–100. https://doi.org/10.2478/jlst-2020-0006.

Waters, T.R., V Putz-Anderson, and A Garg. 1994. *Applications Manual for the Revised NIOSH Lifting Equation.*

Weaver, Kimberly A., Hannes Baumann, Thad Starner, Hendrick Iben, and Michael Lawo. 2010. 'An Empirical Task Analysis of Warehouse Order Picking Using Head-Mounted Displays'. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1695–1704. CHI '10. New York, NY, USA:

Association for Computing Machinery. https://doi.org/10.1145/1753326.1753580.
Webster, Scott, Robert A. Ruben, and Kum-Khiong Yang. 2012. 'Impact of Storage Assignment Decisions on a Bucket Brigade Order Picking Line'. *Production and Operations Management* 21 (2): 276–90. https://doi.org/10.1111/j.1937-5956.2011.01267.x.
Wildemuth, Barbara M. 2016. *Applications of Social Research Methods to Questions in Information and Library Science, 2nd Edition*. ABC-CLIO.
Winkelhaus, S., F. Sgarbossa, M. Calzavara, and E. H. Grosse. 2018. 'The Effects of Human Fatigue on Learning in Order Picking: An Explorative Experimental Investigation'. *IFAC-PapersOnLine*, 16th IFAC Symposium on Information Control Problems in Manufacturing INCOM 2018, 51 (11): 832–37. https://doi.org/10.1016/j.ifacol.2018.08.442.
Winkelhaus, Sven, Eric H. Grosse, and Stefan Morana. 2021. 'Towards a Conceptualisation of Order Picking 4.0'. *Computers & Industrial Engineering* 159 (September): 107511. https://doi.org/10.1016/j.cie.2021.107511.
Wright, T. P. 1936. 'Factors Affecting the Cost of Airplanes'. *Journal of the Aeronautical Sciences* 3 (4): 122–28. https://doi.org/10.2514/8.155.
Yeow, Paul H. P., and David T. Goomas. 2014. 'Ergonomics Improvement in Order Selection in a Refrigerated Environment'. *Human Factors and Ergonomics in Manufacturing & Service Industries* 24 (3): 262–74. https://doi.org/10.1002/hfm.20374.
Zangaro, F, S Finco, Daria Battini, and I Zennaro. 2019. 'An Optimization Model for the Storage Assignment of the Reference under Ergonomics Constraints', Proceedings of the Summer School Francesco Turco, .
Zhang, Jun, Feng Liu, Jiafu Tang, and Yanhui Li. 2019. 'The Online Integrated Order Picking and Delivery Considering Pickers' Learning Effects for an O2O Community Supermarket'. *Transportation Research Part E: Logistics and Transportation Review* 123 (March): 180–99. https://doi.org/10.1016/j.tre.2019.01.013.
Zhang, Xiaoli, Jelle De Vries, René De Koster, and ChenGuang Liu. 2021. 'Fast and Faultless? Quantity and Quality Feedback in Order Picking'. *Production and Operations Management*, August.
Zhao, Xiaosong, Na Liu, Shumeng Zhao, Jinhui Wu, Kun Zhang, and Rui Zhang. 2019. 'Research on the Work-Rest Scheduling in the Manual Order Picking Systems to Consider Human Factors'. *Journal of Systems Science and Systems Engineering* 28 (3): 344–55. https://doi.org/10.1007/s11518-019-5407-y.