Juggling Extra Limbs: Identifying Control Strategies for Supernumerary Multi-Arms in Virtual Reality

Hongyu Zhou
School of Computer Science
The University of Sydney
Sydney, NSW, Australia
hzho4130@uni.sydney.edu.au

Tom Kip
The School of Computer Science
The University of Sydney
Sydney, NSW, Australia
vkip4651@uni.sydney.edu.au

Yihao Dong
School of Computer Science
The University of Sydney
Sydney, NSW, Australia
yihao.dong@sydney.edu.au

Andrea Bianchi
Industrial Design
KAIST
Daejeon, Republic of Korea
School of Computing
KAIST
Daejeon, Republic of Korea
andrea.whites@gmail.com

Zhanna Sarsenbayeva
School of Computer Science
The University of Sydney
Sydney, NSW, Australia
zhanna.sarsenbayeva@sydney.edu.au

Anusha Withana
School of Computer Science
The University of Sydney
Sydney, NSW, Australia
anusha.withana@sydney.edu.au

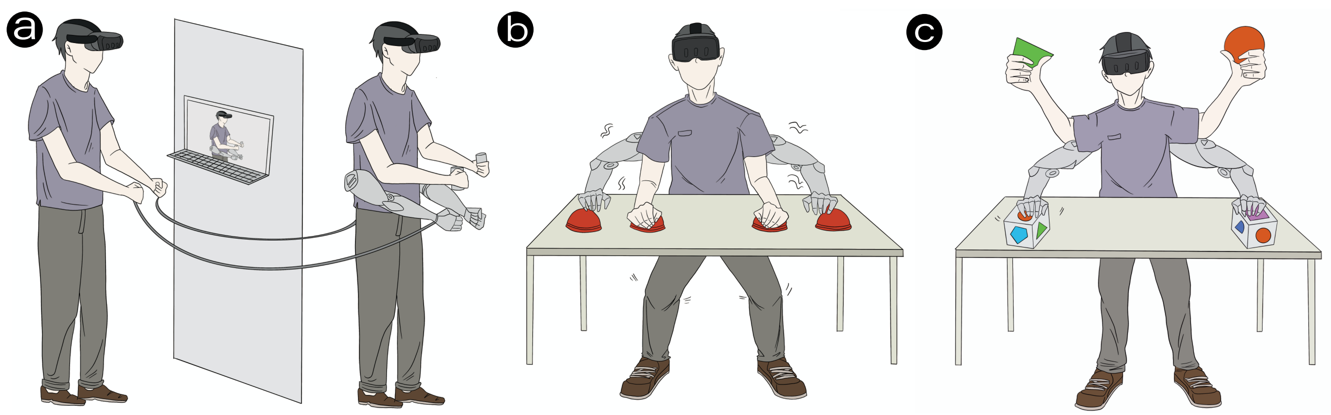

Figure 1: Wizard-of-Oz approach and experimental tasks. (a) The participants' VSLs were controlled by a human operator behind a partition to simulate semi-autonomous supernumerary limbs. (b) Basic Control Task: In this task, participants used both their own arms and the VSLs to press buttons. (c) Factory Task: Participants used their own arms and the VSLs to collaboratively insert found objects into the appropriate holes.

Abstract

Using supernumerary multi-limbs for complex tasks is a growing research focus in Virtual Reality (VR) and robotics. Understanding how users integrate extra limbs to achieve shared goals is crucial for developing efficient supernumeraries. This paper presents an exploratory study (N=14) investigating strategies for controlling virtual supernumerary limbs with varying autonomy in VR object manipulation tasks. Using a Wizard-of-Oz approach to simulate semi-autonomous limbs, we collected qualitative and quantitative data. Results show participants adapted control strategies based on task complexity and autonomy, affecting task delegation, coordination, and body ownership. Based on these findings, we propose guidelines—commands, demonstration, delegation, and labeling—to improve multi-limb interaction design by adapting autonomy to user needs and fostering context-aware experiences.

CHI '25, Yokohama, Japan

ACM ISBN 979-8-4007-1394-1/25/04
https://doi.org/10.1145/3706598.3713647

CCS Concepts

• **Human-centered computing** → **Virtual reality**; **Empirical studies in HCI**; **User studies**.

Keywords

Embodied Interaction, Virtual/Augmented Reality, Empirical study that tells us about how people use a system, Interaction Design

ACM Reference Format:

Hongyu Zhou, Tom Kip, Yihao Dong, Andrea Bianchi, Zhanna Sarsenbayeva, and Anusha Withana. 2025. Juggling Extra Limbs: Identifying

Control Strategies for Supernumerary Multi-Arms in Virtual Reality. In *CHI Conference on Human Factors in Computing Systems (CHI '25), April 26–May 01, 2025, Yokohama, Japan.* ACM, New York, NY, USA, 16 pages. https://doi.org/10.1145/3706598.3713647

1 INTRODUCTION

Leveraging multiple supernumerary limbs to perform complex tasks is becoming a popular research domain in application areas such as Virtual Reality (VR) and robotics [50, 56–58]. To harness the potential of supernumerary limbs in addressing complex tasks, prior research has predominantly employed fixed control mappings for operating the supernumerary limbs, where specific commands are assigned to the movement of particular body parts [65, 70]. However, while fixed control mappings can be effective for certain tasks, they lack the flexibility required for dynamic interactions in complex environments, especially when supernumerary limbs are given *a level of autonomy.*

To facilitate the exploration of dynamic control mappings, it is crucial to understand how users would effectively coordinate these semi-autonomous supernumerary limbs with their own limbs to expand human capabilities and address the unique challenges of interaction design and user experience [19] in a multiple supernumerary limbs setting. In this study, we examine the use of supernumerary limbs within a VR environment to simulate interactions that could inform future applications. VR provides a valuable platform for studying and evaluating the integration of supernumerary limbs, offering a controlled, immersive space to simulate complex scenarios that are difficult to replicate in real-world robotic applications [3, 41, 64]. This environment allows users to engage in a "suspension of disbelief [52]," which supports the study of embodied interaction in novel body configurations. In this paper, we deal with "virtual supernumerary limbs (VSLs)" to explore VR-based interactions rather than real-world robotic applications. Although there are clear differences between VR and real-world environments, insights gained from VR-based studies can potentially inform future developments not only of VR applications, but also physical mechanical limbs, as shown in previous work [62, 82].

Particularly, this paper focuses on how VSLs, when equipped with different and increasing degrees of autonomy, can assist users in managing non-repetitive, multi-layered tasks in VR environments. Instead of simply mapping another limb (e.g. foot movement [67]) or executing predefined commands [66], autonomous VSLs in our study learn, adapt, and act semi-independently to respond to changing task requirements and conditions. This adaptability is especially crucial for scenarios involving real-time decision-making and task variations, where the user needs more responsive control than what fixed mappings can provide [67]. This shift from fixed mappings to autonomous, context-sensitive control can elicit unique user interaction strategies, highlighting the influence of different levels of autonomy in VSLs on user behavior.

To this end, we implemented a Wizard-of-Oz setup to simulate autonomous VSLs that respond to user input and task complexity. Integrating these autonomous capabilities allowed us to freely explore the design space in breadth and depth without restrictions imposed by limitations of today's technologies. Our study focused on key research questions surrounding interaction strategies and system autonomy of multiple VSLs control: **RQ1**: *What control strategies do users employ when interacting with semi-autonomous VSLs in different scenarios?* **RQ2**: *How does the level of autonomy in VSLs affect users' ability to manage tasks?* and **RQ3**: *How do users' perceptions of embodiment and control change with level of autonomy?*

To capture a comprehensive understanding of user interaction, our study blended qualitative and quantitative research methods. Qualitative data were gathered through think-aloud and semi-structured interviews, providing insights into the participants' thought processes. Quantitative data on task performance and the awareness of embodiment were collected to quantify the effectiveness of different control strategies and their impact on user performance and perception. Furthermore, we analyzed operator (or wizard) reaction patterns to validate the operator's behavior, ensuring their actions were consistent and adhered to predefined patterns. This analysis supports the conclusion that our task performance results are based on participant behavior, unaffected by operator errors. This mixed approach revealed several key findings: users favored task delegation to the VSLs for repetitive or lower-priority actions, adjusted their strategies to switch between manual control and automation depending on task complexity, and developed more efficient coordination as they grew familiar with the system. Additionally, users found object labeling and categorization to be crucial in managing VSLs effectively, and they experienced shifts in embodiment based on the level of system autonomy. These findings underscore the need for adaptive control mechanisms that help users balance attention, streamline task-switching, and enhance the sense of embodiment, informing design recommendations for future systems employing multiple VSLs.

2 RELATED WORK

2.1 Virtual Reality and Robotic Body

Recent advancements in robotics and VR technologies have enabled novel interactions where users can experience the detachment, attachment, and exchange of body parts, even while retaining their natural limbs. For instance, Iwasaki et al. [40] introduced the concept of a "detachable body", where a third and fourth robotic limb can be seamlessly attached and detached, allowing users to operate these limbs remotely across different locations. Additionally, research on supernumerary robotic limbs has explored task-specific customization of limb end-effectors, enhancing adaptability and functionality in varied contexts. VR and wearable Shape-Memory Alloys (SMA) have further expanded these possibilities by enabling shared body experiences, where two individuals can collaboratively control a single body(e.g., [66, 75, 91]. For example, Saraiji et al. created a system where wearable robotic arms are remotely operated by another person, facilitating a shared embodiment [35, 66]. Takizawa et al. extended this exploration into VR, investigating how different perspectives impact the shared body experience [75].

Building on these developments, VR provides unprecedented opportunities to explore more abstract states of being a digital cyborg, such as inhabiting an invisible body [44], existing as separated body parts [43], or even duplicating oneself [50]. These immersive experiences push the boundaries of body ownership and self-perception, offering new avenues to understand how users adapt to technologically augmented bodies. Recent studies providing the experience of acquiring virtual robotic limbs [24, 39]. Arai et al. [3] has demonstrated that augmenting a participant's avatar with an extra body part in a virtual environment can evoke the sense of embodiment, leading participants to perceive the virtual limb as an extension of their own body.

Inspired by these findings, our research leverages VR technology to investigate how the addition of multiple semi-autonomous VSLs influences user interaction. To further quantify the effects, we employed an additional questionnaire to measure the sense of agency and ownership [3].

2.2 Supernumerary Robotic Limbs

Research on supernumerary robotic limbs (SRL) has explored various forms of augmenting the human body [34, 63, 65, 81], such as robotic legs [59], fingers [34], and arms [81]. Our research focuses on virtual supernumerary limbs, which add extra limbs to enhance functionality in a VR environment.

Initially, SRL were developed for industrial applications, such as assisting workers in construction and assembly tasks [56–58]. The field has since expanded to include designs like robotic tails for balance [47, 53] and additional limbs for complex tasks [18, 63, 77, 86]. For example, Maekawa et al. [47] introduced a wearable robotic tail to aid in balance, showing its potential in physically demanding activities.

Control mechanisms for SRL have been a focal point of investigation, with various methods proposed to enable intuitive and effective operation. Techniques such as remapping body motions (e.g., foot or shoulder movements) [65, 67, 70] and utilizing brain-computer interfaces for direct limb control [60] have been explored. Despite the promise of these approaches, the majority of studies rely on fixed mapping schemes that limit adaptability in dynamic environments. Moreover, there is a noted lack of comprehensive user studies, with many systems validated through proof-of-concept demonstrations involving single users [65, 80, 81]

To address these limitations, our research investigates the interaction dynamics involved in controlling multiple sumernumeary limbs within a VR environment s) using a non-fixed mapping scheme for supernumerary arms capable of different levels of autonomy. In our study, participants will utilize voice and gestural commands to dynamically switch control between two VSLs while performing semi-realistic tasks. This approach aims to provide new insights into flexible and responsive multiple VSLs control for practical usages, contributing to a deeper understanding of human-machine interaction in increasingly complex contexts.

2.3 Mapping Control Strategies in SRL

Existing research has primarily focused on fixed-mapping control schemes, where SRL are governed by pre-determined mappings, such as proprioceptive interfaces like foot pedals [67]. While these methods have demonstrated effectiveness in controlled environments, they struggle to accommodate the dynamic nature of real-world scenarios, where user intentions and tasks may shift rapidly. Similarly, another study on artificial limb substitution added two robotic arms controlled by leg movements, expanding the user's capacity to four arms [66]. However, such systems still lack the flexibility required for real-time adaptation, particularly when users must manage multiple limbs simultaneously. Similarly, efforts to enhance physical capabilities through SRL, such as the development of robotic tails for balance augmentation, underscore their potential but fall short in addressing the complexities of dynamic, multi-limb control [47]. Additionally, studies on wearable robotic arms examined user-preferred interaction styles but focused on controlled environments, leaving the need for flexible, real-time control largely unaddressed [51].

Our study design is inspired by existing research on user-preferred interaction methods and the control of multiple VSLs. Building on these foundations, our research shifts the focus to non-fixed mapping, enabling users to dynamically manage and seamlessly switch control among multiple VSLs in real-time.

2.4 Wizard-of-Oz Methodology in Supernumerary Robotic Limbs Research

The Wizard-of-Oz (WoZ) methodology is widely used in human-computer interaction research to simulate autonomous systems, enabling the study of user interactions before full prototypes are available [2, 84]. For example, Diederichs et al. used a WoZ vehicle to explore human interaction with AI-driven cars, providing insights into user trust and adaptability [16]. Similarly, Hu et al. developed "Wizundry," a platform for simulating speech-based interfaces with multiple wizards, facilitating the study of complex human-AI collaboration [33]. In SRL research, Muehlhaus et al. used the WoZ approach to simulate semi-autonomous functionality, analyzing user strategies for controlling additional limbs [51]. This methodology has also been applied to such contexts as automated vehicles [28], offering insights into user acceptance and trust. This approach is particularly valuable for SRL research, allowing researchers to investigate user behavior and interaction patterns in controlled environments. By leveraging WoZ setups, researchers can collect rich data that informs the design of SRLs and other advanced robotic systems, bridging the gap between early prototypes and user-centered design.

Our research builds on these insights by leveraging the WoZ methodology to simulate varying levels of autonomy in VSLs in a VR environment. This approach enables us to systematically study user behavior, control strategies, and embodiment under controlled conditions, mitigating the technical constraints of fully autonomous systems. By focusing on VR, we extend the applicability of WoZ in SRL research, providing foundational insights into user interaction patterns while acknowledging the need for future validation in physical settings.

3 DESIGN AND IMPLEMENTATION

This section outlines the design considerations and implementation details of our Wizard-of-Oz setup for VSLs in a VR.

3.1 Design Considerations

In HCI, shared manipulation in virtual environments has been explored through frameworks like the Level of Manipulation (LoM)[5]. In our VR study, we implemented "distributed control across different LoM to form a cohesive entity" [73] using a four-arm setup, where two VSLs were controlled by the participant and two by the human operator. This setup allowed both to independently control separate LoM, collectively contributing to the operation of a single virtual entity for coordinated interaction.

The control mechanism in our study mapped the user-controlled avatar's limbs and the VSLs to separate controllers, with both the participant and operator wearing VR headsets. The participant directly controlled the avatar's limbs, while the operator managed its VSLs. Non-arm-related actions, such as walking, were excluded to maintain focus on arm manipulation.

To address the concerns of over-reliance on human operators in Wizard-of-Oz setups [26], we introduced *random error injections* and *delayed responses* in the operator-controlled limbs to simulate the unpredictability of imperfect autonomous systems [61]. These errors were designed to mimic real-world limitations in autonomous control, requiring the participant to adapt to system imperfections. Additionally, during critical tasks, the operator introduced *temporary loss of control* in certain limbs to simulate failures in the system, forcing the participant to find alternative strategies, thus enhancing the realism and complexity of the simulated autonomy.

Since both the participant and operator were responsible for controlling a total of four virtual limbs for the same avatar, the question of how virtual vision will be shared in a collaborative setting is an important consideration in the system design. To facilitate seamless collaboration, we synchronized the visual experience within the virtual environment. Both the participant and operator views were anchored to the same avatar, ensuring a unified first-person perspective. The participant's head movements controlled the viewpoint, minimizing motion sickness and maintaining alignment between the two users' actions.

The participant's Unity interface was shared with the operator via, allowing the operator to monitor actions in real-time. The operator could also view the participant's perspective directly through the VR headset, ensuring effective coordinated interaction within the virtual environment.

3.2 Implementation

We used the Unity platform to implement our study. This setup enabled both the participant and a human operator to share a single virtual environment, with their real-time interactions synchronously reflected on the same avatar. Our setup was based on two Oculus Quest 3 head-mounted displays and controllers. The application was developed using Unity version 2022.3.7f1 and maintained a stable frame rate of 90 Hz throughout the experiment. Each pair of devices was connected to a dedicated computer via USB, and both computers were physically connected to the same local network to minimize latency [20, 37]. This allowed us to reduce the delay between the participants actions to 5ms. In the virtual environment, the participant and the human operator were simultaneously tethered to a single anthropomorphic avatar.

To avoid biases, the participant and the human operator were positioned two meters apart in the same room but separated by a physical partition, ensuring the participant was unaware of the operator's presence, as shown in Figure 1. This setup allowed the operator to clearly hear the participant's instructions while remaining undetected, thereby maintaining the integrity of the Wizard-of-Oz methodology. To enhance the realism of autonomous interaction, we implemented *blinded control conditions*, where the operator received incomplete task information, simulating limited system awareness. This forced the participant to adjust for the system's imperfect decision-making. Combined with *error injection and control delays*, this approach captured user behavior in scenarios that mimicked real-world autonomous system challenges. We selected the "Photon Unity Networking" (Photon PUN) framework to implement the task environment due to its robust capabilities in supporting multiplayer interactions. Photon PUN's integration with Photon Cloud, a software-as-a-service (SaaS) solution, facilitated the development of our multiplayer environment. Additionally, we utilized an on-premises Photon Server to host the task locally, which further minimized network latency and ensured real-time responsiveness between the participant's actions and the system's feedback.

4 EXPERIMENTAL METHOD

The aim of this study is to explore interaction strategies for controlling multiple semi-autonomous VSLs in non-fixed mapping VR scenarios. To ensure the tasks reflect real-world use cases for supernumerary limbs, we designed two VR tasks representing distinct categories of interaction: targeted reaching named *Basic Control Task* and object manipulation named *Factory Task.*

4.1 Experimental Conditions and Procedure

4.1.1 Participants. We invited 14 participants with prior VR experience (6 female, 8 male; M = 25.6 yrs; SD = 7.9 yrs; no color vision deficiencies) to participate in the study. This sample size is consistent with standards in HCI research [10, 51]. Participants with VR experience were selected to avoid distractions from learning controls [76]. 7 out of 14 participants used VR regularly in professional settings, while the other 7 used it for entertainment or gaming.

4.1.2 Experimental Conditions. We employed a within-subjects experimental design. All participants completed the experiment in both autonomy conditions — *low-autonomy level* and *high-autonomy level.* These two autonomy levels, as theoretically outlined by Kim et al. [42], provide the understanding of how different degrees of robot independence influence human-robot interaction. Detailed descriptions of autonomy levels, including participant instructions, system capabilities, and operational guidelines, are provided in Appendix A. The autonomy settings, which encompass *description*, *capabilities*, and *role*, were designed following the practical framework proposed by Lima et al. [15]. To minimize any potential fatigue or learning effects, we counterbalanced the order of conditions presented to the participants with a break between each task and condition.

To formulate the differentiation of the two autonomy levels, we classified plausible instructions for VSL control into several categories informed by principles from robot learning and multi-agent coordination frameworks [4, 22, 38, 90]. They start with two high-level categories: ***action instructions*** and ***meta-action instructions***. *Action instructions* are instructions issued by the participants that specify how the VSLs should operate or proceed. These include ***Command Instructions***, which involve direct physical actions, such as pressing or rotating objects [38]; ***Demonstration Instructions***, involving participant-led imitation, derive from the robot learning from demonstration (LfD) paradigm, where robots replicate observed behaviors [4]; and ***Delegation Instructions***, which issue abstract commands for autonomous planning, using conditional generation and reinforcement learning to enable systems to infer strategies for multi-step tasks [90]. *Meta-instructions*, on the other hand, do not specify a course of action but are used to simplify or speed up communication between the operator and the VSLs. For example, these include ***Labeling Instructions***, which allow participants to use properties like color or size to specify or indicate objects, hence simplifying naming conventions for future action commands, which is important for contextual recognition in robotic task sequencing [22]. We used this classification to define the two autonomy levels:

Low-autonomy level: At this level, the VSLs require explicit, step-by-step guidance from participants, responding to instructions categorized as ***Command***, ***Labeling***, or ***Demonstration instructions***. Each command triggers a single action from the operator (controlling the VSLs) without independent planning or task abstraction. For example, a ***Command instruction*** like "Press here" results in simple, discrete movements reflecting immediate task execution [38], i.e., pressing a button. A ***Labeling instruction***, e.g., "This is a red ball," involves identifying and categorizing objects based on specified properties, consistent with robotic object classification frameworks [22]. A ***Demonstration instruction***, such as rotating an object by 90 degrees, requires the operator to replicate the participant's real-time actions, exemplifying robot learning from demonstration [4]. This level limits the VSLs to reactive behaviors, entirely dependent on participant guidance.

High-autonomy level: At this level, the VSLs demonstrate advanced capabilities, including object recognition, multi-step task execution, and autonomous planning. The key difference in this level is the inclusion of ***Delegation instructions***, such as "Sort all blocks by colour" prompts the operator to autonomously identify object properties, develop sorting plans, and execute multi-step sequences, reflecting decentralized task planning strategies [90]. By incorporating planning and abstraction, this level provides enhanced efficiency and flexibility for managing complex tasks.

4.1.3 Procedure. The study procedure began by collecting demographic information, including participants' age, gender, VR proficiency, prior VR gaming experience, and any color vision deficiencies. Participants were then briefed on the experiment's purpose, and written informed consent was obtained. Participants were explicitly informed that the VSLs responded to specific verbal commands and demonstrations using hand gestures.

Next, we provided participants with detailed descriptions of the two autonomy levels (see Appendix A) and verbally explained each component, covering the *description*, *capabilities*, and *role*. To ensure consistency, the operator's actions followed predefined rules (see Appendix B), developed during pilot studies to standardize responses across participants. In the low-autonomy condition, the operator performed tasks step-by-step based on explicit participant instructions, whereas in the high-autonomy condition, tasks were executed as single actions based on participant commands.

Participants were then assisted with putting on the VR headsets and provided with controllers. Before beginning the main tasks, they engaged in a 5-minute warm-up exercise based on the relaxed think-aloud protocol (RTA) [29, 30].

This exercise encouraged them to verbalize their thought processes, enabling us to gather richer qualitative data on their decision-making and interaction strategies. Throughout the experiment, participants were encouraged to continue thinking aloud, describing their actions, reasoning, and adjustments during the tasks.

The experimental tasks followed a sequence of *Basic Control Task* → *Factory Task*, conducted once at the *low-autonomy level* and again at the *high-autonomy level*. Each task required the use of both avatar's limbs (controlled by the participant and VSLs, and participants described their strategies and reasoning throughout the tasks. To prevent potential biases, the order of autonomy levels was counterbalanced.

After each autonomy level, participants removed their VR headsets and completed an embodiment questionnaire. Following the *Factory Task*, we conducted a semi-structured interview focusing on their experience of managing the VSLs, their perceptions of task performance under different levels of system autonomy, and their subjective awareness of embodiment effects. The entire experiment, including the briefing, training, data collection, and interviews, took approximately 50 minutes. No monetary compensation was offered to participants for their involvement.

4.2 Experimental Tasks

The tasks were designed to explore participants' control strategies to employ the VSLs in VR, requiring coordination, precision, and timing. They required participants to simultaneously and collaboratively use the avatar's limbs and VSLs.

4.2.1 Basic Control Task. The first task in our study (Figure 2(a)) is a typical "reaching" task, commonly used in VR object manipulation [69]. Participants completed basic coordination through the action of three-dimensional reaching out to a button, a fundamental skill in remote robotics manipulation [31]. The reaching task, chosen as the first task for its straightforward objective, singular focus, and absence of distracting elements. It provides a controlled environment to examine basic multiple-limb coordination. As demonstrated by Guterstam et al. [79], reaching tasks are crucial for exploring how users integrate the VSLs into their body schema.

The task included four stages, with an increasing number of lit buttons at each stage. In stage 1, only one button was lit; Vstage 2 had two; and progressively stage 4 had four buttons. The participants were asked to use both their avatar's limbs and SLs to press lit buttons to turn them off. The task was considered complete once all lit buttons were pressed in each stage, with no repeated attempts or iterations within the stages. Performance metrics such as task completion time and error rate were recorded to assess how effectively participants managed the tasks across different stages. Participants had the freedom to choose which limb to use for each button, but in the later stages, with three and four buttons lit, participants needed to engage both their avatar's limbs and VSLs simultaneously to complete the task.

4.2.2 Factory Task. The second task is a fast-paced robotic-control task (Figure 2(b)) involving grasping, rotating, and reaching [49]. It builds on the earlier reaching task by adding precision, timing, and coordination requirements. To ensure the task reflected natural behavior, it requires bilateral coordination, involving both limbs for

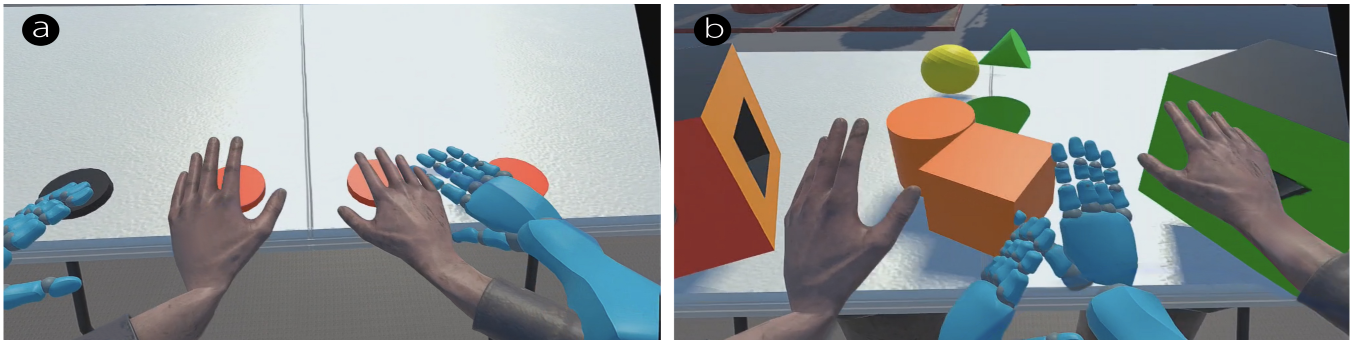

Figure 2: Two tasks used in the user study: the two avatar's limbs were controlled by the participants, while the VSLs were controlled by a human operator. (a) Basic Control Task: Participants used the avatar's limbs and VSLs to simultaneously actuate illuminated buttons in a sequence. (b) Factory Task: One hand grasps a falling target, while the other rotates a Shape Sorter to match and insert the target into the correct hole based on shape and color.

simultaneous or alternating movements. Specifically, we focused on asymmetrical bilateral coordination, where each limb performs a distinct action, such as one hand grasping while the other rotates or positions an object [27]. Our task design aligns with key insights from Zhang et al. [89], who emphasize the role of SRLs in high-pressure, multitasking environments such as industrial settings. This task was chosen as the second for incorporating these elements, reflecting real-world applications of SRLs in enhancing human performance in repetitive, demanding tasks.

The task mechanism involved a *Shape Sorter*, a cubic object with four sides, each containing differently shaped holes. Two Shape Sorters were placed on the table, with the same set of shapes on each, but arranged with different color associations and orientations. While the shapes on both sorters were identical, the colors corresponding to each shape varied between the two sorters. For example, a triangular hole might be on the green side of one sorter and the yellow side of the other.

The task structure featured ten objects, with matching shapes and colours of the *Shape Sorter* faces. The objects were dropped one by one on the table every three seconds (Fig. 2). The participants were asked to insert the objects to the shape sorter matching the shape (e.g., such as triangles, cylinders, and cones) and color. Participants needed to grab the falling target, rotate the Shape Sorter to align both the shape and the color, and insert the target into the corresponding hole. This design tested participants' ability to coordinate the avatar's limbs and VSLs simultaneously, managing the tasks of grasping, rotating, and inserting the target under time pressure. Participant's earned points for each correct insertion, and task completion time and accuracy measured performance.

4.3 Experimental Setup

We conducted the experiment in a 2.5 x 3.5 meter laboratory room, which was divided into two sections by a partition to ensure that the participant was unaware of the human operator's presence, maintaining the integrity of the Wizard-of-Oz setup. A GoPro camera recorded the participant's activities for later analysis. The participant wore a VR headset and provided voice commands, which the operator could hear, but the operator remained silent and did not communicate directly with the participant during the experiment.

4.4 Experimental Data

4.4.1 Qualitative data. We audio and video recorded the think-aloud session and the answers to the interviews as qualitative data.

4.4.2 Performance. We measured the participants' performance using two variables: task completion time and error rate. Specifically, for the *Basic Control Task*, the task completion time was calculated from the moment the button illuminated until the participant successfully actuated the button. For the *Factory Task*, performance was assessed by measuring the time taken for each action, from grasping the target to placing it into the correct hole, along with the number of accurate target insertions. To avoid biases, participants were not shown their scores during the experiment.

4.4.3 Embodiment Experience. As a measure of the degree of shared virtual body experience, we utilized the Avatar Embodiment Questionnaire (AEQ) to assess evaluation of embodiment [25]. The participants completed the questionnaire at the end of tasks associated with each level of system autonomy (i.e., after completing the low-autonomy level, and again after completing the high-autonomy level). This approach mirrors the two critical stages identified in [3], which is pre- and post-learning stages. Participants removed their VR headsets to complete the questionnaire.

5 RESULTS

In this section, we present both quantitative and qualitative findings, focusing on embodiment, control and task management strategies against different levels of autonomy of multiple VSLs.

5.1 Quantitative Findings

Figure 3 outlines our quantitative findings in terms of task completion time, error rate and participants' response to the Avatar Embodiment Questionnaire (AEQ). Since the data did not follow a normal distribution, we used the Aligned Rank Transform (ART) test to perform non-parametric comparisons.

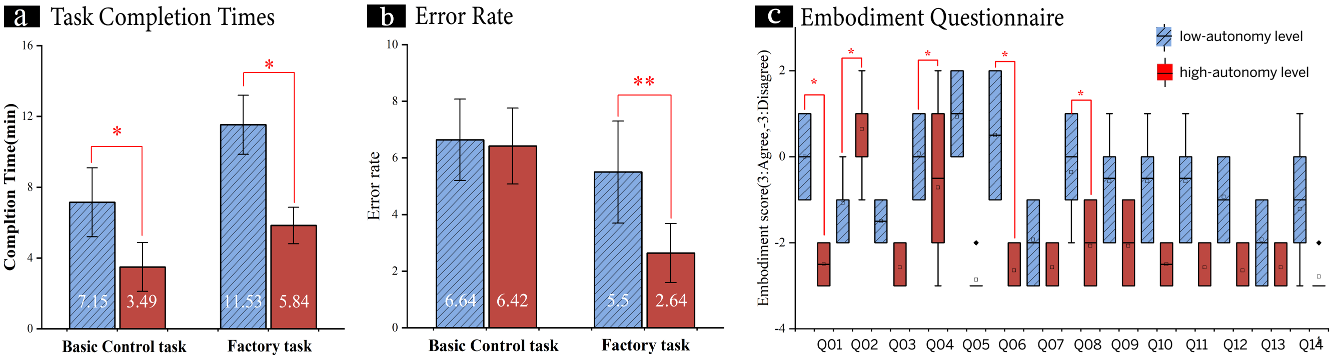

Figure 3: The averages for *low-autonomy* level and *high-autonomy* levels in (a) the task completion time in the basic control task and factory task; (b) the error rates in the basic control task and factory task; (c) the embodiment questionnaire. * indicates p < .05 and ** indicates p < .01.

As shown in Figure 3(a), participants on average took significantly longer time to complete the tasks in *low-autonomy level* compared to *high-autonomy level* in both *Basic Control task* ($z = -1.05, p < 0.05$) and *Factory task* ($z = -2.09, p < 0.05$). For the error rate, we did not observe a statistically significant difference between the autonomy levels in the *Basic Control task* ($z = -0.75,\ p = 0.44$). However, as shown in Figure 3(b), in the *Factory task*, error rate in the *low-autonomy level* was significantly higher compared to that of *high-autonomy level* ($z = -2.93,\ p < 0.01$)

As shown in the Figure 3(c), results from the embodiment questionnaire administered after each autonomy level revealed statistically significant differences (using Wilcoxon–Mann–Whitney test) to several questions. Statistically significant differences were observed in several key questions: Specifically, Q1: "I felt as if the virtual extra limbs/arms were my limbs/arms" ($p < 0.05$); Q2: "It felt as if the virtual extra arms/limbs I saw were someone else's" ($p < 0.05$); Q3, "It seemed as if I might have more than two limbs/arms" ($p < 0.05$); Q4, "It felt like I could control the virtual extra arms as if they were my own arms" ($p < 0.05$); Q6, "I felt as if the movements of the virtual extra arm were influencing my own movements" ($p < 0.05$); Q8, "I felt as if my arms were located where I saw the virtual extra arms" ($p < 0.05$). As a trend, the participants tend to feel a higher sense of body ownership (Q1, 2), a sense of agency (Q4, 6), and a sense of self-location (Q8) after *low-autonomy level* compared to that of *high-autonomy level.*

To validate the reliability of our quantitative measures, we analyzed completion times for different tasks and how they relate to instruction types to ensure their actions were consistent and adhered to guidelines (Figure 4). Since the data conforms to a normal distribution, we employed a One-Way ANOVA to compare the operator's average reaction time for instructions.

Across the study, participants issued a total of 672 commands. Specifically, ***Command*** - 192 times, ***Labeling*** - 160, ***Demonstration*** - 176, and ***Delegation*** - 144 times.

For ***Command Instructions***, shorter commands (1–3 seconds) resulted in an average operator reaction time of $M = 2.50$ ($SD = 0.31$) seconds, while longer commands (4–6 seconds) required significantly more time at $M = 4.53$ ($SD = 0.40$) seconds ($z = -3.21, p < 0.01$). Similarly, for ***Labeling Instructions***, shorter commands (1–2 seconds) corresponded to a reaction time of $M = 2.02$ ($SD = 0.23$) seconds, whereas longer commands (3–5 seconds) took $M = 3.84$ ($SD = 0.31$) seconds. For more complex ***Demonstration Instructions***, the operator's reaction time also scaled proportionally with command length. Commands lasting 3–4 seconds elicited an average reaction time of $M = 4.52$ ($SD = 0.39$), while those lasting 6–8 seconds increased significantly to $M = 8.03$ ($SD = 0.48$) seconds ($z = -3.52, p < 0.001$). Lastly, ***Delegation Instructions*** exhibited the longest reaction times, with shorter commands (2–3 seconds) averaging $M = 3.46$ ($SD = 0.30$) seconds and longer commands (6–8 seconds) reaching $M = 9.04$ ($SD = 0.61$) seconds.

Notably, as shown in Figure 4, ***Command***, ***Demonstration***, and ***Delegation Instructions*** were consistently followed by an operator reaction, with reaction times roughly aligning with the length of the commands. In contrast, ***Labeling Instructions*** were followed by shorter reaction times since no physical movement was required. Additionally, under Low Autonomy, ***Delegation Instructions*** did not trigger any reaction as this feature was unavailable, while in high Autonomy, ***Delegation Instructions*** consistently resulted in reactions of equivalent durations.

5.2 Qualitative Findings

We conducted an inductive thematic analysis [6] to explore how participants controlled the VSLs and how their experiences varied with different autonomy levels. This analysis aimed to understand participants' interaction strategies, focusing on how they balanced manual and autonomous control, as well as the challenges of managing multiple VSLs. The analysis was conducted by two independent coders, ensuring diverse perspectives and reducing individual bias. To enhance consistency, inter-rater reliability was measured using Cohen's Kappa ($\kappa = 0.74$)..

Following Braun and Clarke's framework [7], we began by familiarizing ourselves with the data through a review of all interview transcripts to identify key elements of participants' feedback. In the initial coding phase, we assigned descriptive labels to relevant passages, focusing on body ownership, agency, task delegation, labeling strategies, and managing multiple tasks. Independent coding was followed by collaborative discussions to refine and consolidate

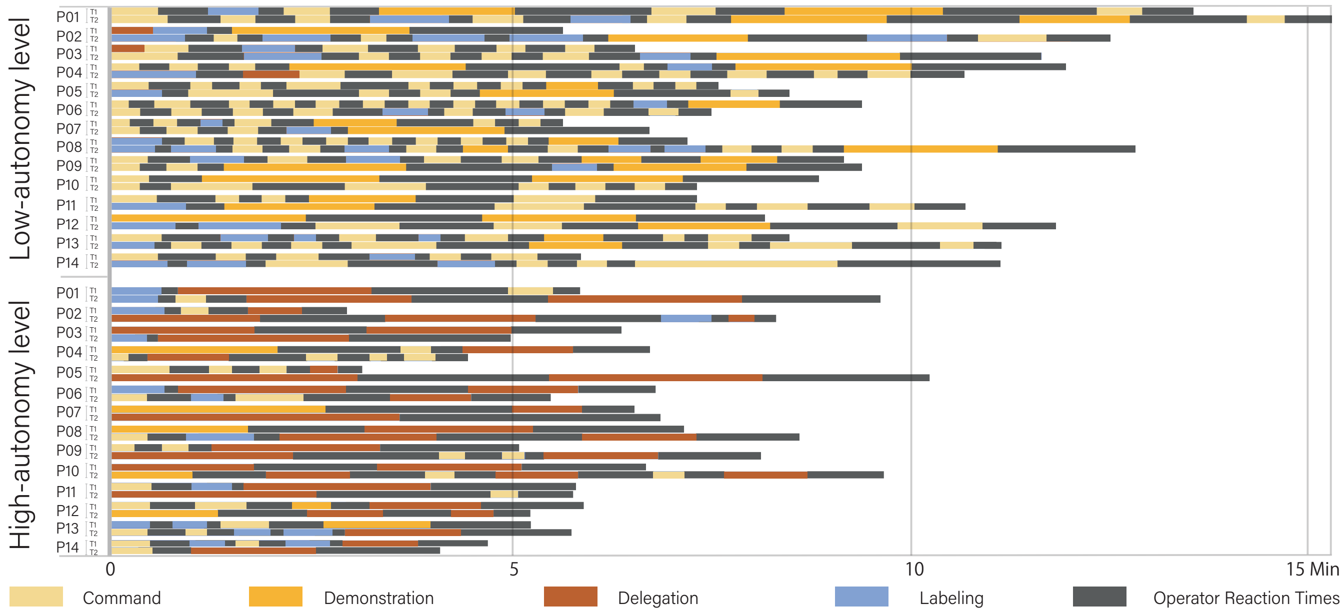

Figure 4: The participant-issued instructions, including *Command Instructions, Demonstration Instructions, Delegation Instructions* and *Labeling Instructions*, along with the WoZ operator's reaction times to these four instruction types. A single colored segment may represent multiple instructions of the same type. It also compares the reaction times across two distinct autonomy levels for both the Basic Control Task and the Factory Task.

codes into initial categories reflecting participants' strategies for handling task complexity and adapting control methods.

We refined and grouped the codes into higher-order themes as patterns emerged, such as "Body Ownership, Agency, and Control Embodiment" and "Control Strategies and Their Perceived Performance". These themes helped us build a framework for understanding participants' experiences in adapting to multitasking demands with VSLs. Our iterative coding approach consistently identified key themes such as "Balancing Cognitive Load and Task Complexity" and "Trust, Ownership, and Control Embodiment" across participant data. We followed reflexive thematic analysis principles [7], allowing ongoing reflection and refinement to shape the analysis.

The following sections outlines the key themes that emerged from participants' experiences, providing insights into how they adapted their control strategies and perceptions across different tasks and levels of autonomy.

5.2.1 ***"It felt weird, like they were someone else's arms" — Body Ownership, Agency, and Control Embodiment.*** Having two additional limbs was an unusual experience for participants, with many (N = 8) initially describing a sense of disconnection. As P11 remarked, *"It felt weird, like they were someone else's arms"*, especially when the VSLs didn't fully align with their intended movements. However, participants were surprisingly able to develop a sense of autonomy over time: *"At first, they felt awkward, but the more I used them, the more natural"*, noted P08.

Sense of Body Ownership: 8 out of 14 participants felt a stronger connection to the VSLs when they had direct control, particularly during low-autonomy level. *"When I was guiding the arm, it made me feel more real"*, P02 said. This manual control reinforced their sense of ownership, particularly in VR settings, where repeated interactions allowed participants to adapt to the virtual embodiment of the arms, rather than perceiving them as physical robotic systems. *"The more I used it, the more natural it felt"*, explained P05.

In contrast, 9 out of 14 participants found that as the system's autonomy increased, their sense of ownership diminished. *"When the arms acted on their own, it was more like watching but not controlling"*, P05 said. Although the system efficiently performed tasks, the lack of direct involvement left participants feeling disconnected, which "didn't feel like my arm anymore," noted P07, and this is also mirrored in our quantitative findings in Section 5.1.

Sense of Agency: Participants (10 out of 14) reported that direct control enhanced their sense of agency, providing a feeling of mastery, especially during precision tasks. *"I felt more control when I moved the arms"*, said P03. Conversely, increased autonomy reduced participants' sense of being in control, with 8 out of 14 participants reporting a shift to an observer role. *"When the arms worked, I felt like more of an observe"*, shared P02. The quantitative findings (Section 5.1) also mirrored this trend. While autonomy allowed efficiency, it diminished participants' perception of control.

Sense of Self-Location: 9 out of 14 participants experienced a sense of self-location when the VSLs closely synchronized with their real bodies. *"I didn't really feel like they were extra or separate"*, said P07. However, when the system's movements failed to meet expectations, it disrupted this connection. *"If the arms didn't follow my action in time, I felt disconnected"*, noted P02. 7 out of 14 participants, managing multiple VSLs simultaneously was challenging for their spatial awareness, particularly during complex tasks.

Addressing Research Question: These findings directly address RQ3, revealing how participants' perceptions of embodiment—specifically body ownership, agency, and self-location—shifted with varying levels of autonomy. The stronger sense of ownership and agency in low-autonomy conditions contrasted with the diminished embodiment in high-autonomy settings, highlighting the critical role of direct control in shaping user experience.

5.2.2 *"I felt like I was juggling!" — Control Strategies and Their Perceived Performance*.

Controlling a single limb in virtual reality allows participants to intuitively manage actions. However, controlling multiple VSLs highlights how often coordination is required and how difficult it is to maintain. 11 out of 14 participants in our study highlighted how managing several limbs at once required them to constantly shift focus and attention, making it difficult to allow smooth coordination between VSLs. *"It felt like juggling tasks"*, P05 remarked, emphasizing the complexity of maintaining precise, synchronized movements when operating multiple VSLs simultaneously.

Direct Control and Autonomous Repetition: 8 out of 14 participants used various strategies to delegate tasks to the VSLs. Some (N=13) preferred for direct control, where the VSLs mirrored their movements, especially at *low autonomy* level. *"Left arm, mirror my movements"*, P11 said, who employed a specific strategy to teach the VSLs to follow their avatar's limbs' actions. P03 described their approach, saying, *"I choose to do a simple action, and ask it to learn and follow"*. *"It is easier"*, they added, *"It [is] like another arm that did learn what I was doing, but I need to look at two arms"*. 6 out of 14 participants named the action before demonstrating it. For example, P08 said *"This is rotate, rotate this way"* to teach the VSLs. This allowed for real-time adjustments, but it became mentally exhausting when managing multiple VSLs. *"It required so much more effort and guidance, which was tiring"*, noted P02.

6 out of 14 participants opted to teaching the VSLs a set of actions to repeat autonomously. *"Once I showed the arm how to do something, I didn't have to guide it anymore—it just kept going"*, P05 said. While this method was efficient, it sometimes required several tries to ensure correct execution. P06 clearly described this idea: *"I wanted to demonstrate a movement and let left arm learn it. I taught it a simple action, then keep it do it again and again. Just kept going until I say stop"*. In contrast, at higher levels of autonomy, all participants opted for more direct command strategies rather than demonstrations to control the VSLs. As P04 issued the command, *"Robotic [virtual supernumerary] Arm, press the red button"*, the system executed the task autonomously. Instead of showing the VSLs how to perform an action, 8 out of 14 participants simply gave verbal commands and let the system execute the task autonomously.

Organizing Strategies for Efficiency: Pre-planning and labelling tasks by object properties played a crucial role in improving task management for 13 out of 14 participants. 10 out of 14 participants in the factory task tried labeling objects by color or shape to improve task management and streamline interactions with the VSLs. P06 said, *"Robotic [virtual supernumerary] arms, this is light [Illuminated button], always press the light one [Illuminated button]"*, this participant explained it by adding, *"I [labelled] the target how I felt like, and yeah, it's red, but I'm ... used to calling stuff that lights up a 'light button,' so that's what I said"*. However, 4 out of 14 participants felt that too much pre-planning disrupted their natural workflow. *"I spent more time trying to label thing[s]"*, explained P09, indicating that excessive organization could slow down task execution.

In addition to labeling, 5 out of 14 participants categorized objects by size or function, which enhanced coordination. *Once everything was sorted, I didn't have to keep thinking about it"* (P08). This strategy reduced decision-making during the task, but 4 out of 14 participants found it challenging to balance planning with keeping their momentum. One participant remarked, *[Planning] helped, but I sometimes lost track of the main task when I was too focused on organizing,"* said P11.

At the *high-autonomy* level, 6 out of 14 participants chose not to label targets themselves. Instead, they followed conventional rules when issuing commands, rather than assigning custom names to the targets. *"I could say something like 'rotate the box 180 degrees' or 'press the red button', and they'd just do it"*, noted P04.

Managing Task Control Approaches: Sequential handling allowed for better control but took more time, especially at *low-autonomy* levels. *"Doing one thing at a time... , it gave me more control, but it just slowed me down"*, noted P08. In contrast, multitasking offered speed but led to more errors. *"When I tried doing both at once, it felt faster, but..., I made way more mistakes"* (P01). At *high-autonomy* level, 7 out of 14 participants demonstrated different strategies for managing control between themselves and the VSLs. For high-priority, delicate tasks, 4 out of 14 participants preferred keeping full control to ensure accuracy. *"I just liked handling those tasks myself, the ones that needed constant adjusting"* said P03. Others found it more efficient to delegate even complex tasks to the autonomous VSLs, especially when juggling multiple tasks. *"I'd let the arms handle the boring stuff, and even some of the tricky bits, just so I could focus on hard thing, like find the shape"*, noted P05.

Switching between sequential and simultaneous actions was tough for 5 out of 14 participants, while 6 out of 14 participants adapted by letting the VSLs handle repetitive tasks. 6 out of 14 participants often switched control modes based on task complexity. *"For the simpler stuff, I'd just ... teach it to take over... but when it was more detailed, I'd switch back to doing it manually"*, explained P09.

Addressing Research Question: These observations answer RQ1 by demonstrating how users dynamically shifted between teaching-by-demonstration, labeling, and delegation strategies based on specific task demands. Furthermore, the reduced error rates and significantly improved efficiency in the high-autonomy level (Sec. 5.2.1) address RQ2, showing that higher autonomy enhanced task management for complex scenarios but required users to adapt their coordination approaches.

5.2.3 *Summary*.

Our study examined participant interactions with semi-autonomous VSLs across two autonomy levels. Qualitative results showed that participants preferred direct control in *low-autonomy* level, which strengthened their sense of body ownership, agency, and self-location. At this level, participants used strategies like having the VSLs follow their actions, repeat simple tasks, or even labeling objects themselves to improve task management. In *high-autonomy* level, participants found it more efficient to delegate tasks to the system, relying on verbal commands and pre-defined rules rather than manual control or custom

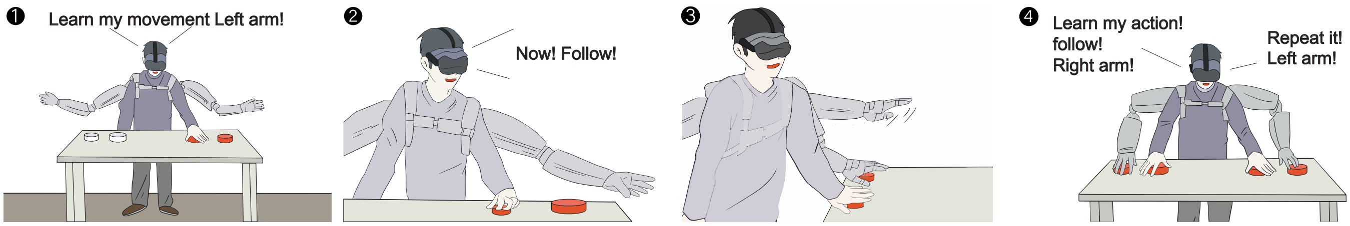

Figure 5: The flow of the follow and repeat strategies performed by participants.

labeling. However, this shift led to a diminished sense of control and connection to the limbs.

6 DISCUSSION

Building on our findings, we have identified key implications for designing and improving autonomous VSLs in VR. This section explores how users adapt control strategies, manage task complexity, and experience embodiment when coordinating VSLs. Additionally, we discuss challenges and opportunities in integrating system autonomy to enhance user experience and system responsiveness in future VR applications.

6.1 Implications for Control strategies and VSLs Control

Our findings reveal that selecting effective control strategies is critical for optimizing the use of VSLs. Participants adapted their strategies based on task complexity and autonomy levels, shifting between verbal commands, teaching by demonstration, and manual control to best suit task demands. While low autonomy settings benefited from precise, step-by-step instructions, higher autonomy allowed for efficient task delegation. However, balancing manual input and autonomy presented challenges, particularly in multitasking scenarios, where task-switching disrupted workflow.

We selected VR as a research platform for simulating and evaluating VSLs interactions due to its capacity to safely and flexibly replicate scenarios that are challenging to achieve with physical robotic systems. Prior research suggests that VR can serve as an effective prototyping environment for understanding body schema extension and control strategies, with potential applications in physical settings [3, 41]. Although VR-based embodiment involves a level of user "suspension of disbelief" [52], it still enables valuable insights into user adaptation to novel body configurations. While our findings focus on VR, they provide foundational insights that may guide future exploration with physical robotic limbs, assuming that VR-to-reality transferability is feasible.

While previous research has explored multimodal interaction and task management [14, 36], further development is needed in the following directions to enable smoother transitions of control.
Verbal Control: Our results demonstrate that continuous verbal control was most effective in *low-autonomy* settings, where participants needed precise, step-by-step instructions to guide the VSLs, aligned with prior research [32]. However, as task complexity or autonomy increased, continuous verbal control became less efficient, reflecting the challenges of maintaining precision when managing VSLs in real-time.

With higher autonomy, participants adopted more efficient strategies such as task delegation or pre-labeled ensemble actions (Figure 6), which reduced reliance on continuous verbal input by allowing the system to execute tasks autonomously after initial instructions. Previous studies on multimodal interaction similarly suggest that as autonomy increases, users prefer intuitive, low-effort methods like task abstraction or labeling [14]. Controlling VSLs under non-fixed mapping further exacerbated challenges, as simultaneous tracking and the lack of consistent mapping made precise verbal input difficult.

Participants often combined labeling with control by example to simplify task management, streamlining control and enabling autonomous handling of multi-step tasks. However, unlike studies employing pre-defined control schemes [12, 13], the reliance on customized labels introduced challenges, particularly the repeated guidance needed when the system failed to learn correctly. To address these issues, future systems should reduce user intervention required for correcting learning failures.
Control Choice as a Function of Task or Autonomy: Participants adapted their control strategies based on task complexity, with simpler, repetitive tasks delegated to the autonomous VSLs and more complex, high-stakes tasks managed manually, as shown in figure 5. Research in task management, such as Rasmussen's Skill-Rule-Knowledge Framework, supports this approach, where simple tasks are categorized as skill-based and easily automated, while complex tasks, which are knowledge-based, require manual control for greater precision [46]. The consistent preference for manual control in critical tasks, even when autonomy was available, reflects the importance of precision and control for participants.

Task-switching, intended to facilitate multitasking, often ended up disrupting the flow of tasks instead. Switching frequently between manual and autonomous control, especially in complex tasks, interrupted their workflow, making it harder to maintain smooth task execution. While autonomy improved efficiency for simpler tasks, more complex tasks required careful oversight. Rather than treating autonomy and manual control as separate modes, future systems should integrate these approaches more seamlessly, allowing for dynamic shifts based on task complexity without disrupting the user's workflow.

6.2 Optimizing Control and Performance in Multiple VSLs Systems

Managing multiple VSLs revealed important insights into the intersection of body ownership, agency, and task efficiency. While

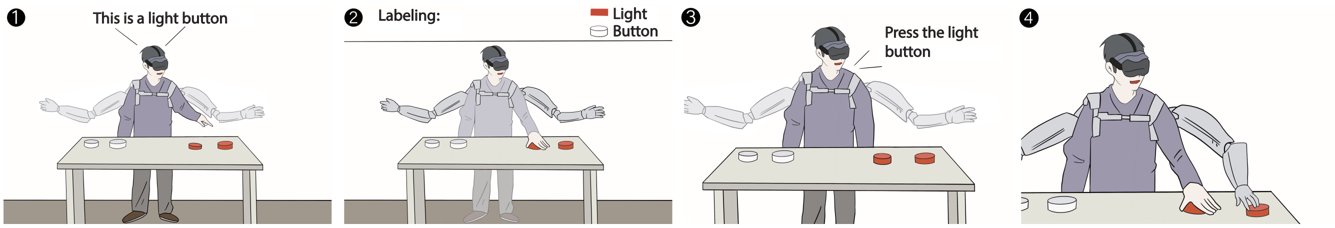

Figure 6: The organizing strategies performed by participants.

participants faced significant challenges in coordination across autonomy levels, the study highlighted opportunities for optimization. Participants noted the potential of multiple VSLs when designed with a balance of manual and autonomous control. However, achieving optimal efficiency and sustained engagement requires careful attention to synchronization, attention allocation, and user feedback mechanisms. This section explores these factors and identifies strategies for improving multiple VSLs systems.

Sense of Body Ownership in Autonomous Multiple VSLs Systems: In *low-autonomy* conditions, participants felt a stronger sense of body ownership, perceiving the VSLs as natural extensions of their body. Coordinated control reinforced this connection, but higher autonomy levels introduced a noticeable disconnect, especially when autonomous actions disrupted limb coordination, as similarly observed in previous research [48, 91]. Participants favored partial autonomy, which allowed manual control of key movements while efficiently delegating routine tasks, balancing control and assistance. Precise visual alignment between the avatar's limbs and VSLs remained crucial for maintaining ownership across all autonomy levels. Future systems should leverage real-time visual feedback to continuously reinforce this connection.

Sense of Agency in Autonomous Multiple VSLs Systems: Higher autonomy shifted participants' roles from direct control to supervision, transforming agency into a strategic interaction—a contrast to the findings of Firlej and Taeihagh [23]. While partial control over one VSL, combined with monitoring the other, helped maintain participants' sense of influence even in *high-autonomy* conditions, agency was not perceived as a simple binary state [54]. Instead, it emerged as a dynamic experience, with participants continually balancing delegation and responsibility. However, unpredictability in the system's responses introduced cautious delegation, which, in turn, affected participants' confidence in the autonomous features.

Attention Management and Task Efficiency in Multiple VSLs Control: Task efficiency depended on how participants allocated cognitive attention [21, 71, 72, 72]. Attention theories, such as Kahneman's model of attention [9], suggest that humans allocate limited mental resources to tasks based on demand and importance. Dividing attention between manual and autonomous tasks optimized performance, allowing participants to focus on complex actions while VSLs handled routine tasks [1]. However, frequent task-switching disrupted flow and reduced efficiency [11]. Sustained attention became increasingly challenging in high-demand tasks, with monitoring autonomous VSLs leading to fatigue. Future designs should prioritize attention management by assigning tasks based on graded attention levels.

Opportunities of Using Multiple VSLs: The potential of multiple VSLs systems spans various fields as supernumerary limbs approach real-world applications [8, 74, 78]. Prior studies have explored augmenting human capabilities through robotic systems in controlled settings like manufacturing and healthcare [83, 87, 88]. Our research builds on this by showing how VR-developed control strategies can guide VSLs implementation in dynamic scenarios. In industries like manufacturing and logistics, delegating repetitive tasks to VSLs allows workers to focus on precision-driven actions [17, 68], reflecting our findings where participants assigned simpler tasks to autonomous VSLs while managing complex ones manually. This approach could enhance productivity by aligning with selective attention needs in high-demand environments. In healthcare, multiple supernumerary limbs systems could support injury recovery, physical therapy, and assistive technologies for individuals with disabilities [45]. Studies highlight adaptability as crucial in rehabilitation technologies [55], and our findings suggest VSLs could further personalize therapy.

6.3 Challenges and Opportunities of Non-Fixed Mapping

The non-fixed mapping of control in multi-arm systems posed both challenges and opportunities. *High-autonomy* conditions reduced mental load by allowing VSLs to handle simpler tasks, but the lack of fixed associations introduced uncertainty. In *low-autonomy* settings, frequent switching and manual coordination increased cognitive strain. Even in *high-autonomy* scenarios, participants hesitated to fully trust the system's autonomous decisions, especially when precision was needed. This section explores the impact of non-fixed mapping on workload and autonomy.

Workload Distribution and Cognitive Strain in Autonomous Non-Fixed Mapping Systems: In *low-autonomy* conditions, participants experienced higher cognitive load due to frequent task switching and manual coordination between the VSLs. The absence of fixed task-to-limb associations required continuous monitoring and role adjustments, increasing mental strain and reducing efficiency. Conversely, *high-autonomy* conditions alleviated much of this burden by delegating simpler tasks to the VSLs, allowing participants to focus on more complex actions, consistent with prior findings [85]. However, participants still had to monitor the VSLs, and the non-fixed mapping introduced uncertainty in balancing manual control and trust in the system.

Trust and Control in Autonomous Multiple VSLs Systems: Managing multiple VSLs in *low-autonomy* settings was challenging, particularly when precision was required, as participants had to

Table 1: This table presents key findings from our study alongside corresponding design suggestions for enhancing multiple VSLs control systems. The table is organized into categories reflecting the main themes discussed in our results and discussion sections. Each row aligns a specific strategy with its main findings, followed by actionable design recommendations aimed at improving task management, user experience, and system efficiency in multiple VSLs VR interactions.

Control strategies and VSLs Control (Section 6.1)		
Strategists	**Main Finding**	**Design Suggestions**
Verbal Control	Frequent switching between VSLs interrupted workflow and added cognitive load.	Enable smoother switching between control modes. Provide flexible task role reassignment mechanisms.
Task Demonstration	Direct mirroring required high user attention and frequent recalibration.	Incorporate hybrid control (manual and automated) systems. Offer real-time feedback on positioning.
Task management	Labeling improved coordination but pre-defined categories reduced flexibility.	Support dynamic object recognition. Allow on-the-fly adjustments to categories and labels.
Control and Performance in Multiple VSLs Systems (Section 6.2)		
Body Ownership	Participants felt disconnected from VSLs during high autonomy and poor alignment with their body perception.	Provide real-time visual feedback. Gradually adjust control autonomy.
Agency	Users preferred manual control in critical tasks, and labeling improved trust in autonomous actions.	Balance autonomy and manual control with user-defined labels. Offer options for customizable task assignment.
Task Efficiency	Synchronizing multiple VSLs was challenging, and interruptions impacted performance.	Introduce visual and auditory cues for better synchronization. Enable dynamic task distribution based on VSLs usage.
Challenges and Opportunities of Non-Fixed Mapping (Section 6.3)		
Workload	Manual control in *low-autonomy* settings increased workload; *high-autonomy* shifted users to a supervisory role.	Provide proactive feedback mechanisms. Let users toggle between manual and autonomous modes.
Autonomy	High autonomy caused uncertainty and reduced user confidence, while manual control was seen as more predictable.	Ensure autonomy is predictable and transparent. Implement clear trust-building mechanisms.

manually coordinate each limb. In *high-autonomy* settings, trust became a key issue as participants were hesitant to rely fully on the system's autonomous decision-making when control switched between VSLs without direct input. While autonomy reduced the manual burden, unpredictable decisions led to reduced confidence. As emphasized by Hancock et al. [85], mismatches between user trust and system reliability can result in either overtrust or distrust. Future systems should offer more transparent and predictable autonomous decisions to address these challenges in non-fixed mapping scenarios.

7 Limitations and Future Work

Our study has several limitations that should be noted to inform future research. In this study, the autonomy of the VSLs was simulated using a Wizard-of-Oz approach, where the system's autonomous behavior was covertly controlled. Although participants believed they were interacting with fully autonomous VSLs, the system's responses were manually guided. This allowed us to explore a range of interaction scenarios, but the simulation may not fully capture how users would behave with truly autonomous systems. Specifically, participants were limited to interacting with the VSLs through mediated commands without the ability for direct physical interaction. The absence of real-time system errors or unexpected actions typically encountered with fully autonomous VSLs could influence the user's experience. As a result, these findings may not entirely generalize to real-world autonomous systems. Future studies with actual autonomous systems are needed to further validate how users would engage with and trust such technologies in uncontrolled environments.

While VR served as an accessible and controlled platform for studying supernumerary limb interactions, we recognize that our findings are still speculative regarding their application to physical robotic systems. Although past studies indicate potential transferability from VR to real-world settings [3, 41], our findings require further validation with physical prototypes before they can be confidently applied to actual supernumerary robotic limbs. Consequently, while this study provides initial insights, future work is necessary to comprehensively assess the relevance of these insights in real-world contexts.

We focused on two distinct levels of autonomy to examine how users interact with the VSLs, offering valuable insights into basic control strategies. However, this limited scope may not capture the full range of system autonomy or task complexity that could arise in more dynamic, real-world scenarios. Future work should explore a broader spectrum of autonomy levels and task complexities to better understand how users adapt to varying system autonomy. Similarly, while our tasks were designed to isolate specific coordination and control challenges, they may not represent the full variety of real-world applications for supernumerary limbs. Expanding the scope of tasks in future studies will help validate findings across a wider range of use cases.

While the sample size meets the typical standards for exploratory studies in Human-Computer Interaction (HCI) to gather qualitative insights and identify key interaction patterns [10], it may still limit the generalizability of the findings. Similar sample sizes have been used in comparable exploratory studies, such as those by Muehlhaus et al. [51], to effectively investigate interaction strategies. The study's small sample size, while suitable for qualitative exploration, limits the generalizability of the quantitative findings. The quantitative metrics, such as task completion time and error rate, should be interpreted as exploratory trends rather than definitive results. Furthermore, certain user diversity factors, such as color blindness or other visual impairments, were not considered. Future studies should incorporate a larger, more diverse participant pool to address potential accessibility challenges and ensure the system's inclusivity.

8 CONCLUSION

In this work, we explored how varying levels of autonomy in VSLs affected user interaction, control strategies, and embodiment in a VR environment. Through two task-based VR activities— the *Basic Control Task* and the *Factory Task*—we examined how users adapted to low and high autonomy using the VSLs across two tasks. Our findings demonstrate that participants preferred assigning repetitive or lower-priority tasks to the VSLs while maintaining direct control over more critical actions. Additionally, participants reported a stronger embodiment and control with a *low-autonomy* level, while a *high-autonomy* level allowed for greater task efficiency but a weakened embodiment of the VSLs. We have discussed the key implications of our themes on future research and design. In addition, we have provided a discussion of use cases for control strategies and a table of concrete suggestions. Our findings are based on VR-based simulations of VSLs and provide insights into virtual embodiments. While these insights offer potential directions for real-world robotic systems, further research is needed to validate these results in physical environments. This understanding is crucial for designing autonomous systems that enhance user experience, improve task performance, and maintain a strong embodiment, with potential applications in gaming, rehabilitation, and assistive technologies.

Acknowledgments

This project was supported by the Australian Research Council Discovery Early Career Award (DECRA) - DE200100479. Dr. Anusha Withana is the recipient of a DECRA fellowship funded by the Australian Government. We also acknowledge financial support from the University of Sydney's Digital Sciences Initiative through the DSI Research Pilot Project grant scheme. Furthermore, we are grateful for the support provided by the Neurodisability Assist Trust and Cerebral Palsy Alliance, Australia - PRG04219. Dr. Andrea Bianchi was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2024-00337803). We extend our gratitude to all user study participants for their valuable time and contributions. Special thanks to Dr. Brandon Syiem for his invaluable support and insightful comments. Additionally, we appreciate the members of the AID-LAB for assisting us in various ways.

References

[1] Lubna Ahmed and Jan W De Fockert. 2012. Focusing on attention: The effects of working memory capacity and load on selective attention. (2012).

[2] Günter Alce, Mattias Wallergård, and Klas Hermodsson. 2015. WozARd: a wizard of Oz method for wearable augmented reality interaction—a pilot study. *Advances in Human-Computer Interaction* 2015, 1 (2015), 271231.

[3] Ken Arai, Hiroto Saito, Masaaki Fukuoka, Sachiyo Ueda, Maki Sugimoto, Michiteru Kitazaki, and Masahiko Inami. 2022. Embodiment of supernumerary robotic limbs in virtual reality. *Scientific reports* 12, 1 (2022), 9769.

[4] Brenna D Argall, Sonia Chernova, Manuela Veloso, and Brett Browning. 2009. A survey of robot learning from demonstration. *Robotics and autonomous systems* 57, 5 (2009), 469–483.

[5] Peter Bayliss. 2007. Beings in the Game-World: Characters, Avatars, and Players. In *Proceedings of the 4th Australasian Conference on Interactive Entertainment* (Melbourne, Australia) *(IE '07)*. RMIT University, Melbourne, AUS, Article 4, 6 pages.

[6] Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative research in psychology* 3, 2 (2006), 77–101.

[7] Virginia Braun and Victoria Clarke. 2021. One size fits all? What counts as quality practice in (reflexive) thematic analysis? *Qualitative research in psychology* 18, 3 (2021), 328–352.

[8] Trevor L Bruns, Andria A Remirez, Maxwell A Emerson, Ray A Lathrop, Arthur W Mahoney, Hunter B Gilbert, Cindy L Liu, Paul T Russell, Robert F Labadie, Kyle D Weaver, et al. 2021. A modular, multi-arm concentric tube robot system with application to transnasal surgery for orbital tumors. *The International Journal of Robotics Research* 40, 2-3 (2021), 521–533.

[9] Brian Bruya and Yi-Yuan Tang. 2018. Is attention really effort? Revisiting Daniel Kahneman's influential 1973 book attention and effort. *Frontiers in psychology* 9 (2018), 1133.

[10] Kelly Caine. 2016. Local standards for sample size at CHI. In *Proceedings of the 2016 CHI conference on human factors in computing systems*. 981–992.

[11] Hao Mark Chen, Liam Castelli, Martin Ferianc, Hongyu Zhou, Shuanglong Liu, Wayne Luk, and Hongxiang Fan. 2024. Enhancing Dropout-based Bayesian Neural Networks with Multi-Exit on FPGA. *arXiv preprint arXiv:2406.14593* (2024).

[12] Rawiphon Chotikunnan, Kittipan Roongprasert, Phichitphon Chotikunnan, Pariwat Imura, Manas Sangworasil, and Anuchart Srisiriwat. 2023. Robotic Arm Design and Control Using MATLAB/Simulink. *International Journal of Membrane Science and Technology* 10, 3 (2023), 2448–2459.

[13] Ganesh Choudhary and Chethan Ram BV. 2014. Real time robotic arm control using hand gestures. In *2014 International Conference on High Performance Computing and Applications (ICHPCA)*. IEEE, 1–3.

[14] Andrew Correa, Matthew R Walter, Luke Fletcher, Jim Glass, Seth Teller, and Randall Davis. 2010. Multimodal interaction with an autonomous forklift. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 243–250.

[15] Edirlei Soares De Lima and Bruno Feijó. 2019. Artificial intelligence in human-robot interaction. *Emotional Design in Human-Robot Interaction: Theory, Methods and Applications* (2019), 187–199.

[16] Frederik Diederichs, Lesley-Ann Mathis, Valeria Bopp-Bertenbreiter, Benjamin Bednorz, Harald Widlroither, and Frank Flemisch. 2021. A Wizard-of-Oz vehicle to investigate human interaction with AI-driven automated cars. In *Proceedings of the DSC*.

[17] Wojciech Dudek and Tomasz Winiarski. 2020. Scheduling of a robot's tasks with the tasker framework. *IEEE Access* 8 (2020), 161449–161471.

[18] Jonathan Eden, Mario Bräcklein, Jaime Ibáñez, Deren Yusuf Barsakcioglu, Giovanni Di Pino, Dario Farina, Etienne Burdet, and Carsten Mehring. 2022. Principles of human movement augmentation and the challenges in making it a reality. *Nature Communications* 13, 1 (2022), 1345.

[19] Hesham Elsayed, Kenneth Kartono, Dominik Schön, Martin Schmitz, Max Mühlhäuser, and Martin Weigel. 2022. Understanding Perspectives for Single-and Multi-Limb Movement Guidance in Virtual 3D Environments. In *Proceedings of*

the 28th ACM Symposium on Virtual Reality Software and Technology. 1–10.

[20] Hongxiang Fan, Martin Ferianc, Miguel Rodrigues, Hongyu Zhou, Xinyu Niu, and Wayne Luk. 2021. High-performance FPGA-based accelerator for Bayesian neural networks. In *2021 58th ACM/IEEE Design Automation Conference (DAC)*. IEEE, 1063–1068.

[21] Martin Ferianc, Hongxiang Fan, Divyansh Manocha, Hongyu Zhou, Shuanglong Liu, Xinyu Niu, and Wayne Luk. 2021. Improving performance estimation for design space exploration for convolutional neural network accelerators. *Electronics* 10, 4 (2021), 520.

[22] Lennon Fernandes and BR Shivakumar. 2020. Identification and Sorting of Objects based on Shape and Colour using robotic arm. In *2020 Fourth International Conference on Inventive Systems and Control (ICISC)*. IEEE, 866–871.

[23] Mikolaj Firlej and Araz Taeihagh. 2021. Regulating human control over autonomous systems. *Regulation & Governance* 15, 4 (2021), 1071–1091.

[24] Masaaki Fukuoka, Adrien Verhulst, Fumihiko Nakamura, Ryo Takizawa, Katsutoshi Masai, and Maki Sugimoto. 2019. FaceDrive: Facial expression driven operation to control virtual supernumerary robotic arms. In *SIGGRAPH Asia 2019 XR*. 9–10.

[25] Mar Gonzalez-Franco and Tabitha C Peck. 2018. Avatar embodiment. towards a standardized questionnaire. *Frontiers in Robotics and AI* 5 (2018), 74.

[26] Ken Gu, Madeleine Grunde-McLaughlin, Andrew McNutt, Jeffrey Heer, and Tim Althoff. 2024. How do data analysts respond to ai assistance? a wizard-of-oz study. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*. 1–22.

[27] Yves Guiard. 1987. Asymmetric division of labor in human skilled bimanual action: The kinematic chain as a model. *Journal of motor behavior* 19, 4 (1987), 486–517.

[28] Azra Habibovic, Jonas Andersson, Maria Nilsson, V Malmsten Lundgren, and Jan Nilsson. 2016. Evaluating interactions with non-existing automated vehicles: three Wizard of Oz approaches. In *2016 IEEE intelligent vehicles symposium (IV)*. IEEE, 32–37.

[29] Morten Hertzum. 2016. A usability test is not an interview. *interactions* 23, 2 (2016), 82–84.

[30] Morten Hertzum, Pia Borlund, and Kristina B Kristoffersen. 2015. What do thinking-aloud participants say? A comparison of moderated and unmoderated usability sessions. *International Journal of Human-Computer Interaction* 31, 9 (2015), 557–570.

[31] Leigh R Hochberg, Daniel Bacher, Beata Jarosiewicz, Nicolas Y Masse, John D Simeral, Joern Vogel, Sami Haddadin, Jie Liu, Sydney S Cash, Patrick Van Der Smagt, et al. 2012. Reach and grasp by people with tetraplegia using a neurally controlled robotic arm. *Nature* 485, 7398 (2012), 372–375.

[32] Brandi House, Jonathan Malkin, and Jeff Bilmes. 2009. The VoiceBot: a voice controlled robot arm. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 183–192.

[33] Siying Hu, Hen Chen Yen, Ziwei Yu, Mingjian Zhao, Katie Seaborn, and Can Liu. 2023. Wizundry: A cooperative Wizard of Oz platform for simulating future speech-based interfaces with multiple Wizards. *Proceedings of the ACM on Human-Computer Interaction* 7, CSCW1 (2023), 1–34.

[34] Yuhan Hu, Sang-won Leigh, and Pattie Maes. 2017. Hand development kit: Soft robotic fingers as prosthetic augmentation of the hand. In *Adjunct Proceedings of the 30th Annual ACM Symposium on User Interface Software and Technology*. 27–29.

[35] Yongquan Hu, Mingyue Yuan, Kaiqi Xian, Don Samitha Elvitigala, and Aaron Quigley. 2023. Exploring the design space of employing ai-generated content for augmented reality display. *arXiv e-prints* (2023), arXiv–2303.

[36] Yongquan Hu, Shuning Zhang, Ting Dang, Hong Jia, Flora D Salim, Wen Hu, and Aaron J Quigley. 2024. Exploring large-scale language models to evaluate eeg-based multimodal data for mental health. In *Companion of the 2024 on ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 412–417.

[37] Xincheng Huang, James Riddell, and Robert Xiao. 2023. Virtual reality telepresence: 360-degree video streaming with edge-compute assisted static foveated compression. *IEEE Transactions on Visualization and Computer Graphics* (2023).

[38] Pham Ngoc Hung and Takashi Yoshimi. 2016. Extracting actions from instruction manual and testing their execution in a robotic simulation. *ASEAN Engineering Journal* 6, 1 (2016), 47–58.

[39] Masahiko Inami, Daisuke Uriu, Zendai Kashino, Shigeo Yoshida, Hiroto Saito, Azumi Maekawa, and Michiteru Kitazaki. 2022. Cyborgs, human augmentation, cybernetics, and JIZAI body. In *Proceedings of the Augmented Humans International Conference 2022*. 230–242.

[40] Yukiko Iwasaki, Kozo Ando, Shuhei Iizuka, Michiteru Kitazaki, and Hiroyasu Iwata. 2020. Detachable Body: The impact of binocular disparity and vibrotactile feedback in co-presence tasks. *IEEE Robotics and Automation Letters* 5, 2 (2020), 3477–3484.

[41] Ziyi Jiang, Yanpei Huang, Jonathan Eden, Ekaterina Ivanova, Xiaoxiao Cheng, and Etienne Burdet. 2023. A virtual reality platform to evaluate the effects of supernumerary limbs' appearance. In *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*. IEEE, 1–5.

[42] Stephanie Kim, Jacy Reese Anthis, and Sarah Sebo. 2024. A taxonomy of robot autonomy for human-robot interaction. In *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. 381–393.

[43] Ryota Kondo, Yamato Tani, Maki Sugimoto, Masahiko Inami, and Michiteru Kitazaki. 2020. Scrambled body differentiates body part ownership from the full body illusion. *Scientific reports* 10, 1 (2020), 5274.

[44] Ryota Kondo, Sachiyo Ueda, Maki Sugimoto, Kouta Minamizawa, Masahiko Inami, and Michiteru Kitazaki. 2018. Invisible Long Arm Illusion: Illusory Body Ownership by Synchronous Movement of Hands and Feet.. In *ICAT-EGVE*. Limassol, 21–28.

[45] Simone Leone, Luigi Giunta, Vincenzo Rino, Simone Mellace, Alessio Sozzi, Francesco Lago, Elio Matteo Curcio, Doina Pisla, and Giuseppe Carbone. 2024. Design of a Wheelchair-Mounted Robotic Arm for Feeding Assistance of Upper-Limb Impaired Patients. *Robotics* 13, 3 (2024), 38.

[46] Chiuhsiang Joe Lin, Wei-Jung Shiang, Chun-Yu Chuang, and Jin-Liang Liou. 2014. Applying the skill-rule-knowledge framework to understanding operators' behaviors and workload in advanced main control rooms. *Nuclear Engineering and Design* 270 (2014), 176–184.

[47] Azumi Maekawa, Kei Kawamura, and Masahiko Inami. 2020. Dynamic assistance for human balancing with inertia of a wearable robotic appendage. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 4077–4082.

[48] Pat Marion, Maurice Fallon, Robin Deits, Andrés Valenzuela, Claudia Pérez D'Arpino, Greg Izatt, Lucas Manuelli, Matt Antone, Hongkai Dai, Twan Koolen, et al. 2018. Director: A user interface designed for robot operation with shared autonomy. *The DARPA Robotics Challenge Finals: Humanoid Robots To The Rescue* (2018), 237–270.

[49] Qaid Mohammed Marwan, Shing Chyi Chua, and Lee Chung Kwek. 2021. Comprehensive review on reaching and grasping of objects in robotics. *Robotica* 39, 10 (2021), 1849–1882.

[50] Reiji Miura, Shunichi Kasahara, Michiteru Kitazaki, Adrien Verhulst, Masahiko Inami, and Maki Sugimoto. 2021. Multisoma: Distributed embodiment with synchronized behavior and perception. In *Proceedings of the Augmented Humans International Conference 2021*. 1–9.

[51] Marie Muehlhaus, Marion Koelle, Artin Saberpour, and Jürgen Steimle. 2023. I need a third arm! eliciting body-based interactions with a wearable robotic arm. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. 1–15.

[52] Janet H Murray. 2020. Virtual/reality: how to tell the difference. *Journal of visual culture* 19, 1 (2020), 11–27.

[53] Junichi Nabeshima, MHD Yamen Saraiji, and Kouta Minamizawa. 2019. Arque: artificial biomimicry-inspired tail for extending innate body functions. In *ACM SIGGRAPH 2019 Posters*. 1–2.

[54] Merel Noorman and Deborah G Johnson. 2014. Negotiating autonomy and responsibility in military robots. *Ethics and Information Technology* 16, 1 (2014), 51–62.

[55] Susanne Palmcrantz, Anneli Wall, Katarina Skough Vreede, Påvel Lindberg, Anna Danielsson, Katharina S Sunnerhagen, Charlotte K Häger, and Jörgen Borg. 2021. Impact of intensive gait training with and without electromechanical assistance in the chronic phase after stroke–a multi-arm randomized controlled trial with a 6 and 12 months follow up. *Frontiers in Neuroscience* 15 (2021), 660726.

[56] Federico Parietti and Harry Asada. 2016. Supernumerary robotic limbs for human body support. *IEEE Transactions on Robotics* 32, 2 (2016), 301–311.

[57] Federico Parietti and Harry H Asada. 2013. Dynamic analysis and state estimation for wearable robotic limbs subject to human-induced disturbances. In *2013 IEEE International Conference on Robotics and Automation*. IEEE, 3880–3887.

[58] Federico Parietti and H Harry Asada. 2014. Supernumerary robotic limbs for aircraft fuselage assembly: body stabilization and guidance by bracing. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 1176–1183.

[59] Federico Parietti, Kameron C Chan, Banks Hunter, and H Harry Asada. 2015. Design and control of supernumerary robotic limbs for balance augmentation. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 5010–5017.

[60] Christian Penaloza, David Hernandez-Carmona, and Shuichi Nishio. 2018. Towards intelligent brain-controlled body augmentation robotic limbs. In *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 1011–1015.

[61] John Sören Pettersson and Malin Wik. 2014. Perspectives on Ozlab in the cloud: A literature review of tools supporting Wizard-of-Oz experimentation, including an historical overview of 1971-2013 and notes on methodological issues and supporting generic tools. (2014).

[62] Elisa Prati, Valeria Villani, Margherita Peruzzini, and Lorenzo Sabattini. 2021. An approach based on VR to design industrial human-robot collaborative workstations. *Applied Sciences* 11, 24 (2021), 11773.

[63] Domenico Prattichizzo, Maria Pozzi, Tommaso Lisini Baldi, Monica Malvezzi, Irfan Hussain, Simone Rossi, and Gionata Salvietti. 2021. Human augmentation by wearable supernumerary robotic limbs: review and perspectives. *Progress in*

Biomedical Engineering 3, 4 (2021), 042005.
[64] David Putrino, Yan T Wong, Adam Weiss, and Bijan Pesaran. 2015. A training platform for many-dimensional prosthetic devices using a virtual reality environment. *Journal of neuroscience methods* 244 (2015), 68–77.
[65] MHD Yamen Saraiji, Tomoya Sasaki, Kai Kunze, Kouta Minamizawa, and Masahiko Inami. 2018. Metaarms: Body remapping using feet-controlled artificial arms. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology*. 65–74.
[66] MHD Yamen Saraiji, Tomoya Sasaki, Reo Matsumura, Kouta Minamizawa, and Masahiko Inami. 2018. Fusion: full body surrogacy for collaborative communication. In *ACM SIGGRAPH 2018 Emerging Technologies*. 1–2.
[67] Tomoya Sasaki, MHD Yamen Saraiji, Charith Lasantha Fernando, Kouta Minamizawa, and Masahiko Inami. 2017. MetaLimbs: Multiple arms interaction metamorphism. In *ACM SIGGRAPH 2017 emerging technologies*. 1–2.
[68] Christina Schmidbauer. 2022. Adaptive task sharing between humans and cobots in assembly processes. *TU Wien* (2022).
[69] Ian Sharp, Felix Huang, and James Patton. 2011. Visual error augmentation enhances learning in three dimensions. *Journal of neuroengineering and rehabilitation* 8 (2011), 1–6.
[70] Hideki Shimobayashi, Tomoya Sasaki, Arata Horie, Riku Arakawa, Zendai Kashino, and Masahiko Inami. 2021. Independent control of supernumerary appendages exploiting upper limb redundancy. In *Proceedings of the Augmented Humans International Conference 2021*. 19–30.
[71] Brandon Victor Syiem, Ryan M Kelly, Jorge Goncalves, Eduardo Velloso, and Tilman Dingler. 2021. Impact of task on attentional tunneling in handheld augmented reality. In *Proceedings of the 2021 CHI conference on human factors in computing systems*. 1–14.
[72] Brandon Victor Syiem, Ryan M Kelly, Eduardo Velloso, Jorge Goncalves, and Tilman Dingler. 2020. Enhancing visitor experience or hindering docent roles: attentional issues in augmented reality supported installations. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 279–288.
[73] Philipp Sykownik, Katharina Emmerich, and Maic Masuch. 2018. Exploring patterns of shared control in digital multiplayer games. In *Advances in Computer Entertainment Technology: 14th International Conference, ACE 2017, London, UK, December 14-16, 2017, Proceedings 14*. Springer, 847–867.
[74] Krzysztof Adam Szczurek, Raul Marin Prades, Eloise Matheson, Jose Rodriguez-Nogueira, and Mario Di Castro. 2023. Multimodal multi-user mixed reality human–robot interface for remote operations in hazardous environments. *IEEE Access* 11 (2023), 17305–17333.
[75] Ryo Takizawa, Adrien Verhulst, Katie Seaborn, Masaaki Fukuoka, Atsushi Hiyama, Michiteru Kitazaki, Masahiko Inami, and Maki Sugimoto. 2019. Exploring perspective dependency in a shared body with virtual supernumerary robotic arms. In *2019 IEEE international conference on artificial intelligence and virtual reality (AIVR)*. IEEE, 25–257.
[76] Adele Tong, Praneeth Perera, Zhanna Sarsenbayeva, Alistair McEwan, Anjula C De Silva, and Anusha Withana. 2023. Fully 3D-printed dry EEG electrodes. *Sensors* 23, 11 (2023), 5175.
[77] Yuchuang Tong and Jinguo Liu. 2021. Review of research and development of supernumerary robotic limbs. *IEEE/CAA Journal of Automatica Sinica* 8, 5 (2021), 929–952.
[78] Albert Tung, Josiah Wong, Ajay Mandlekar, Roberto Martín-Martín, Yuke Zhu, Li Fei-Fei, and Silvio Savarese. 2021. Learning multi-arm manipulation through collaborative teleoperation. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 9212–9219.
[79] Kohei Umezawa, Yuta Suzuki, Gowrishankar Ganesh, and Yoichi Miyawaki. 2022. Bodily ownership of an independent supernumerary limb: an exploratory study. *Scientific reports* 12, 1 (2022), 2339.
[80] Vighnesh Vatsal and Guy Hoffman. 2017. Wearing your arm on your sleeve: Studying usage contexts for a wearable robotic forearm. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 974–980.
[81] Vighnesh Vatsal and Guy Hoffman. 2018. Design and analysis of a wearable robotic forearm. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 5489–5496.
[82] Eric Wade and Carolee J Winstein. 2011. Virtual reality and robotics for stroke rehabilitation: where do we go from here? *Topics in stroke rehabilitation* 18, 6 (2011), 685–700.
[83] Baicun Wang, Pai Zheng, Yue Yin, Albert Shih, and Lihui Wang. 2022. Toward human-centric smart manufacturing: A human-cyber-physical systems (HCPS) perspective. *Journal of Manufacturing Systems* 63 (2022), 471–490.
[84] Astrid Weiss, Regina Bernhaupt, Daniel Schwaiger, Martin Altmaninger, Roland Buchner, and Manfred Tscheligi. 2009. User experience evaluation with a wizard of oz approach: Technical and methodological considerations. In *2009 9th IEEE-RAS International Conference on Humanoid Robots*. IEEE, 303–308.
[85] Mahisorn Wongphati, Yushi Matsuda, Hirokata Osawa, and Michita Imai. 2012. Where do you want to use a robotic arm? And what do you want from the robot?. In *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 322–327.
[86] Bo Yang, Jian Huang, Xinxing Chen, Caihua Xiong, and Yasuhisa Hasegawa. 2021. Supernumerary robotic limbs: A review and future outlook. *IEEE Transactions on Medical Robotics and Bionics* 3, 3 (2021), 623–639.
[87] Geng Yang, Zhibo Pang, M Jamal Deen, Mianxiong Dong, Yuan-Ting Zhang, Nigel Lovell, and Amir M Rahmani. 2020. Homecare robotic systems for healthcare 4.0: Visions and enabling technologies. *IEEE journal of biomedical and health informatics* 24, 9 (2020), 2535–2549.
[88] Muhammad Hamza Zafar, Even Falkenberg Langås, and Filippo Sanfilippo. 2024. Exploring the synergies between collaborative robotics, digital twins, augmentation, and industry 5.0 for smart manufacturing: A state-of-the-art review. *Robotics and Computer-Integrated Manufacturing* 89 (2024), 102769.
[89] Kailing Zhang, Yi Long, and Xiaofeng Luo. 2023. Review of Supernumerary Robotic Limbs. In *Journal of Physics: Conference Series*, Vol. 2456. IOP Publishing, 012004.
[90] Mengyang Zhang, Guohui Tian, Ying Zhang, and Peng Duan. 2021. Service skill improvement for home robots: Autonomous generation of action sequence based on reinforcement learning. *Knowledge-Based Systems* 212 (2021), 106605.
[91] Hongyu Zhou, Treshan Ayesh, Chenyu Fan, Zhanna Sarsenbayeva, and Anusha Withana. 2024. CoplayingVR: Understanding User Experience in Shared Control in Virtual Reality. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 8, 3 (2024), 1–25.

A Autonomy level details

Before starting the study, we introduced the participants to an action guide detailing the operational behaviors of the virtual supernumerary limbs. The operator adhered to this predefined guide, which specified responses to participant commands across different autonomy levels.

A.1 Low-autonomy level

Description: At this level, the virtual supernumerary limbs have basic capabilities, including the ability to perceive what is happening in their environment. They can understand and respond to simple instructions and learn from user interactions. However, while they can recognize general elements in the environment, they cannot automatically determine where or how to place specific objects without explicit user guidance. The arms rely on you to guide them step by step on how to complete tasks, categorizing items based on your instructions.

Capabilities: While they understand their surroundings, they depend on you to guide their actions. They respond well to step-by-step instructions where you guide each action.

Role: You will need to provide detailed, step-by-step guidance through verbal or gestural commands on how the virtual supernumerary limbs should sort and handle objects. These arms can only recognize and act upon the tasks you have specifically taught them. They cannot automatically identify or categorize items they haven't learned before, so they rely entirely on your instructions to understand and execute each action.

A.2 High-autonomy level

Description: This version of the virtual supernumerary limb system is significantly more advanced, with improved language comprehension and the ability to understand complex tasks. It also automatically recognizes things in the environment, reducing the need for teach them, and allows you to demonstrate and execute complex tasks and instructions. However, they do not inherently know what actions to take; they need you to instruct them on what to do.

Capabilities: The virtual supernumerary limbs can autonomously recognize their surroundings. However, they still require your instructions to determine what actions to take. Once directed, they can efficiently execute tasks with minimal guidance.

Role: You will need to provide clear instructions on what actions the arms should take. Whether it's sorting objects, placing them in specific locations, or any other task, the virtual supernumerary limbs depend on your directives to know what to do. The arms are designed to follow your commands accurately, but they rely on you to define the task and guide their actions accordingly.

B Operator Action Guidelines

The operator followed a predefined set of action rules to ensure consistency across participants and tasks. These guidelines were tailored to the study's four instruction types and autonomy levels, aligning with principles from multi-agent coordination and robot learning frameworks [4, 22, 38, 90]:

B.1 Command Instructions

Low-autonomy: Perform each step based on explicit participant guidance (e.g., grasp, move, rotate). For instance, the operator can rotate a block incrementally in response to specific participant instructions for each action. If a series of commnads given, only execute the first. If a complex command such as a combined task given, do not perform any actions. *High-autonomy*: Execute the full sequence or complex actions based on the command's intent (e.g., directly move a labeled object to a specified position).

B.2 Demonstration Instructions

Low-autonomy: Mimic participant's demonstrations step-by-step. If the participant rotated an object, the operator replicated the action. *High-autonomy*: Learn the demonstrated action and execute autonomously for similar objects. For instance, stacking two blocks based on a demonstration was applied to subsequent blocks independently.

B.3 Delegation Instructions

Low-autonomy: No action was taken, as Delegation Instructions were not supported under this autonomy level. *High-autonomy*: Plan and execute tasks sequentially after an instruction. A single command like "Organize objects by size in a row" prompted the operator to complete the task autonomously.

B.4 Labeling Instructions

Low-autonomy: Identify and label objects explicitly based on participant input. For example, the operator must wait for the participant to confirm "The red squar" before proceeding. *High-autonomy*: Automatically identify attributes and proceed with the complex tasks without additional guidance.