Adaptive Paradox: Balancing Empowerment and Equity in Technology-Mediated Spatial Production

Arina Rahma [1] and Timo Hartmann [1]
a.rahma@tu-berlin.de; timo.hartmann @tu-berlin.de
[1] Technische Universität Berlin

Abstract

Space is not a neutral container but socially produced (Lefebvre, 1991). Industry 5.0 (Commission, 2021) has reshaped this production by introducing technology as an active intermediary, transforming a human-environment dynamic into a human-system-environment interaction. As technology becomes organisational infrastructure, it stabilises around the designers' assumptions (Pinch & Bijker, 1984), and the space it produces reflects those assumptions. When a system's representational logics and underlying datasets encode assumptions of dominant user groups as default, the resulting spatial arrangement may function well for some while imposing disproportionate loads on others, not by intention, but through structural absence. This paper introduces the "adaptive paradox" to describe how technology, once infrastructuralised for universality, generates inequitable spatial burdens on users whose profiles were under-represented at the moment of design. We examine this in Voice-Directed Warehousing (VDW), a pick-by-voice order-picking technology deployed in a Portuguese warehouse which employs a linguistically and demographically diverse workforce. Drawing on spatial production (Lefebvre, 1991), Social Construction of Technology (SCOT) (Pinch & Bijker, 1984), and feminist technology studies (FTS) (D'Ignazio & Klein, 2020); (Massey, 2009; Wajcman, 2004), we analyse how human–VDW collaboration produces differential spatial practices across diverse operators. Methods include ethnographic fieldwork during two spans of five-day fieldwork in 2025 with 20 participants and analysed through thematic analysis (Braun & Clarke, 2006) referring to theoretical frameworks. Three findings illuminate the adaptive paradox across conceived design, the perceived operational outcomes, and workers' everyday lived responses. Together, they suggest human–VDW collaboration shift physical into cognitive and spatial loads that fall unevenly across groups, calling for spatial governance grounded in participatory design and spatial diversification.

Keywords: spatial production, equity, voice-recognition, technology-mediated space, industry 5.0

1. Introduction

When technology integrates to industries, it does not merely act as an efficiency driver, it becomes infrastructure. Recent scholars of socio-technical systems have debated whether infrastructure's significance lies primarily in organisational arrangements (Lee & Schmidt, 2018) or foregrounding materiality risks technological determinism (Corchia & Borghini, 2025; Degens et al., 2022). In ethnography framework, infrastructure shapes and is shaped by community practices and social arrangements (Star, 1999), which in turn spatially produce its materiality. The materiality defines the channels through which knowledge and power circulate, shaping who can move where and under what conditions (Graham, 2010).

Therefore, space, as Lefebvre (1991) argued, is not a passive container but a social product, constructed through the interplay of institutional design and everyday human practices. In industries, how workspace is designed and how it influences workers' workaround are acts of spatial production. Industry 5.0 (Commission, 2021) has fundamentally altered this production, introducing technology as an active intermediary and shifting the relationship from human-environment to human-technology-environment within workforce. The shift, while promises human-centric collaboration, often prioritises aggregate efficiency metrics (Grosse et al., 2015) over differential lived realities, showing a critical gap between plans and situated practices (Suchman, 2006).

This paper investigates such differential spatial production in a Portuguese food distribution centre that the human-environment relationship is mediated by Voice-Directed Warehousing (VDW) system. We introduce the concept of the "adaptive paradox" to name a structural pattern that emerges from the chain of spatial production: the system is designed to be universally empowering, is stabilised in practical level, yet generates differential spatial outcomes across user profiles. Two research questions guide the study: *(1) How does the representation of VDW and warehouse space produce differential spatial practices and lived experiences across users with different linguistic, gender, and experiential profiles?; (2) What adaptive spatial practices do users develop in response to systemic barriers, and in what ways can technology empower equity while mitigating the barriers?*

The paper proceeds as follows. Section 2 situates the theoretical framework. Section 3 describes the methodology. Section 4 presents three empirical findings. Section 5 discusses implications and contributions. Section 6 addresses limitations, and Section 7 presents the conclusion.

2. Theoretical Framework

2.1. Warehouse Operations and VDW Systems as Spatial Infrastructure

Order-picking is the core operation of food distributin centre and known as their most labour-intensive activity, accounting for over half of supply-chain time expenditure (Frazelle, 1996; Grosse et al., 2017; Winkelhaus et al., 2018). Order-picking is a process of retrieving items from storage locations, requiring operators to traverse warehouse areas and locating products across numerous locations while maintaining accuracy under time pressure, hence being labour-intensive. Studied have explored time-efficiency gains through storage-system manipulation (Zhang, 2016), layout and routing optimisation (Đukić et al., 2010; Mendes, 2011; Rakesh & Adil, 2015), and picking-method improvements.

Paper-based lists (PB) were the earliest method which employ manual record on paper. Handheld radio-frequency (RF) scanners replaced paper with digital displays but still occupy operators' hands. Pick-by-light uses illuminated indicators but require significant infrastructure investment. Among these, pick-by-voice systems have demonstrated superior performance improving accuracy up to 25% over PB and RF (Dujmešić et al., 2018).

VDW is a pick-by-voice technology providing warehouse operators with hands-free, audio-based instructions for order-picking task. Operators hear item locations and quantities, navigate to the location, and confirm verbally. It connects to a Warehouse Management System (WMS) that allocates tasks, calculates routes, and records productivity metrics in real time. By enabling hands-free operation and deliver real-time navigation, it aims to reduce physical strain while increase picking speed and accuracy (Dujmešić et al., 2018; Grosse et al., 2015, 2017; Lucchese & Mummolo, 2024).

Despite these operational gains, VDW is not neutral. Warehouse efficiency research typically reports aggregate performance data without accounting for user profile diversity (Dujmešić et al., 2018). Speech recognition trained on dominant phonological norms can produces systematic misrecognition for non-native speakers (Feng et al., 2024). Wayfinding studies suggest that alphanumeric coordinate system, the spatial logic on which warehouse layouts are built, support survey-based navigators which may impose additional cognitive burden on route-based users. Research shows this navigation preference follows gendered patterns (Koulouri et al., 2012; Lawton, 1994; Morag & Parush, 2024), suggesting spatial learning labour distributes unevenly across genders. Merit-based metrics obscure the differential conditions under which productivity is achieved (Garcia et al., 2025). This representation of "universality" can introduce systematic bias against under-represented groups, although whether these patterns hold in any given setting remains an empirical question that this study addresses.

2.2. Spatial Production: An Analytical Framework for Applied Technology

Lefebvre (1991) distinguished spatial production in three registers: (1) conceived space (representations of space), the space of planners encoded in layouts, algorithms, and governance; (2) perceived space (spatial practice), the embodied dimension of everyday navigation and work rhythms; and (3) lived space (representational space), the subjective meanings where identities, resistance, and belonging are contested. This production is iterative: existing condition shapes what users do, and what users do reconfigure space in turn, showing how space is creates reciprocal interplay between planners – space, planners – users and users. The technology mediates these relationships.

Applied to VDW, the triad maps directly onto the warehouse. The conceived space encompasses the speech-recognition architecture, alphanumeric layout, and WMS productivity metrics. Through its underlying algorithm, the system constitutes institutional assumptions of space about who the worker is and how the warehouse should be navigated. The perceived space is the embodied experience of moving through aisles, interpreting audio commands to spatial locations, and sustaining a picking rhythm. The lived space captures workers' subjective dimension, where they negotiate their sense of belonging and fairness in the gaps between the system's design intentions and their daily reality.

Technologies, however, do not produce space uniformly. As Suchman (2006) argued, socio-material arrangements embody assumptions about users, and the same system yields different outcomes depending on who is using it. SCOT (Pinch & Bijker, 1984) conceptualised this as interpretative flexibility: technologies are experienced differently across groups, and those differences are structural rather than incidental. When it exists, context-specific tactics (De Certeau, 1984) on how to navigate space emerges. In Lefebvrian terms, this forms iteration between conceived and perceived space. Users whose needs were not anticipated within the conceived space develop situated responses to reconfigure the technology to their circumstances (Suchman, 2006). These tactics are not individual coping mechanisms but collective spatial practices, underscoring the importance of inclusive and human-centric engineering (Bolton, 2022).

2.3. Equity Discourses in Technology-Mediated Space

SCOT (Pinch & Bijker, 1984) demonstrates how technologies stabilised around dominant groups encode their assumptions as normative. FTS argued this universality politically naturalises inequities through structural absence rather than deliberate exclusion (Criado-Perez, 2019; D'Ignazio & Klein, 2020; Wajcman, 2010; Wajcman, 2004), as social axes such as gender, language, and ethnicity are constitutive of spatial production (Massey, 2009; Pradeep et al., 2024). The interactions with technological infrastructures alter users' positionalities depending on their power relation to those infrastructures (Corchia & Borghini, 2025; Suchman, 2006).

In the VDW context, three elements interlock within equity discourses: (1) human as workforce, (2) VDW as technology, and (3) warehouse as spatial context. First, the workforce consists of heterogeneous profiles across gender, linguistic background, experience and spatial approach. These continuously shape differential outcomes in human-system-space interaction (Stephanidis et al., 2019). Second, a VDW system was not designed with this heterogeneity in mind. This may result to imbalanced spatial performance and experiencing of non-native and accent variants speakers due to impaired recognition when interacting (Feng et al., 2024). Third, a warehouse spatial environment whose standardised zoning, routing, and alphanumeric codification (Đukić et al., 2010) may demand uneven cognitive investment from different operators. Literatures show distributional tendencies in navigation preference across gender, with men perform better in survey-based strategies (i.e., cardinal directions, metric distances, spatial configuration), whereas women prefer route-based strategies more often (landmarks, sequential, left-right cues) (Koulouri et al., 2012; Lawton, 1994; Morag & Parush, 2024). It is important to note that these are distributional tendencies documented in the literature, which individual variation remains substantial and context-dependent.

From this convergence, the "adaptive paradox" emerges as a concept linking technological design and spatial production. We trace how design intentions (conceived) may produce differential operational realities (perceived), experienced as equitable or inequitable (lived) depending on the worker's profile. This study integrates three levels to examine how VDW's system architecture and spatial logic interact or create barriers for users, and what equitable design process looks like in response.

This study adopts an organisational ethnographic approach (Barley, 1986; Ybema, 2009), examining VDW as a technological infrastructure within the spatial and social infrastructure of a warehouse. Reporting addresses 27 of 32 item in the Consolidated Criteria for Reporting Qualitative Research (COREQ) (Tong et al., 2007), while remining five were considered irrelevant.

3. Methodology

3.1. Data Collection

We conducted five days fieldwork in February 2025 at a Portuguese food distribution centre, combining direct observation with semi-structured interviews, all conducted in person. Observation documented human–machine and human–human interactions, spatial navigation practices, and workaround behaviours during active order-picking shifts. Interview guidelines followed Spradley (2011) and incorporated an empathy-map structure (Gray et al., 2010) to discover device usability, situational awareness, learning curves, trust in technology, fairness perceptions, and adaptive strategies. Interviews lasted between 20 and 75 minutes and were conducted in English or Portuguese with internal interpretation support.

Participants were recruited through purposive sampling, facilitated by operational managers. Selection criteria included: (1) minimum six months of VDW experience, (2) diverse linguistic and cultural backgrounds, and (3) balanced gender representation. The final sample comprised 20 participants: 14 operators (7 women, 7 men) and 6 managers (2 women, 4 men). Among them, 12 were native Portuguese speakers and 8 were non-native speakers, with experience ranging from 9 months to 24 years. Participants were informed the aim of studies, and that participation was voluntary. All names used in this paper are pseudonyms.

3.2. Data Analysis

Analysis followed theoretical thematic analysis(Braun & Clarke, 2006) with coding iteration anchored in the human-fatigue dimensions, order-picking challenges ((Grosse et al., 2017; Winkelhaus et al., 2018) and Lefebvrian (1991) triad. Interviews were transcribed and translated using Happyscribe and coded in Atlas.ti through three iterations. First, they were coded descriptively following interview guidelines. After initial coding, they are refined into organised themes on technology usability and its relation to workloads: cognitive load, mental load, and physical load. The coding also consists empathy map to distinguish participant's own experience and their observations. The last iterations are reviewed against the spatial triad framework.

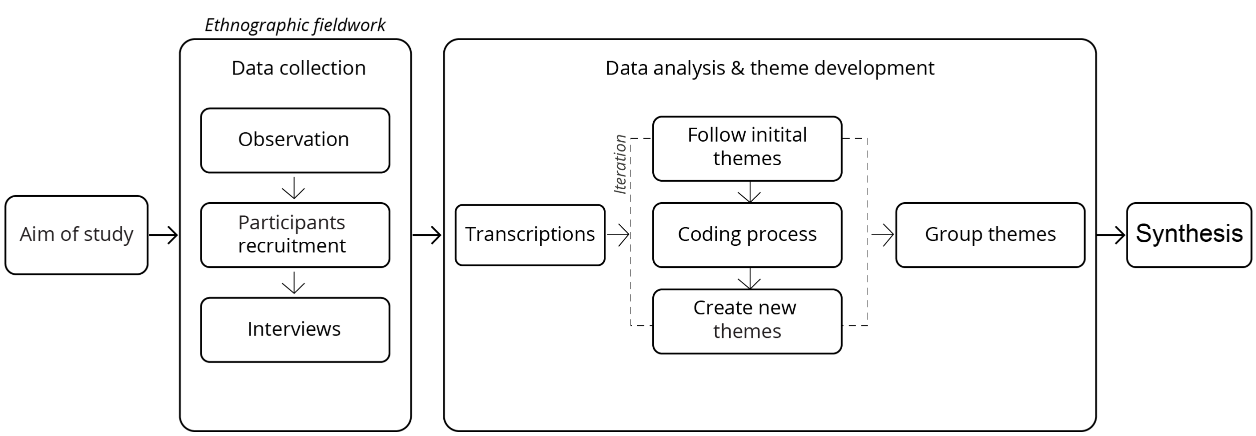

Figure 1: Research Methodology

4. Key Finding

4.1. Voice as Spatial Interface and Paradox in Counter-Representation

- *Conceived Space: The Default Speaker as Spatial Norm*

  VDW is the first dimension of the conceived space: a spatial navigation system whose interface is voice. Audio commands, directing operators to storage locations through alphanumeric cues (i.e., "go to SPL 304"). Operators verify arrival by confirming a check-digit number (i.e., "56, ready"), and VDW response with task quantity (i.e.,"drop 5"). Voice is not a parallel channel alongside movement but the prerequisite for it: without being recognised, operators cannot advance through warehouse space. In conceived space , whose phonological profiles system is calibrated on determines who the default subject is and, those whose movement is less likely interrupted. In this site, that default subject is a native, Portuguese speaker, withwhile English available is available as a derivative option. Management confirms this design boundary:

  > ***Samuel, Manager:*** *"The voice system is the same for everyone. The only difference is Portuguese or English. We thought about introducing the Indian or the Punjabi language, but we couldn't support that because if someone had a problem, we couldn't help the people, so we stick to Portuguese and English."*

  With the technology stabilised around the native speakers, it renders non-native speakers as subsidiary. Their positionalities are compromised as their linguistic patterns were under-represented in the system's representational logic. Interactive speech features such as sequential positioning prompts and re-recording are available to all operators in principle, yet their benefit depends on whether the system can first recognise the user's speech. For them, persistent misrecognition undermines these features before they can function. In Lefebvrian terms, this produces a warehouse whose conceived legibility is available only to those whose voices match the system's training data.

- *Perceived Space: Misrecognition as Spatial Disruption*

  In this site, VDW's speech recognition consistently underperformed for non-native and accented speakers, consistent with Feng et al.'s 2024 findings on automated speech recognition. Because voice is also the navigation interface, each misrecognition event is a moment of spatial disruption. Perceived space constitutes the embodied picking rhythms: receiving an audio command, translating it to a coordinate, navigating to the location, confirming verbally, and proceeding to the next task. The rhythm fractures at each failure, causing operator to halt, redirect cognitive attention from navigation to diagnostic troubleshooting, and restart the interaction. Each failure imposes a dual-burden: operators must diagnose the system failure while simultaneously navigate solutions, disrupting the spatial rhythm within physical space.

  > ***Elon (M, non-native):*** *"When I just start the work, I put the Talkman. But when it says 'ready?' I need to reply many times, 'ready, ready, ready', because sometimes it does not catch (my voice). This is a very big problem. I need to restart it again.."*

  > ***Gaby (F, non-native):*** *"In the first two weeks, I experienced some difficulties with Talkman…because it was my first time worked with Voice-picking, so I didn't really know the steps. When I said the*

numbers, it was always saying something else, so I repeated them over and over again. I had to speak a lot of times the same thing. I rerecorded my voice after three or four weeks."

Table 1: Breakdown Types and Remedial Strategies Observed.

Breakdown	**Possible Commands or Strategies**
Misrecognition	Repeating commands
Pronunciation	"Repeat", "say again", modulating pitch, change recording script
Ambient noise	"Repeat", "pause", "say again"
Peer conversations	"Repeat", "pause", "say again", "last preparation", "last put"
Hardware or network malfunctions	Waiting for network, change device

Failures present as opaque breakdowns with multiple potential causes and require different remedial strategies as shown in Table 1. New operators record their voice on the second day of induction, then receive two weeks of assisted practice. the system reliably recognises an operator's voice during this period directly shapes how quickly they develop spatial fluency. For non-native speakers, persistent misrecognition risks extending this adaptation beyond the training window, while native speakers reported to settle the rhythm within two to three weeks.

> ***Amana (F, native):*** *For me it was two, three weeks... I started doing faster than the ones that came before with me, because we don't have the same rhythm."*

- *Lived Space: Counter-Representation as Spatial Agency*

  Finding 1 shows the most significant spatial production in operators' active renegotiation tactics to the system's spatial representation. Facing persistent misrecognition, operators subtitute prescribed vocabulary with individual alternatives to improve recognition reliability.

> ***Olivia (F, non-native):*** *In my case, one is 'uno'. Number one. Because of that 'one' is 'un', and the sound doesn't go very well. And many people change it. Some put letters, some say 'one', some say uno."*
>
> ***Carla (F, non-native):*** *And then, for example, my colleague. His voice is totally Creole. If I get him to speak here, his manner and his voice, the way he commands, it's all in Creole, most of the time, if you pay attention, it's in Creole. So he recorded it in Creole."*

These substitutions constitute counter-representations of space through acts of rewriting conceived design to reclaim spatial access that excludes their (vocal) identities. Operators do not passively accept the system's representational logic; instead, they renegotiate it from within. This renegotiation is, however, invisible to management, uncompensated in productivity metrics, and unrecognised as skill. Non-native operators must maintain their individual vocabulary layer as an additional cognitive burden on top of the primary navigation task; a structural consequence of having been absent from the conceived design of technological space.

4.2. Finding 2: Standardised Layout and Gendered Spatial Learning

- *Conceived Space: The Warehouse as a Codified Terrain*

  Warehouse layout constitutes a second dimension of conceived space: standardised parallel aisles, zoning system, and dedicated routes designed for throughput efficiency and traffic-collision mitigation. The signage employs alphanumeric systems (i.e., "SPL 304") that encodes every position as an abstract coordinate which VDW delivers it as audio commands. With minimal distinctive landmark cues aside from positioning the alphanumeric signage in odd-even patterns, the design is consistent to support survey-based navigation cues.

  Although numbers encode zone and position information (i.e., "584" means the drinks section (5), on the right side), the physical environment does not specify the zone visually. The warehouse displays as codified terrain, which only navigable for those who have memorised the structure and built mental maps linking signage to physical locations. Consequently, we suggest that the design demands more spatial cognitive labour to adapt for route-navigators, which is unevenly distributed across the workforce.

- *Perceived Space: Extended Adaptation and Differential Legibility*

  Wayfinding studies (Koulouri et al., 2012; Lawton, 1994; Morag & Parush, 2024) argued that navigation performance across gender differ depending on cue preferences, where men prefer survey-based while women prefer route-based strategies. The warehouse design of grided aisles is using metrics and cardinal direction, thus according to studies, is structurally optimised for survey-based navigators. Route-based navigators may find aisles appear identical, requiring them to construct mental maps from limited spatial anchors such as the memorising odd-even number patterns on left-right side and learn which zone and aisle sections the number falls. In the accounts gathered in this study, men are described to have faster adaptation compared to women who started at the same time. A senior operator who trained new recruits, described differential learning trajectories, not only because of the interaction with VDW, but also in understanding the workflow:

Figure 2: Numerical Signage in the Warehouse

> ***Bram (M, native):*** *I trained four people, two men, two women. One of them already has a lot of have experience from the same work. The other one was like a lot of confusion, but I teach him. (and then) he was good. The two guys, it's okay, they are working now. but (for) the two women they are not catching. The first one had more difficulty to understand all the work. It's not only the voice, but she never worked in our house, so she's very difficult (for her) to catch (understand)"*

Accordingly, we argue that survey-based navigators map alphanumeric commands onto their mental models more efficiently, whereas route-based navigators must sustain an additional cognitive translation. Linking to gendered patterns, warehouse's spatial design influence longer learning trajectories on women since it may not be designed to support their orientation preferences. Evidence shows that only women who reported extended adaptation phases up to two months, well beyond the formal two-week training period, while most men in the sample reported achieving comfort within the training window or shortly after.

> ***Carla (F, non-native):*** *"...after a week I think was okay. But to feel really comfortable or to get the real top of the mountain, I can say two months. Then I know that I am prepared."*
>
> ***Olivia (F, non-native):*** *"I think I reached productivity in about two months. I need to become more adept with the palettes."*
>
> ***Francis (M, native):*** *"I think maybe two, three days I can go by ourselves...In fact, if more than two, three days, everything is the same. So we need to go alone to work. then we learn by ourselves."*

This reflects structural mismatch between warehouse's universal spatial logic and diverged users' navigation strategies. However, we interpret this pattern cautiously: although the finding consistent with gendered pattern wayfinding studies we cannot isolate navigation style from confounding factors such as prior experience, job familiarity, or individual learning differences.

- *Lived Space: Gendered Geography on Belonging and Precarity*

  Finding 2 concerns on how the warehouse design structured for universal legibility actually illegible for some, reflecting to gendered geography (Massey, 2009). It distributes uneven cognitive efforts and produce differential belonging around competence among operators. Productivity metrics assessed during and after the training period register disadvantage without accounting for structural causes. Additional spatial learning labour may cause productivity deficiency which could influence employment decision. The conceived space thus may participate in making certain operators' positions more precarious.

4.3. Finding 3: Meritocratic Governance and Tactical Manipulation

- *Conceived Space: The Algorithm as Spatial Authority*

  A third dimension of conceived space is the WMS: an institutional representation as a fair, objective and universal. Tasks are allocated in randomised sequences to prevent supervisor bias, and performance is measured in boxes per hour as a standardised metric. The algorithm functions as spatial authority, deciding where operators go and how their output is counted without reference to the social, physical, or cognitive conditions under which that output is produced. This meritocratic framing performs a double spatial operation. It imposes standardised conditions that may reproduce inequity, then frames those conditions as evidence that inequity does not exist. Operators who bear the productivity cost of speech misrecognition (Finding 1) or extended navigation adaptation (Finding 2) have those costs attributed to individual performance deficit (having less boxes per hour) rather than systemic spatial limitations. The representation of the warehouse as fair and objective is itself a form of spatial production, concealing the imbalances it reproduces.

- *Perceived Space: Tactical Manipulation of Spatial Rhythms*

  Operators distinguish between "good pallets" (lighter, higher-quantity; i.e., 100 boxes of cereals) and "bad pallets" (heavier, lower-quantity; i.e., 30 boxes of soda), recognising randomised allocation does not prevent the accumulation of physically demanding tasks over time which link to how those yield to their productivity scores. This distincition reveals how operators read the system's logic and act within it. When they experience the gap between claimed fairness and lived condition, they respond through tactical spatial practices (De Certeau, 1984): deliberately manipulating of movement rhythms and break scheduling to increase the chance of receiving more favourable task assignments.

  > ***Nuno (M, non-native):*** *"To make productivity, we need to put more boxes. We need to make fast or be smart and try to get some good pallets, like a pallet with a high number, for example 100 boxes. For that, you can go to the bathroom and wait for a better one."*

  The claimed 'universality' apparently raised a fairness issue. Weight, however, was raised only by women participants, which may reflect either an actual differential physical burden or gendered norms around disclosing physical difficulty.

  > ***Mariana (F, native):*** *"...the strength difference is more difficult here. I think for some women, it's a little bit more difficult because of the body strength... men should be in a place where it is heavier, otherwise it's difficult. Some women have more strength, but others, no. I have more strength than some guys, so it's not a problem. Normally, girls have a little strength, but more mindful. So I think stay clever (for women)."*

  > ***Laura (F, native):*** *"...here it's the same in all areas. I am saying that's not fair, because, as I said, some boxes are heavy, and some boxes are not so heavy. So, on average, it should be different."*

- *Lived Space: The Moral Equity of the Warehouse*

  Finding 3 reveals how fairness is diverged between system's claim and operators' lived experience. Tactical manipulations through slowing down or timing break are not individual misconducts but rational spatial agencies, reflecting collective knowledge that the metric does not deliver what it claims. However, these tactics carry consequences. Operators whose productivity metrics are already compromised by misrecognition or extended navigation adaptation may face greater repercussions when engaging in tactical manipulation. This may generate a vicious cycle of resistance where operators must reproduce the very space that previously excluded them and deploy spatial agency that ultimately results in reprimands against themselves.

5. Discussion

5.1. Adaptive Paradox as a Transferable Conceptual Lens

Across the three findings, the adaptive paradox operates through consistent Lefebvrian lens: conceived spaces encoded with normative assumptions produce perceived experiences that diverge by demographic profile, generating lived conditions of inequity and contested belonging. The paradox is structural rather than incidental, emerging from the absence of minority users from the conceived space at the moment of design. As Star (1999) argued, infrastructure becomes visible upon breakdown thus its apparent neutrality conceals the social relations it encodes. Findings illustrate axis of this absence: voice as spatial interface produces immobilisation for non-native speakers whose vocal profiles were absent from the design; codified territory produces illegibility for workers whose spatial literacy requires extended cognitive investment; and merit-based metric compress differential spatial knowledge into number which operators deem as injustice.

These findings extend Lefebvre's spatial production into technology-mediated work, suggesting that algorithmic system and spatial infrastructures co-produce differential environments for different social subjects. They provide empirical illustrations of interpretive flexibility (Pinch & Bijker, 1984), indicating that VDW is not experienced neutrally but in ways that, in this site, favoured dominant profiles. Findings also ground FTS' claims (D'Ignazio & Klein, 2020; Wajcman, 2010; Wajcman, 2004) in an organisational setting, illustrating how default design assumptions can reproduce socio-spatial hierarchies in everyday practice.

We propose that the adaptive paradox is not specific to VDW. It describes a structural tendency that may recur in socio-technical systems calibrated to dominant user profiles across sectors, unless diverse spatial and user knowledge is incorporated from the outset of design. Nevertheless, the extent of generalisation requires further investigation.

5.2. Spatial Production Triad as Governance Framework

Each register of the spatial triad yields governance implications (Table 1). At the conceived level, the default assumptions embedded in system architecture must be restated. For VDW, speech recognition should be trained on linguistically diverse profiles, introduce differential navigation, wayfinding or signage strategies for warehouse, and productivity metrics should account for differential conditions.

At the perceived level, workers whose profiles diverge from the default bear disproportionate adaptation burdens that remain invisible. Multimodal interfaces combining audio with visual or haptic cues, transparent diagnostics on failures, and workflow customisation could reduce cognitive overhead and allow workers to develop spatial practices with greater autonomy.

At the lived level, workers' experiential knowledge of system failures, workarounds, and inequitable outcomes represent expertise that aggregate metrics cannot capture. Organisations should formalise channels for this knowledge through co-design workshops or structured feedback mechanisms so operators contribute actively to system design rather than passive recipients. Critically, this engagement should precede deployment because the adaptive paradox originates when certain user profiles are excluded at the moment of design closure.

Table 2: *Synthesis of Findings Across the Lefebvrian Spatial Triad*

Finding	Conceived Space	Perceived Burden	Lived Response	Design Implication
1. Voice as spatial interface	Native Portuguese as default speaker; system calibrated on dominant phonological profiles	Persistent misrecognition fractures picking rhythm; dual-task cognitive burden; extended adaptation for non-native speakers	Phonological substitutions (e.g., 'x' for 'one'; Creole commands) as counter-representation of space	Train speech recognition on linguistically diverse profiles; transparent diagnostics for recognition failures
2. Spatial learning labour	Alphanumeric grid layout with minimal landmarks; survey-based navigation as spatial norm	Extended adaptation for women participants; perhaps due to requiring additional cognitive translation in navigation	Spatial uncertainty; differential belonging; potential productivity penalty during evaluation period	Advance visual landmarks; multimodal navigation support; customise workflow
3. Tactical resistance	Randomised task allocation; boxes-per-hour as universal metric; algorithm as spatial authority	Uniform metrics mask differential physical and cognitive conditions; 'good' vs. 'bad' pallet distinction	Tactical manipulation of breaks and timing; collective understanding of systemic unfairness; risk of reprimands	Capacity-aware task allocation; participatory metric redesign; formalise worker feedback channels

6. Limitation and Future Research

Fieldwork was conducted over five days in a single Portuguese distribution centre. The short window means some dynamics were captured as snapshots rather than processes. Interviews in Portuguese were conducted with interpretation support, which may not only introduced nuance of misinterpretation but also may influence participants' willingness to disclose critical views. We attempted to mitigate this by assuring confidentiality, employed in-interview verification and corrections. We acknowledge the exploratory nature of this study and position the findings presented here as preliminary insights from an exploratory study rather than conclusive generalisations. Future research should pursue comparative studies across organisational and national contexts to identify which features of the adaptive paradox are specific to VDW and which generalise to other technology-mediated work environments. Intersectional analyses combining gender with linguistic background, age, disability, and socioeconomic status would enrich the equity framework. Testing the adaptive paradox as a conceptual lens in other project-based and digitally mediated settings is also a promising direction. Larger samples with more systematic demographic data would enable quantitative triangulation of the qualitative patterns identified here.

7. Conclusion

This paper examined equity challenges in Voice-Directed Warehousing through the adaptive paradox: the pattern by which conceived system designed for universal efficiency produces disproportionate spatial burdens on users whose profiles were absent from the moment of design. Drawing on Lefebvre's spatial triad, SCOT, and FTS, we traced how designed neutrality becomes operational inequity and lived injustice.

Across three findings (spatial immobilisation through speech-recognition bias, uneven spatial learning labour through standardised codification, and concealed inequality through merit-based metric), human–VDW collaboration emerges as socially contested rather than neutral. The system's asymmetries were rendered invisible by the universality rhetoric, concealing how default assumptions favoured dominant user profiles and produced differential spatial experiences for others. Addressing the adaptive paradox requires beyond technical optimisation. Decisions about whose voice a system recognises, whose movement patterns it accommodates, and whose performance it measures are spatial and political questions. The adaptive paradox offers a transferable lens, a conceptual foundation towards equitable human-centricity beyond efficiency metrics.

References

Barley, S. R. (1986). Technology as an Occasion for Structuring: Evidence from Observations of CT Scanners and the Social Order of Radiology Departments. *Administrative Science Quarterly, Mar., 1986, Vol. 31, No. 1 (Mar., 1986), pp. 78-108*, (Sage Publications, Inc. on behalf of the Johnson Graduate School of Management, Cornell University).

Bolton, M. L. (2022). Humanistic Engineering: Engineering for the People. *IEEE Technology and Society Magazine*, *41*(4), 23–38. https://doi.org/10.1109/mts.2022.3219132

Braun, V., & Clarke, V. (2006). Using Thematic Analysis in Physchology. *Qualitative Research in Psychology*, *3*(2), 77–101. Retrieved December 4, 2025, from https://doi.org/10.1191/1478088706qp063oa

Commission, E. (2021). Industry 5.0: Towards a sustainable, human-centric and resilient European industry. https://data.europa.eu/doi/10.2777/308407

Corchia, L., & Borghini, A. (2025). Infrastructure as a sociological category: Concept, applications, and paradigmatic turns? *Journal of Classical Sociology*, *25*(2), 123–151. https://doi.org/10.1177/1468795X251327051

Criado-Perez, C., author. (2019). *Invisible women : Data bias in a world designed for men*. New York : Abrams Press, [2019] ©2019. https://search.library.wisc.edu/catalog/9912685038702121
Includes bibliographical references (pages 322-391) and index.

De Certeau, M. (1984). *The Practice of Everyday Life*. Berkeley: University of California Press.

Degens, P., Hilbrich, I., & Lenz, S. (2022). Analyzing Infrastructures in the Anthropocene [Version Number: 1.0]. *Historical Social Research*, *47*, 728. https://doi.org/10.12759/HSR.47.2022.36
SeriesInformation Historical Social Research, 47(4), 7-28.

D'Ignazio, C., & Klein, L. F. (2020). *Data Feminism*. MIT Press. Retrieved December 4, 2025, from https://direct.mit.edu/books/book/4660/Data-Feminism

Dujmešić, N., Bajor, I., & Rožić, T. (2018). Warehouse Processes Improvement by Pick by Voice Technology. *Tehnicki vjesnik - Technical Gazette*, *25*(4). https://doi.org/10.17559/TV-20160829152732

Đukić, G., Česnik, V., & Opetuk, T. (2010). Order-picking Methods and Technologies for Greener Warehousing. *Strojarstvo*, *52*(1), 23–31.

Feng, S., Halpern, B. M., Kudina, O., & Scharenborg, O. (2024). Towards inclusive automatic speech recognition. *Computer Speech & Language*, *84*, 101567. https://doi.org/10.1016/j.csl.2023.101567

Frazelle, E. H. (1996). *World-Class Warehousing and Material Handling*. Atlanta, Ga. : Logistics Resources International.

Garcia, S. C., Dhanani, L. Y., & Wiese, C. W. (2025). Merit thrives under evidence-based DEI practices and disparate impact protections. *Industrial and Organizational Psychology*, *18*(3), 320–328. https://doi.org/10.1017/iop.2025.10021

Graham, S. (2010). *Disrupted Cities: When Infrastructure Fails* (1st). Routledge.

Gray, D., Brown, S., & Macanufo, J. (2010). *Gamestorming: A Playbook for Innovators, Rulebreakers, and Changemakers*. O'Reilly Media, Incorporated. https://books.google.de/books?id=3OCbAgAAQBAJ

Grosse, E. H., Glock, C. H., Jaber, M. Y., & Neumann, W. P. (2015). Incorporating human factors in order picking planning models: Framework and research opportunities. *International Journal of Production Research*, *53*(3), 695–717. https://doi.org/10.1080/00207543.2014.919424

Grosse, E. H., Glock, C. H., & Neumann, W. P. (2017). Human factors in order picking: A content analysis of the literature. *International Journal of Production Research*, *55*(5), 1260–1276. https://doi.org/10.1080/00207543.2016.1186296

Koulouri, T., Lauria, S., Macredie, R. D., & Chen, S. (2012). Are we there yet?: The role of gender on the effectiveness and efficiency of user-robot communication in navigational tasks. *ACM Transactions on Computer-Human Interaction*, *19*(1), 1–29. https://doi.org/10.1145/2147783.2147787

Lawton, C. A. (1994). Gender differences in way-finding strategies: Relationship to spatial ability and spatial anxiety. *Sex Roles*, *30*(11-12), 765–779. https://doi.org/10.1007/BF01544230

Lee, C. P., & Schmidt, K. (2018). A bridge too far? Critical remarks on the concept of 'infrastructure' in CSCW and IS. *Socio-Informatics: A Practice-based Perspective on the Design and Use of IT Artifacts*, (Oxford: Oxford University Press), 177–217.

Lefebvre, H. (1991). *The Production of Space* (D. Nicholson-Smith, Trans.; Reprinted). Blackwell.

Lucchese, A., & Mummolo, G. (2024). Human-Centric Assistive Technologies in Manual Picking and Assembly Tasks: A Literature Review. *Management and Production Engineering Review*. https://doi.org/10.24425/mper.2024.151132

Massey, D. B. (2009). *Space, place, and gender* (6. print). Univ. of Minnesota Press.

Mendes, P. (2011). *Demand Driven Supply Chain: A Structured and Practical Roadmap to Increase Profitability*. Springer Berlin, Heidelberg. Retrieved March 30, 2026, from https://doi.org/10.1007/978-3-642-19992-9

Morag, I., & Parush, A. (2024). The effect of verbal instructions while using digital indoor wayfinding devices on gender, performance, and self-reported strategies. *Applied Ergonomics*, *114*, 104160. https://doi.org/10.1016/j.apergo.2023.104160

Perez, Caroline Criado. (2019). *Invisible Women*. New York: Abrams Press.

Pinch, T. J., & Bijker, W. E. (1984). The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology might Benefit Each Other. *Social Studies of Science*, *14*(3), 399–441. https://doi.org/10.1177/030631284014003004

Pradeep, K., Jisha Vg, Baiju Paul, & Paul T Benziker. (2024). Gendered Space and Spatial Discourses in Everyday Life: Exploring The Notion of Henri Lefebvre. *International Research Journal on Advanced Engineering and Management (IRJAEM)*, *2*(11), 3273–3281. https://doi.org/10.47392/IRJAEM.2024.0482

Rakesh, V., & Adil, G. K. (2015). Layout Optimization of a Three Dimensional Order Picking Warehouse. *IFAC-PapersOnLine*, *48*(3), 1155–1160. https://doi.org/10.1016/j.ifacol.2015.06.240

Spradley, J. P. (2011). *The ethnographic interview* (Nachdr). Wadsworth.

Ethnographic Interviews(1/6/2025, 3:14:20 PM) "Participant Observation (Spradley 1980)," (Spradley, 2011, p. iv) "My interest in this approach began from a rather simple observation: some tasks are best accomplished before other tasks when doing ethnography." (Spradley, 2011, p. iv) "Part Two, " The Developmental Research Sequence,"" (Spradley, 2011, p. iv) "People everywhere learn their culture by observing other people, listening to them, and then making inferen ces." (Spradley, 2011, p. 8) "reasoning from evidence (what we perceive) or from premises (what we assume)." (Spradley, 2011, p. 8) "In doing field work, ethnographers make cultural inferences from three sources: (1) from what people say; (2) from the way people act; and (3) from the artifacts people use." (Spradley, 2011, p. 8) "None of the sources for making inferences—behavior, speech, artifacts—are foolproof," (Spradley, 2011, p. 8) "The ethnog rapher must then make inferences about what people know by listening carefully to what they say, by observing their behavior, and by studying artifacts and their use." (Spradley, 2011, p. 9) "the ethnographer must draw the generalizations for himself, must formulate the abstract statement without the direct help of a native informant" (Spradley, 2011, p. 9).

Star, S. L. (1999). The Ethnography of Infrastructure. *American Behavioral Scientist, Vol. 43 No. 3*, *Sage Publications, Inc.*(November/December 1999), 377–391.

Stephanidis, C., Salvendy, G., Antona, M., Chen, J. Y. C., Dong, J., & Duffy, V. G. (2019). Seven HCI Grand Challenges. *International Journal of Human–Computer Interaction*, *35*(14), 1229–1269. https://doi.org/10.1080/10447318.2019.1619259

Suchman, L. (2006). *Human-machines reconfigurations: Plans and situated actions* (2nd ed). Cambridge University Press.

Tong, A., Sainsbury, P., & Craig, J. (2007). Consolidated criteria for reporting qualitative research (COREQ): A 32-item checklist for interviews and focus groups. *International Journal for Quality in Health Care*, *19*(6), 349–357. https://doi.org/10.1093/intqhc/mzm042

Wajcman, J. (2010). Feminist theories of technology. *Cambridge Journal of Economics*, *34*(1), 143–152. https://doi.org/10.1093/cje/ben057

Wajcman, J. (2004). *TechnoFeminism* (repr). Polity.

Winkelhaus, S., Sgarbossa, F., Calzavara, M., & Grosse, E. (2018). The effects of human fatigue on learning in order picking: An explorative experimental investigation. *IFAC-PapersOnLine*, *51*(11), 832–837. https://doi.org/10.1016/j.ifacol.2018.08.442

Ybema, S. (Ed.). (2009). *Organizational ethnography: Studying the complexities of everyday life*. SAGE.

Zhang, Y. (2016). Correlated Storage Assignment Strategy to reduce Travel Distance in Order Picking□□Zhang thanks the financial support of the china scholarship council (CSC). *IFAC-PapersOnLine*, *49*(2), 30–35. https://doi.org/10.1016/j.ifacol.2016.03.006