UNIVERSITY OF PADUA

PH.D. SCHOOL IN BRAIN, MIND AND COMPUTER SCIENCE
Curriculum in Neuroscience, Technology and Society
XXXV Cycle

Human motion as a natural control of industrial robots in VR: insights on users' performance and workload

Candidate:
Federica Nenna

Supervisor:
Prof. Luciano Gamberini
University of Padua

Co-Supervisor:
Prof. Mauro Conti
University of Padua

Abstract

Recently, the industrial sector has witnessed a massive shift of general interest from the machine to the human, who have become the core of the current industrial evolution. The manifest of the Industry 5.0 is indeed the 'human-centric manufacturing', which places the worker's well-being at the center of the production process. In this framework, a prominent topic is also covered by the increasingly aging workforce, which is bringing particular attention to the senior worker. Indeed, aging brings cognitive and physical decay and is typically related to decreasing flexibility and adaptation to new technologies, which may play a counterposing role to the industrial technological progress. Cognitive research is thus gaining room in all discussions involving workers and machines, and it is supporting the human-centric approach by providing workers' performance and workload assessments.

One of the most significant innovations endorsed by Industry 5.0 is leveraging Virtual Reality (VR) for teleoperating or simulating industrial robots. However, VR industrial applications are fairly recent, and workers' psychophysical aspects have been regrettably marginalized until now. Consequently, whether and how VR robot teleoperations are beneficial from human factors' perspective is still under debate, and the role of cognitive science in this respect is now essential.

With this Ph.D. thesis, we thus aim to provide a broad overview of performance and workload of users simulating robotic teleoperations in VR. We conducted 5 experimental studies, whose common thread is the industrial robot UR10e that was here purposely reproduced in VR via Unity. The VIVE Pro Eye VR headset was also deployed in all experiments; it is provided with an integrated eye tracker, which offers an exceptional opportunity for continuous workload monitoring during robotic teleoperation. Furthermore, the last study was conducted at the Berlin Mobile Brain/Body Imaging (MoBI) Laboratory, which provides dedicated tools and approaches for measuring Electroencephalography (EEG) during free motion.

The strength of the presented studies thus resides in the combination of multiple metrics for analyzing human behaviors and brain activity during simulated teleoperations, as well as in the assortment of knowledge coming from cognitive science, human factors, human-robot interaction and computer science sectors. Overall, our results significantly contribute to the state of the art on VR-based telerobotics, particularly offering a multimodal and multifaceted overview of human performance and workload when guiding an industrial robotic arm in VR.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Industry 4.0, and more recently 5.0, are removing the perceived boundary of a manufacturing environment and allowing humans to interact with machines in a fluid-tight manner, also remotely. This is the case of **human-robot collaborative systems**, which are also called collaborative robots, or cobots (Krüger et al., 2009). Compared to traditional robots, cobots are provided with a manual guidance feature, which allows to manually drive them by pushing/pulling specific points of their structure. This means that humans are allowed to use their own hands to drive the robot over desired positions. In this way, the amount of time typically spent on robot programming is significantly streamlined and reduced. Moreover, the traditional support function that robots have for humans is taken to the next level, as now cobots are able not only to support operators in carrying out a given task, but they literally share the task with them (Faccio et al., 2022; Nenna et al., 2022a).

However, there are **hazard environments** that don't allow humans to approach machines for safety reasons, such as underwater, underground, in nuclear facilities, aerial or space (Szczurek et al., 2022). Other examples of hazard factory environments and tasks are the followings: radioactive product disassembly, chemical product assembly, bio-material production (Liu and Wang, 2020). An analogous scenario sees operators who are unable to perform inspection, maintenance or repair of robotic systems in person, as happened quite often during the latest pandemic. Notably, with the advent of the COVID pandemic, interest in remote control and VR has turned into an urgent necessity even in the manufacturing field (Melluso et al., 2020). The spread of the pandemic has indeed emphasized the weaknesses of current industrial systems and shed light on the importance of teleoperation and remote control of production sites. In these situations, **telerobotics** can help driving robots remotely. Specifically, telerobotics (Niemeyer et al., 2016) aims to replicate human manipulative skills over any distances and at any scale to a remote workplace. In this regard, the concept of telepresence is particularly relevant. It can be defined as "the feeling of being present at another location than the physical location of one's body" (Minsky, 1980). Ideally, operators teleoperating a robot should have the impression of being physically on the remote side, in a way that the interaction with the robot feels as natural and intuitive as possible.

Even in such remote interactions, the **manual guidance** feature would be useful. Indeed, there are examples of remote human-robot collaborative production systems that aim at allowing manual manipulation of a physical replica of an industrial cobot connected with its local counterpart (Liu and Wang, 2020). Recently, similar collaborative frameworks involving physical interactions between humans and robots were also developed by leveraging immersive Virtual Reality (VR) (Du et al., 2021; Wonsick and Padir, 2020), which is the highest interactive digital tool per excellence.

The close integration of technology, robots, automated factory lines and intelligent manufacturing generated a significant **increase in skill demands** for the operator (Doolani et al., 2020). Indeed, workers are now called to master the technical complexity of the novel industrial systems and to cope with cognitive loads to work efficiently. Therefore, despite the latest evolution of manufacturing brought positive outcomes, the related increased complexity also led to new challenges as a collateral effect (Nguyen Ngoc et al., 2022).

This was one of the motives that promoted the establishment of **Industry 5.0**. This paradigm provided a proper response to the new challenges that arose after the technical advancements in smart and intelligent manufacturing of Industry 4.0, which was mainly characterized by an overturning of the human-machine relation. Indeed, with Industry 5.0, the general interest of researchers and stakeholders massively shifted from the machine to the human, who has become the core of such an industrial revolution. Not without reason, the manifest of Industry 5.0 is the **'human-centric manufacturing'** (Lu et al., 2022), which placed the worker's well-being at the center of the production process. In this framework, the urge to better understand the single worker in the field, with his/her individual needs, attitudes and capabilities, thus arose.

A 'human-centered design' incorporates knowledge and techniques coming from **human factors and ergonomics** fields to make systems usable (ISO, 2019). Specifically, it is to notice that, in the human-robot collaborative framework, the design complexity can run in two directions: from the user to the robot, and vice versa (Nguyen Ngoc et al., 2022). In the first case, the human must be able to effectively cooperate with all components of a given physical system, while simultaneously exchanging data with the system for hybrid decision-making. In the second case, the design of such smart systems must be capable of sensing and responding to the humans they interact with.

This highlights the urge of **multidisciplinary research** when addressing issues that concern human-robot collaborative frameworks, as both the technical aspects of the machine and the cognitive aspects of the human are essential in such interplay. In its survey on Human-Robot Interaction (HRI), Goodrich et al. (2008) claimed that:

"*[...] there are a number of accepted practices that are emerging in HRI. A key practice is to include experts from multiple disciplines in research efforts. These disciplines frequently include robotics, electrical and mechanical engineering, computer science, human–computer interaction, cognitive science, and human factors engineering. Other relevant disciplines include design, organizational behavior, and the social sciences. Importantly, some conferences encourage multidisciplinary submissions are working to establish the practice of having all papers refereed by* *reviewers representing different disciplines.*".

Nonetheless, to date, the majority of the research works that were produced on Industry 4.0 and 5.0 areas demonstrate how higher attention is always dedicated to the technical aspects, while the analysis of human factors, and even more of cognitive factors, is consistently marginalized (Grandi et al., 2020).

The **scarce consideration of human factors** in the design of human-robot collaborative frameworks is alarming. Indeed, as telerobotics is a wide area that finds applications in various everyday scenarios, it might lead to the outbreak of complex problems in any of those scenarios, with often unknown consequences. For instance, as accurately explained by Nguyen Ngoc et al. (2022), consequences of systems that neglect the human performance and cognitive states are "nuclear accidents, market failures in new product development, robotic-surgery-related adversities, technological accidents during machine manipulation, and interaction issues among humans and smart systems". This goes to show that assessing, monitoring and improving performance and cognitive loads of operators dealing with smart manufacturing interfaces definitely matters.

Human factors research can help in this direction. Specifically, the human factors' field embraces **engineering psychology, ergonomics, cognitive ergonomics, and accident analysis** (Goodrich et al., 2008). Over the years, it produced concepts such as mental workload (Hart, 2006), situation awareness (Endsley et al., 2003) and trust in automation (Lee and See, 2004), but also very important theories of cognitive ergonomics, such as Wickens's Multiple Resource Theory (Wickens, 2002). All these concepts and frameworks are now largely accepted in the most various applied fields, and became essential to the understanding of any human-robot complex. One of the most common practices in human factors is indeed to provide workers' performance and workload measurements, which would definitely support the human-centric approach in human-robot collaborative frameworks (Kaufeld and Nickel, 2019).

On a higher level, **neuroergonomics** is a multidisciplinary field that integrates theories and principles from human factors, ergonomics and neuroscience (Parasuraman and Rizzo, 2006; Parasuraman and Wilson, 2008). Therefore, it embraces all concepts addressed in this thesis. Its main purpose is to better understand the relationships that occur between brain function and behavioral outcomes in the work sector and everyday life. More broadly, it can thus be defined as the study of the human brain and performance at work and in everyday settings (Parasuraman and Rizzo, 2006).

Human factors and neuroergonomics research thus started to propose and test different approaches for assessing human mental states while interacting with a robotic system (Villani et al., 2020). Specifically, recent methodological advances allow multimodal assessments of human factors, combining different behavioral and physiological tools to fully understand humans in the field (Dehais et al., 2020a; Ladouce et al., 2017; Matthews et al., 2015;

Wascher et al., 2020). Mobile eye-tracking technology (Fuhl et al., 2016) and Mobile Brain/Body Imaging (MoBI) (Jungnickel et al., 2019) are just two examples of recent methodologies that allow ecologically valid assessment of human behavioral, cognitive and neural dynamics, also in free motion, and that will be tackled in the present work.

In this respect, **VR** is also becoming an exceptionally valuable methodological tool for different reasons. First, as emphasized by the recent expansion of the **Metaverse** (Mystakidis, 2022), it is increasingly permeating the most varied social contexts, like practical work (Pérez et al., 2019) and entertainment fields (Ansari et al., 2022). Therefore, any research conducted in VR, if uses a reasonably realistic virtual environment, could be easily transposed to practical contexts. Second, and most importantly for cognitive scientists, VR devices allow **full tracking** of humans' actions, interactions, physical movements, and lately even eye movements, without motion constraints. Those characteristics make it possible to collect tons of behavioral data during free action that, if correctly processed and interpreted, can be helpful in explaining or even predicting human intentions, making it a powerful research tool too.

1.1 Research questions, motivations and contributions

The present thesis tackles all the above-mentioned questions by presenting a series of research works that touch different aspects of telerobotics, always emphasizing the users' performance, experienced effort, mental workload and motor control.

What primarily motivated this work is the very recent establishment of Industry 5.0. While its aims are clear, systematic human-centered research with a cognitive characterization is still scarce. We opted for addressing specifically the topic of VR-based telerobotics, as it is increasingly becoming a promising sector, also in manufacturing. An extra gear is given by the VR headset model chosen for our investigations, which is additionally provided with eye-tracking technology. Such technology can help significantly in detecting both implicit (e.g., eye movements) and explicit (e.g., actions and interactions) behaviors of a potential teleoperator, informing on his/her level of workload and fatigue throughout the tasks, without disruptions. Remarkably, while VR was extensively applied in many sectors, and also in telerobotics, its potential in combination with eye-tracking technology was uncovered. Similarly, for increasing the variety of the registered human metrics even more, in the last study we further leveraged mobile EEG to gain brain data during simulated teleoperation, and thus matched behavioral, eye-tracking and self-reports metrics with brain dynamics. The acquisition of brain data in free motion (see MoBI, Mobile Brain-Body Imaging) has proven enormous benefits in various research areas, and particularly the spatial cognition and navigation fields. However, its potential in the fields of industry and telerobotics has yet to be fully explored.

We thus treasured these gaps and followed the ongoing shift of general interest toward the human interacting with robotic systems, rather than on robotic systems themselves. Hence, we gave particular emphasis to performance, fatigue, motor control and workload aspects of the worker in and through VR. Particularly, in the next subsection, the five experimental studies we conducted are briefly introduced, indicating the main research question, motivations and contributions relative to each of them. Each of these studies was or will soon be unfolded in a different scientific publication, as also indicated below. Please, note that, in this dissertation, some passages have been quoted verbatim, and some figures have been reused from the following works - all coauthored by the author of the thesis - after approval of all co-authors: Nenna et al. (2022a); Nenna and Gamberini (2022); Nenna et al. (2022b).

1.1.1 Study 1

Research question. Does driving a virtual rather than physical robot benefit the user?

Motivation. In each Human-Robot Interaction (HRI) or even Collaboration (HRC), both humans and robots have strengths to be valued and weaknesses to be bridged. In this view, each synergistic collaboration between a human and a robot should meet technical and ergonomic standards to optimize the operator's physical and mental workload (Nachreiner et al., 2006). The aim is always to improve the overall system performance and maximize the benefit of the interaction. However, to date, human psychophysical aspects have been regrettably marginalized. The majority of the papers dealing with smart manufacturing have focused on the feasibility of the technical systems and the efficacy of the related framework (Liu and Wang, 2020), often neglecting the factors affecting the final user (Damiani et al., 2018). Indeed, it is still unknown how guiding a robot through a virtual simulation impacts the user's performance and cognitive workload and how it differs from collaboration with a physical robotic arm.

Contribution. To address this research gap, we first faithfully reproduced the robotic arm UR10e (depicted in Figure 1.1) in VR via Unity (depicted in Figure 1.2). Thereafter, we explored how interacting with the physical robotic arm and with its VR-based counterpart affects the user. More specifically, the virtual simulation was tested compared to the physical robot in participants (n=26) executing a pick-and-place task under low (single-task) and high mental demands (dual-task, concurrently with an arithmetic task). In this framework, the utility of VR was corroborated only if the performance of users driving the virtual robot did not decrease compared to driving the physical one. Data collected include operation times and task error as performance measures, changes in pupil size as a function of implicit workload, and self-reported explicit workload.

Publication. Nenna, F., Orso, V., Zanardi, D., & Gamberini, L. (2022). The virtualization of human–robot interactions: a user-centric workload assessment. *Virtual Reality*, 1-19. (**JCR IF 2021: 4.69**)

FIGURE 1.1: The robotic arm UR10e and its workstation

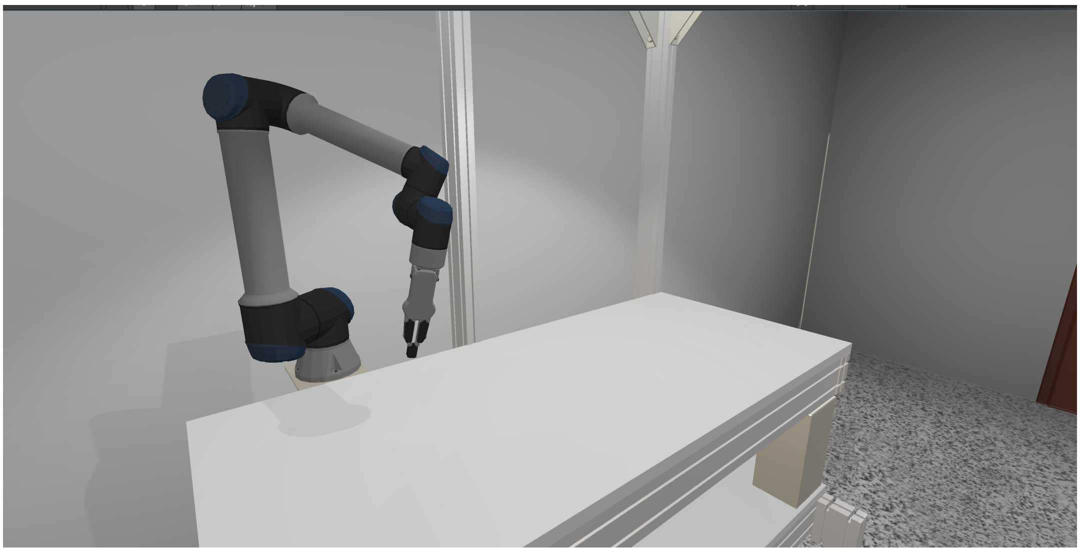

FIGURE 1.2: The VR replica of the robotic arm UR10e and its workstation

1.1.2 Study 2

Research question. Which is the most efficient and less demanding teleoperation modality in VR (i.e., button-based, action-based)?

Motivation. One emerging use case of VR devices in the industrial field is robot teleoperation (Nachreiner et al., 2006; Martín-Barrio et al., 2020; Rosen et al., 2018). As compared to traditional teleoperation means, VR allows to immerse in realistic environments and use interaction modalities that go far beyond the keyboard button press. This makes it possible for example to teleoperate a robot by using gestures or physical actions (Martín-Barrio et al., 2020) rather than using the conventional keyboard, mouse and joysticks (You and Hauser, 2012; Mavridis et al., 2015), increasing the degree of interactivity significantly. The strength of action-based teleoperations is that they leverage natural and embodied controls, allowing users to directly manipulate a replica of the robot in a button-free fashion, and to perform physical actions or gestures similar to those they would perform if manipulating the physical robot itself. On the other hand, it is also true that button-based input modalities are commercially widespread and thus more familiar to the most, likely being easier to learn. Literature on this topic is scarce, and a systematic assessment of mental workload when teleoperating a robot through action-based rather than button-based modalities would be beneficial for clarifying such aspect.

Contribution. We thus conducted a systematic investigation on users (n=24) driving the virtual replica of the robotic arm UR10e through the same pick-and-place task designed for the first study, which simulated robotic teleoperation. We asked participants to perform the same task using different control systems (button-based, action-based) and under different levels of task demand (single-task, dual-task). As a dual-task, they performed the pick-and-place task concurrently with an arithmetic task. Performance was assessed via operation time and error rate at the task, while the user's mental workload was assessed through both traditional self-reports and eye-tracking indexes detected from the VR headset. In this study, we better unfolded aspects regarding the sensibility of VR-gained eye-tracking parameters to workload too. Specifically, we related eye parameters continuously collected throughout the task in VR with the traditionally self-reported workload.

Publication. Nenna, F., Zanardi, D., & Gamberini, L. (2022). Human-centric telerobotics: investigating users' performance and workload via VR-based eye-tracking measures. (arXiv preprint: *arXiv:2212.07345*, Submitted for journal publication)

1.1.3 Study 3

Research question. Does VR-based teleoperation performance change with gender, individual skills and attitudes towards technology?

Motivation. If adopting a user-centric perspective is the first motive of Industry 5.0, understanding how different groups of people respond to robot teleoperations is surely of interest in this sector. Indeed, a valid HRI should be effective for the majority of the population. However, gender, gaming experience, or other individual factors like learnability, problem-solving skills or trust in technology are often likely to affect users' performance when interacting with a robot (Gomer and Pagano, 2011; Hancock et al., 2011; Paperno et al., 2019; Showkat and Grimm, 2018; Welfare et al., 2019). Even though those individual factors may have a significant impact on users' behavior when driving a robotic system, research systematically investigating relations between these individual factors and the actual user's performance in the HRI domain is scarce.

Contribution. We unfolded such aspects by collecting information about gender, gaming experience, learnability skills, problem solving and trust in technology in (n=23) individuals who were called to drive our virtual robotic arm through the same pick-and-place task via both control modalities (i.e, action-based, button-based). In this way, we first clarified whether any of the considered individual factors affected performance and self-reported workload, and second, we additionally investigated whether their effects cut across the two control systems.

Publication. Nenna, F., & Gamberini, L. (2022, March). The Influence of Gaming Experience, Gender and Other Individual Factors on Robot Teleoperations in VR. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 945-949).

1.1.4 Study 4

Research question. Can senior users teleoperate well enough in VR? Does their workload capacity differ from young users?

Motivation. Given the increasingly aging workforce, another research piece was dedicated to who we call *senior workers*. Those workers are more than 50 years old, and are approaching the onset of the age-related gradual functional capacities decrease (Donato et al., 2003; Kenny et al., 2008). Besides causing mental and physical decay, aging is also typically related to decreasing flexibility and adaptation to new technologies (Calzavara et al., 2020; Di Pasquale et al., 2020), which may play a counterposing role in the virtualization of manufacturing systems. Therefore, the age factor definitely needs to be contemplated before considering VR-based telerobotics a good option for all manufacturing workers.

Contribution. We thus systematically assessed: i) whether and how senior users can perform simulated robotic teleoperations via such new interactive technologies; ii) if their workload capacity can withstand repetitive tasks in VR; iii) whether their individual attitudes toward such new technologies differs from younger ones. Young (n=15) and senior (n=15) individuals drove the virtual robot through the same pick-and-place task in VR, both via the button-based and action-based control system, and under low (single-task) and high (dual-task) task demands. Performance and workload aspects were thoroughly examined in terms of operation time and accuracy, eye-tracking parameters (i.e., pupil size variation and perclos) and self-reports (i.e., the NASA-TLX questionnaire, and a self-report on the individual factors).

Publication. Nenna, F., Zanardi, D., Pluchino, P., & Gamberini, L. (*2023*). Can senior workers benefit from VR as well? Examining the role of age on robot teleoperation performance and workload. Submitted for journal publication

1.1.5 Study 5

Research questions. Is higher embodiment of an industrial robotic arm related to better performance and lower workloads? Do the neurophysiological signatures of embodiment and motor control also reveal in our industrial scenario?

Motivation. Literature demonstrated that embodiment can occur in humanoid robots (Farizon et al., 2021) and non-human-looking robots (Aymerich-Franch et al., 2017). Here, we assume that such mechanisms may possibly be triggered even in virtual robotic simulations usually employed in the industrial context, possibly extending to the scenario of Industry 5.0. Remarkably, we were interested in assessing whether the previously assumption of increased teleoperation performance for increasing embodiment (Iwasaki et al., 2022; Toet et al., 2020; Verhagen et al., 2020) actually applies to our industrial VR environment. In such a case, our participants should perform the teleoperation task in a faster and more accurately way in case they felt higher embodiment into the VR replica of the same UR10e employed in all our studies. Differently from the other studies, we provided the participants with a first person view of the workstation, and let them drive the robot by simply moving their own right arm (thanks to the application of a body tracker on their wrist). In this scenario, we introduced temporal and spatial inconsistencies between the robotic arm and the participant's arm with the intention of creating different levels of embodiment (Aymerich-Franch et al., 2016; D'Angelo et al., 2018; Farizon et al., 2021; Ratcliffe and Newport, 2017). In addition to various self-report questionnaires, we collected performance data computed through the VR device, in combination with a Mobile Brain/Body Imaging (MoBI) approach for additionally measuring brain dynamics throughout the task execution. These findings can be important for human-robot interface design, as they would cover an important gap on embodiment and motor control of industrial robots, and would bring significant benefits both to the industrial sector (by clarifying how to increase the efficiency of VR-based telerobotics) and to the final user (by increasing the quality of the VR task and environment, while facilitating motor control and lowering workloads).

Publication. Nenna, F., Gramann, K., Gehrke, L., & Gamberini, L. (*2023*). Is it my arm? Embodiment, sense of agency and ownership of an industrial robotic arm in VR. Work in progress

1.2 Common threads

Methodological tools. All studies elaborated in the present thesis share the same industrial robot UR10e. This industrial robotic arm is a collaborative robot, and as such, it allows close work with the human operator. It is commonly employed in manufacturing, and therefore, represents a concretely applied industrial tool. Furthermore, the same VIVE Pro Eye VR device was used in all experiments. It is provided with an integrated eye-tracker that can measure workload-related indexes within virtual environments without disrupting users' actions. This is an exceptional opportunity for continuous workload monitoring during robotic teleoperation.

Human-centric approach. Also, research questions followed a common thread. In all research works, the main interest was to re-frame the worker as an individual with unique attitudes, preferences and capabilities that, as such, need to have a central role in all HRIs. Indeed, we always measured human performance and workload, which are the human factors of major interest in the workplace. Most interestingly, we always leveraged a multi-method approach for better interpreting the human processes, offering a view on both implicit (e.g., eye parameters) and explicit (e.g., self-reported) dynamics that resulted in and from mental workload changes. Furthermore, we deepened different individual aspects in different experiments, namely gender, gaming experience, learnability skills, problem solving, trust in technology, and finally age.

Natural and embodied cognition approach. Another common strand is characterized by the attention toward the ecological validity of all conducted experiments. Specifically, we emphasized the strengths of real (and even virtual) world 3-dimensionality, always allowing a good degree of free motion and the most natural behaviors in all experimental environments, yet ensuring the proper methodological control. Such an approach follows the concept of natural or embodied cognition (Gallagher, 2006), according to which interactions between brain, body and environment are crucial because bodily experiences shape the way one processes the environment and vice versa. In this view, no matter which human behavior one is interested in investigating, it is essential to investigate it in natural exploration to the greater extent possible (Ladouce et al., 2017).

Intuitive and embodied teleoperations. Finally, it is worth mentioning that the interest in testing more natural and embodied teleoperations persisted throughout all studies. In fact, we aimed to understand to what extent human motion can serve as an intuitive control for robot teleoperations and whether it is preferable over more traditional teleoperation practices (such as button-based teleoperation). With this intent, a button-based control was compared with an action-based control in Studies 2 (Chapter 7), 3 (Chapter 8), and 4 (Chapter 9). Differently, in the last study (Chapter 10), a full natural and hands-free control system was tested, and embodiment dynamics were thoroughly investigated always in a multi-method fashion (i.e., performance, eye-tracking, EEG data).

1.3 An outlook on the thesis structure

Part I focuses on the theoretical framework that served as a foundation of the present thesis. Specifically, a review of various control systems in telerobotics is presented in Chapter 2, particularly delving into VR-based telerobotics, as it is of major interest to this thesis. Chapter 3 then addresses the topic of embodiment in robotic systems, particularly reviewing its main components (i.e., appearance, time and space) and unfolding the theories in support of a direct relation between embodiment, motor control, and teleoperation performance. Chapter 4, instead, addresses the most influential individual factors that were proven to affect users' performance and/or workload in telerobotics, namely age, gender, gaming experience and other secondary individual factors. Finally, Chapter 5 systematically presents the most common human factors metrics typically assessed in the human-robot interaction field, namely performance, workload, embodiment and motor control. Specifically, self-reports, behavioral, eye-tracking and EEG metrics are addressed.

Part II follows by unfolding all five studies that were conducted for the present Ph.D. thesis, namely: Study 1 (Chapter 6), Study 2 (Chapter 7), Study 3 (Chapter 8), Study 4 (Chapter 9), Study 5 (Chapter 10). Each chapter is structured by first presenting the hypotheses and research questions, then explaining the methods, and particularly unfolding characteristics of the sample, technical set-up, procedure, experimental task and design, measurements and statistical analysis. Results of each study are presented and discussed in the relative chapter as well.

Part III concludes with a general discussion (Chapter 11) briefly mentioning all obtained results, and particularly shading light on connections and parallelisms between them. The relevance of our results for the industrial and research sectors is highlighted.

Part I

THEORETICAL FRAMEWORK

Chapter 2

Control systems in telerobotics

Considering that this thesis specifically focuses on VR-based control systems for telerobotics, Section 2.1 briefly describes the types of interfaces that allow human-robot communication and interaction (i.e., 2-D, 3-D technologies and Mater-Slave architectures), while dwelling on VR-based interfaces for human-robot collaborative frameworks specifically. Furthermore, diverse interactive tools available for teleoperating robots are discussed in Section 2.2, from the most conventional ones (e.g., mouse, keyboards) to those allowing a higher level of interactivity (e.g., master-slaves, VR gloves), for finally mentioning Body-Machine Interfaces (BoMI), that allow driving robotic systems through the own body movements directly (e.g., motion capture and gesture recognition technology). Through the literature review, the main interest is directed toward user experience (i.e., sense of presence and individual preferences), and also toward human performance, workload and fatigue elicited by the diverse control systems for telerobotics.

2.1 Interfaces for telerobotics

A human-robot interface can be depicted as a set of elements through which communication and interactions between a human and a robot are enabled. Interfaces allowing to drive robots remotely need to convey an appropriate perception of the robot state, and to provide easy manipulation of the robot. Among the existing human-robot interfaces for telerobotics, there is the **Master-Slave architecture**, which implies the construction of a physical reproduction (Master) of a real robotic system (Slave). Humans can thus interact directly with the Master counterpart of the robot for actually driving the Slave. Among the digital interfaces for telerobotics, instead, the most conventional ones leverage traditional **2D technology**, such as desktop and tablet. Furthermore, there are also human-robot digital interfaces which leverage **3D technology** - such as VR - and allow higher immersion for the user who wants to drive a robotic system.

No matter the technology deployed for allowing teleoperations, the operator's awareness of the robot's environment, the quality of communication link, the robustness of robot's control system and experience of the human operator equally contribute to the effectiveness with which the operator controls a robot. As properly explained by Martín-Barrio et al. (2020):

"[...] a good interface would be clear, concise, familiar, responsive, consistent, aesthetic, efficient and forgiving. Binding such features will give place to an efficient interface while also providing high situational awareness".

2.1.1 VR-based telerobotics

One of the most significant innovations endorsed by the smart manufacturing is leveraging Virtual Reality (VR) for simulating robotic systems (Wonsick and Padir, 2020). Such applications fall within the 3D digital interfaces for human-robot interactions. The attractiveness of VR technologies applied to robotics is related to their flexibility and interactiveness, but also to economic factors. Indeed, they can reduce the cost of expensive specialized systems in manufacturing (Bugalia et al., 2015; Lipton et al., 2017). In industrial robotics, they have been exploited for different purposes, such as education and training (Abidi et al., 2019; Hormaza et al., 2019; Matsas and Vosniakos, 2017; Potticò and Lamberti, 2021), design of user interactions (Fratczak et al., 2019; Hansen et al., 2018; Kaufeld and Nickel, 2019), telerobotics and teleoperations (Linn et al., 2017).

Specifically on telerobotics, VR-based human–robot connections can be particularly useful for operating in physical locations that are inaccessible or that involve physical risks for the operator (e.g., underwater, underground, in nuclear facilities, aerial or space) (Szczurek et al., 2022; Xiao et al., 2020). Operators who are unable to perform inspection, maintenance or repair of robotic systems in person also benefit of telerobotics (Linn et al., 2017; Liu and Wang, 2020). Similar situations happened quite often during the latest pandemic, increasing the interest in remote control and VR also in the manufacturing field (Melluso et al., 2020). In this context, immersive VR devices are valuable mediums as they can supply information in a natural, interactive, and effective way. The feasibility of such telerobotic systems and the efficacy of the related framework have been repeatedly proven. For instance, Wang et al. (2019) designed a virtual teleoperation system for controlling an industrial collaborative robot with satisfactory tracking accuracy. Similarly, Lipton et al. (2017) successfully built a framework for teleoperating a Baxter robot via different control systems in VR. Those are just two examples of successfully implemented VR-based telerobotics frameworks, demonstrating that such technologies are concretely practicable from a technical perspective.

2.1.2 The impact of telepresence

Besides the most technical features of an interface, the user experience (UX) surely affects the quality of any teleoperation or interaction between humans and robots. Telepresence is one of the most relevant UX constructs in this respect. Such a concept emphasizes the possibility for human operators to feel a sense of being physically transported to remote work spaces (Minsky, 1980), and experiencing a feeling of presence at the remote side. Literature shows how the impact of telepresence on teleoperation performance is particularly (but not exclusively) important when teleoperating mobile robots (Opiyo et al., 2021). As outlined by Aracil et al. (2007), every human-robot interface for executing remote tasks should set the goal of evoking the maximum possible degree of telepresence as, in turn, this will increase the performance of the telerobotic system. In order to achieve such experience, information on the remote environment must be displayed to the operator in a natural way (Martín-Barrio et al., 2020) and, as a consequence, high telepresence and situational awareness will be experienced. *Situational awareness* is indeed another key concept for telerobotics. Its definition implies a high time and space perception of elements in the environment, the comprehension of their meaning and the projection of their status in the near future (Endsley, 1988).

When specifically entering VR-based research fields, most researchers also emphasize the concept of **sense of presence** (Lee, 2004), which was instead defined as "*a psychological state in which virtual objects are experienced as actual objects in either sensory or nonsensory ways*". VR, as a fully immersive technology, can promote a feeling of actually being there at the remote side of the teleoperation. This is a point in favor of the adoption of VR technology for robotic teleoperations, which was further supported via systematic research. For instance, Martín-Barrio et al. (2020) specifically assessed whether immersive (3D) interfaces are better compared to conventional ones (2D) to teleoperate hyper-redundant robots. Their hypothesis was supported by their data: the immersive interface was more efficient, it was highly preferred by the user and also led to higher situational awareness compared to the more conventional one. Macchini et al. (2021a) measured the sense of presence as self-reported by individuals driving a non-antropomorphic drone; VR correlated with a higher sense of presence. Interestingly, this was observed only when using VR head-mounted displays, but not other immersive technologies such as a CAVE (James et al., 2011).

Related to that, it is to acknowledge that factors such as limited field of view (FOV), orientation, camera viewpoint, depth perception, degraded video image and motion can significantly impact the teleoperation performance (Martín-Barrio et al., 2020). All such issues are related to the quality of the computer graphics and the immersiveness of the technology involved, and as such, can be overcome by leveraging 3D interfaces for teleoperations. Additionally, issues related to the motion of the robot to teleoperate are also surmountable through VR. Indeed, 3D interfaces allow higher freedom of physical movements and more natural interactions compared to 2D-based human-robot interfaces (e.g., mouse, keyboard). All these points are in favor of leveraging VR for telerobotics, and might contribute to the observation of higher telepresence in 3D compared to 2D environments.

2.2 Interactive tools for telerobotics

The tools that put in contact humans with robots can vary for the degree of interactivity they allow, independently from the level of immersivity of the human-robot interface. Indeed, we here categorized them in **Low Interactivity Control Systems** (LICS), which are the most conventional ones and usually allow indirect HRIs, and **High Interactivity Control Systems** (HICS), which often allow more direct and physical HRIs.

2.2.1 LICS: Low Interactivity Control Systems

While both HICS and LICS allow to interact with a robotic system, LICS keep the interaction more 'indirect'. For instance, keyboard, mouse and joysticks fall into this category. They are the most conventional interaction tools and typically require the user to learn an association between an input (e.g., ctrl+left key button) and an associated response on a robotic system (e.g., the robot moves left). Those interactive tools can be used both in immersive and non-immersive teleoperation environments.

2.2.2 HICS: High Interactivity Control Systems

Differently, all those control systems allowing direct and often physical interactions with a robot are here considered as HICSs. An example is given by Master-Slave frameworks, which consist of controlling a teleoperated robot (slave) through the direct manual manipulation of a physical replica of the robot (master). In such cases, the interaction between the master and the slave is direct and tangible. Also VR allows to manipulate robots through direct interactions. There are kinesthetic or 3D haptic devices that provide force feedback to the users and allow the manipulation of 3D objects in virtual environments. Those solutions for telerobotics allow to literally *get in touch* (in a direct way) rather then to *get connected* (in indirect ways) with a robotic system, and are thus here considered as HICS.

A special case of highly interactive control systems for robotics is constituted by Body-Machine Interfaces (BoMIs). Interactive tools such as *Leap Motion* or other motion capture technologies are prominent in this respect, as they allow to locate the position of a human body, and potentially transpose it into a robotic system. In this way, humans can guide robots in a hands-free fashion, only leveraging their own physical movements. Therefore, they will only have to learn associations between a movement of their own body (e.g., hand up) and an associated response on the robotic system (e.g., the robot moves up). Remarkably, MoBI-based control systems can benefit from natural and embodied mechanisms linked to concepts sush are sense of ownership, agency and self-location, that can significantly influence teeoperation performance. These topic will be further addressed addressed in Chapter 3.

2.2.3 Comparative literature review

Among the studies that directly assessed the effects of different degrees of teleoperation system interactivity on users' performance, many evidenced the advantage of HICSs over LICSs. For example, Vozar (2013), in his own Ph.D. thesis, tested a master-slave in contrast to joystick teleoperation. Participants drove a customized skid-steer robot through a delimited space to reach some boxes and then teleoperated the robot's arm to grab and move some boxes. Results revealed that teleoperating the robot via master-slave significantly improved performance when moving the boxes, it reduced the total time of the task and was rated as more intuitive and easier to use compared to the joystick. Gliesche et al. (2020) compared teleoperation performance in a sample of nurses using a haptic device or a keyboard and mouse for guiding a 7-degree-of-freedom robot manipulator through a desktop-based pick-and-place task. When using the haptic device rather than mouse and keyboard, participants showed shorter task completion times.

A few studies also demonstrated how HICSs outperformed over LICSs in VR. For example, Franzluebbers and Johnson (2019) assessed the performance of users teleoperating a pair of 7-degree-of-freedom robotic arms through a pick-and-place task in VR via two control systems: a stationary 3D mouse and via the VR controllers that tracked participants' movements. Faster execution times were gained when using the VR controller rather than the 3D mouse. Similarly, Martín-Barrio et al. (2020) compared different control systems for teleoperating a robot in VR, including controller, Master-Slave, and physical gestures. Particularly, users maneuvered the robotic arm Kyma with high degrees of freedom to some predetermined positions. Physical gestures were preferred over Master-Slave and controller, and they additionally allowed higher accuracy and faster operation times compared to the controller. On the other hand, in the same study, participants self-reported lower values of workload when using a controller and direct manipulation compared to the master-slave modality.

Besides, there are also studies that did not show an advantage of HICSs over LICSs. For instance, in Rouanet et al. (2009), participants had to teleoperate a zoomorphic robot through a domestic environment to find an object. They were instructed to use three control systems: touchscreen-based buttons, a virtual keyboard on a 2D screen, and arm movements tracked by a hand-held controller. The authors did not observe any performance differences between the input modalities; however, participants preferred to use the touchscreen-based input modality over the other two. In the experiment of Grabowski et al. (2021), instead, participants had to drive a mobile robot equipped with two arms through a pick-and-place task in an immersive large-scale virtual environment. In one condition, they teleoperated the robot via VR controller buttons, while in the other condition they had to physically walk to the target position. In both cases, participants teleoperated the robotic arms by moving their own arms. Results demonstrated a better performance in task completion time and accuracy when using the VR controllers rather than when active walking.

More broadly, it was suggested that human gestures and direct robot manipulations can be advantageous because they are easy to use, robust, fast and can be used in a wide range within the FOV (Hu et al., 2003). Similarly, it was emphasized how they could be a valid alternative to physical master–slave architectures since they can be cheaper and more intuitive (Shirwalkar et al., 2013). Studies assessing users' performance when teleoperating robots via control systems allowing higher or lower interactivity show a general tendency for better performance when using HICSs, even though there is some evidence in contrast with this trend (Rouanet et al., 2009; Grabowski et al., 2021). Differently, literature on the effects of the degree of interactivity on users' workload is still scarce.

Chapter 3

Embodied telerobotics

When interacting with a robot through signals derived from the users' body, we can talk about Body-Machine Interfaces (BoMI). Literature show how they have been extensively employed as application devices in assistive and rehabilitation sectors (Casadio et al., 2012), and in few cases also in telerobotics (Macchini et al., 2021a,b). By leveraging intuitive body control and gestures, these human-robot interfaces appear as imperceptible and almost non-existent to the operator. This characteristic is usually called *transparency*, and it is what makes BoMIs a promising alternative to standard control devices for telerobotics. Specifically, increased transparency is achieved when humans do not even notice that the teleoperation is mediated through a device (Toet et al., 2020). In these situations, the illusory experience that the robot becomes part of the own body/arm/hand will occur, generating a *sense of embodiment*.

This Chapter specifically focus on mechanisms of embodiment applied to robotics and telerobotics. Specifically, Section 3.1 depicts the state of the art, starting from a quick look into the theory of embodiment, then illustrating the single roles of robot appearance (Subsection 3.1.1), temporal (Subsection 3.1.2) and spatial (Subsection 3.1.3) consistencies between the robot body and the human body in generating embodiment, and finally exposing a comparative literature review (Subsection 3.1.4). Thereafter, in Section 3.2, how human-robot interfaces generating embodiment could be advantageous for telerobotics is discussed, and open questions are presented.

3.1 Literature overview

Writ large, the sense of embodiment comprises several underlying subcomponents (Kilteni et al., 2012b) including: **sense of ownership**, which is the feeling that "an object or a body is part of my body"; **sense of agency**, as the feeling that "I can exert control on an object or a body"; **sense of self-location**, for which one can feel that "I am located within an object or a body". De Vignemont (2011) further operationalized sense of ownership, agency and self-location respectively in **affective embodiment** (an individual shows the same affective reactions for the external object as for the own body), **motor embodiment** (the motor system takes the properties of the object/tool as properties of the effector in planning) and **spatial embodiment**, which involves a bodily frame (an object is taken into account within the body's space boundaries), external frame (the localization of a tool/object is processed in the same way as the localization of a part of one's body) and peripersonal frame (an object/tool is processed as peripersonal space).

Pioneering evidence on embodiment is offered by the rubber hand illusion, in which an observer experiences body ownership of a rubber arm and hand when it is stroked simultaneously with the hidden own arm and hand (Botvinick and Cohen, 1998). From that observation, tons of experimental studies investigated mechanisms of embodiment, showing that humans can feel a sense of embodiment for different extracorporeal objects and can integrate them into their body schema (Aymerich-Franch et al., 2017; Schettler et al., 2019; Toet et al., 2020). All published evidence of embodiment agree on one point: a feeling of embodiment can be elicited via **congruent multisensory stimulation**. Specifically, literature shows how the possibility of feeling embodied into an external object (in this case, a robot) is significantly influenced by:

- **Appearance**: physical consistency of the robot and the own's body
- **Time**: visuomotor/visotactile temporal synchrony
- **Space**: spatial co-location of the robot and the own body

However, none of these components seems to be strictly necessary for producing a general sense of embodiment. Indeed, there is evidence of humans embodying external objects/robots/avatars even in absence of visuomotor synchrony (Aymerich-Franch et al., 2016), also when there is no spatial co-location (Miura et al., 2021; Pritchard et al., 2016), and even for non-human-looking objects/tools (Aymerich-Franch, 2012; Aymerich-Franch et al., 2017).

3.1.1 Appearance: physical human-robot consistency

As systematically reviewed by Aymerich-Franch and Ganesh (2016), while a drastic reduction in the illusion occurs when a non-corporeal object is used, it has been demonstrated through both real-world- and VR-based experiments that bodies with extra limbs, fake limbs, robotic hands or arms, mannequins, virtual bodies, and even empty volumes of spaces or invisible bodies can be embodied. One example coming from the robotic domain shows that humans are able to experience a strong sense of embodiment towards a humanoid robot with a non-human-looking metal arm, even when it was covered by a blue plastic cover (Aymerich-Franch et al., 2017). Therefore, there are good chances that a human can also embody a non-anthropomorphic robot, like most of the commercial industrial robots. Nonetheless, when human appearance of the external object to embody is disrupted - like in the case of a non-anthropomorphic industrial robot - it is possible that the weight of visuomotor synchrony and co-location in inducing embodiment increases significantly, or even that they become strictly necessary for provoking a sense of embodiment.

3.1.2 Time: temporal human-robot synchrony

Proper synchrony between the own movement and the provided visual feedback (namely, visuomotor synchronization or correlation) is essential as it is the basis of the sense of agency, and thus motor embodiment. Remarkably, while embodiment is broken to the same extent by incongruities in either visuomotor or visuotactile stimulations (typically used in the rubber hand illusion), synchronous visuomotor stimulations contribute the greatest in establishing general embodiment (Aymerich-Franch and Ganesh, 2016; Kokkinara and Slater, 2014) and lead to a more spatially spread proprioceptive drift (Tsakiris et al., 2006). Some examples of visuomotor synchronization increasing the sense of embodiment can be found in (Aymerich-Franch, 2012; D'Angelo et al., 2018; Farizon et al., 2021; Kokkinara and Slater, 2014; Ratcliffe and Newport, 2017; Sanchez-Vives et al., 2010; Ventre-Dominey et al., 2019). For mentioning an instance, simply synchronizing human movements with object movements allowed to embody non-anthropomorphic 3D shapes in VR (Aymerich-Franch, 2012). Furthermore, in a forearm bisection task, synchronous but not asynchronous visuomotor condition (3-sec delay between the physical and the virtual hand's movements) led to a plastic change of the peri-personal space and the body schema (D'Angelo et al., 2018). Interestingly, Argelaguet et al. (2016) reported a stronger sense of agency for less realistic virtual hands as they provided less mismatch between the participant's movements and the animation of a virtual hand compared to more realistic virtual hands. The authors thus underline that there is no specific need to provide realistic hand looking to induce a sense of agency. Moreover, it seems that the potential of visuomotor synchronization in producing a sense of embodiment (or at least agency) actually exceeds the influence of humanoid looking, which was also confirmed by later research work (Ratcliffe and Newport, 2017).

However, there is evidence that visuomotor synchronization is not essential for stimulating not agency nor a sense of embodiment, particularly when embodying a humanoid robot (Aymerich-Franch et al., 2016). Indeed, participants reported to feel agency when controlling a humanoid robot via joystick (thus lacking visuomotor coupling), and even when the robot's movement was not under direct control of the participant. The authors (Aymerich-Franch et al., 2016) argued that "*even though agency was seemingly not necessary to experience whole body embodiment, we still believe that agency can benefit the sense of embodiment*" (p.106). Also, it is to notice that this evidence comes from an experimental context employing a robot with human semblances, which are known to easily stimulate embodiment as compared to non-anthropomorphic robots or objects (Aymerich-Franch, 2012).

3.1.3 Space: spatial human-robot co-location

When embodying an external object, the sense of self-location is given by the feeling of being located within an object or a body. In other words, it is the feeling that the external object and the own body part are located in the same position. This is particularly important for patients embodying prosthesis and artificial limbs: phantom limbs that are experienced as part of the bodily self can affect prosthesis embodiment only if the phantom and the prosthesis are brought into perceived co-location (Bekrater-Bodmann, 2022). The mechanism underlying the sense of self-location is the proprioceptive drift: as suggested by (Kammers et al., 2009), human proprioception can drift rapidly, leading to overwrite the proprioceptive location information of one's own body part with the visual location information of a virtual object. In this sense, human proprioception is plastic as it adapts to the location of external objects, even when they are dislocated in space.

While some studies suggested that spatial co-location of an external limb and the own limb is a necessary component for the rubber hand illusion (Pavani et al., 2000; Tsakiris et al., 2006), others demonstrated that it is influential but not decisive in inducing a sense of embodiment. For instance, among those claiming that spatial co-location is a necessary element to elicit embodiment, Tsakiris et al. (2006) demonstrated that effects of synchronous vs. asynchronous visuotactile stimulation in a rubber hand illusion paradigm were only present when the rubber hand was shown in a congruent position as the own hand. On the same line, Preston and Newport (2011) demonstrated that competing limb representations that are closer to the own limb are preferred to those that are far from the body. Also when embodying a virtual avatar, the physical co-location between the virtual and the real body (first-person perspective) is known to elicit a stronger sense of self-location than non-collocated perspectives (third-person perspective) (Slater et al., 2010; Petkova et al., 2011). Moreover, it has been reported that the sense of ownership is sensitive to both visual appearance and spatial location of one's own body (Ratcliffe and Newport, 2017).

Differently, among those studies claiming that spatial co-location is not necessary to elicit embodiment, Pritchard et al. (2016) demonstrated that the spatial co-location is only crucial for inducing a sense of self-location. Furthermore, evidence of being embodied in dislocated objects is reported by Newport and Preston (2011), who demonstrated that participants switched ownership in a dislocated hand representation under synchronous stimulation when the co-located hand was stimulated asynchronously. Similarly, Miura et al. (2021) showed that humans can feel embodied in multiple bodies when controlling their movements simultaneously, reporting particularly sense of body ownership and agency - a phenomenon called "distributed embodiment".

3.1.4 Comparative literature

To mention some of the studies that directly compared effects of visual, motion, spatial and/or temporal disruptions, Tsakiris et al. (2006) conducted a series of experiments on the rubber hand illusion demonstrating that differences between synchronous and asynchronous visuotactile conditions were significant only when the rubber hand was in co-location with the human hand. Moreover, significant proprioceptive drifts occurred only when participants watched a visually congruent rubber hand (i.e., left-left) and judged the position of the actually stimulated finger. Newport and colleagues manipulated both congruency of visuotactile stimulations and spatial location of two displayed hands (Newport et al., 2010; Newport and Preston, 2011). They found that ownership and reaching movements were consistent with embodiment of the synchronous hand (Newport et al., 2010), and that participants disowned the hand in the correct spatial location when feedback was asynchronous while taking ownership over the spatially dislocated synchronous hand (Newport and Preston, 2011). In a similar experimental context, (Ratcliffe and Newport, 2017) explored the influence of spatial location, visuomotor synchrony and visual appearance on each embodiment component. They demonstrated that agency was felt only when there was visuomotor synchrony, regardless of visual appearance or spatial location; ownership was reported for synchronous visuomotor conditions regardless of the spatial location, but was strongly influenced by the visual appearance (no ownership was reported if both visual appearance and spatial location were altered); finally, self-location was reported only in case of visuomotor synchrony and spatial co-location (while in case of co-location but visuomotor asynchrony, self-location was reported to be not in the synchronous nor in the asynchronous hand). Seen from a different perspective, visuomotor synchrony was thus crucial for all embodiment components (agency, ownership and self-location), spatial co-location was only critical for self-location and visual appearance for the sense of ownership. In a different study, (Pritchard et al., 2016) observed that both visual appearance of a virtual hand and visuotactile synchrony affected all embodiment subscales, while spatial co-location only influenced the sense of self-location. In a later research work (Kim et al., 2020), embodiment in a human avatar, point-light avatar and out-body point-light was induced through synchronous visuomotion. Interestingly, while the reported sense of agency did not differ between the three avatars, the point-light avatars obtained lower rates of ownership and self-location as compared to the human avatar. These findings corroborate the hypothesis that visual appearance has no influence on agency, but it can be crucial for the other embodiment subcomponents.

3.2 Embodiment in telerobotics

As previously mentioned, human-robot interfaces leveraging intuitive body motion and gestures, such as BoMIs, are likely to trigger mechanisms of embodiment. In these cases, users might forget about the mediator (interface) and act more naturally, thus increasing the transparency of the teleoperation. Furthermore, they are also more likely to feel present at the remote side, and might even change the way they perceive their body in relation to the robot to teleoperate. This translates into higher telepresence (perceived relation between one's self and the remote workspace) and altered body ownership (perceived relation between one's self and one's bodily representation) (Kilteni et al., 2012a). All such mechanisms will most likely improve teleoperation performance (Toet et al., 2020), and therefore, are definitely worth of investigation.

The link between embodiment and telerobotics was clearly proposed by Toet et al. (2020), who reviewed a variety of studies on embodiment and argued that there is experimental evidence in support of the following premises:

1. *"the representation of the body in the brain is malleable and can include non-bodily objects like robotic hands and end effectors"*

2. *"embodiment can be elicitepd through mediated sensorimotor interaction"*

3. *"once established, embodiment is robust against inconsistencies between the robotic system and the operator's body"*

4. *"the strength and robustness of embodiment correlate positively with dexterous task performance"*

In relation to points 1, 2 and 3, as addressed in Section 3.1, literature suggests that if a robot (i.e., a robotic arm) has human form, if it is located in the same place of a human's body part (i.e., the human arm), and if the user feels to can exert control on such robot through the own body (i.e., arm) movements, this robot will likely be embodied into the human's body schema. Specifically in contexts involving non anthropomorphic robots, like in the industrial sector, the first point is thus not met. Therefore, a sense of embodiment can be mainly conveyed by visuomotor synchronization (Aymerich-Franch, 2012; Farizon et al., 2021; Kokkinara and Slater, 2014; Singh et al., 2018) and spatial co-location (Slater et al., 2010; Petkova et al., 2011; Ratcliffe and Newport, 2017; Tsakiris et al., 2006). However, their relation and influence on the feeling of embodiment (and specifically on sense of ownership, agency and self-location) is still understudied, particularly when the human body representation is disrupted (i.e., non-anthropomorphic robots). Furthermore, it needs to be accounted that technical and transmission problems of the human-robot interface, or issues related the communication channel, might occur during any teleoperation. In such cases, the visuomotor synchronization and/or the spatial co-location of the human arm and the robotic arm might be lost, delayed or out of sync, reducing the quality of the interaction (Niemeyer et al., 2016) and, as a consequence, possibly affecting the task performance as well (in terms of speed and/or accuracy).

Additionally, in relation to point 4, there are researches advancing the hypothesis that higher embodiment is related to a better teleoperation performance. However, to the best of our knowledge, proper evidence from telerobotics that support this hypothesis is still missing. One recent study conducted by Iwasaki et al. (2022) specifically addressed the question of how embodying a limb into a robotic limb modulate the attention distribution across limbs. They used a dual-task to test their hypothesis, whereby participants executed a cued button-press with their right hand, while reacting to possible collisions by a moving object with a left robotic hand. Their results demonstrated that embodiment of a robotic limb improves attention allotment for dual-task performance conducted with the robotic hand itself. This is an incredibly important step toward the establishment of a direct relation between embodiment, attention and task performance, which could impact significantly the teleoperation industry. Nonetheless, it is just a first step, and more research is needed, particularly in telerobotics.

Chapter 4

Individual factors in telerobotics

The human-centric approach suggests that machines should be tailored on human needs and capabilities. In this view, it is thus essential to understand how differently skilled or aged users, as well as people with different attitudes towards technology, might respond to the ongoing digitalization proposed by the recent Industry 4.0 and 5.0. In this Chapter, we address these questions by reviewing some of the most well-known individual factors that are likely to impact HRIs and VR-based teleoperations, namely age (Section 4.1), gender (Section 4.2), gaming experience (Section 4.3) and other individual skills and attitudes towards technology (Section 4.4)

4.1 Age

According to the literature on the aging workforce, there is a considerable decrease in an individual's functional capacity following the age of 50 (Di Pasquale et al., 2020; Donato et al., 2003): cardiovascular, respiratory, metabolic and musculoskeletal functional declines are prone to manifest gradually in both men and women (Ilmarinen, 2001). The deterioration of such functions has a strong impact on **physical** labor capacities.

Regarding functional **mental** capacities, a large portion of the age-related losses in cognitive ability may be explained by a decline in the so-called executive functions, namely processing speed, working memory capacity (Verhaeghen and Salthouse, 1997), and the ability to inhibit irrelevant and distracting stimuli (Sweeney et al., 2001). When these functions decline, an individual's ability to manage mental workload can be compromised. These aspects are particularly relevant in working contexts as well: indeed, mental workload is deeply connected with performance (Steinfeld et al., 2006), and its increase can potentially lead to an increased risk of work-related accidents (Hughes et al., 2019).

Aside from the normal aging process, work-related factors may also play a role in deteriorating workers' health and physical state. In particular, assembly and manufacturing jobs that require elderly workers to perform nonstandard work schedules (i.e., evening or night shifts, irregular hours and weekend work) and physically demanding tasks for extended periods create concerns for their physical health (Peeters and van Emmerik, 2008).

Senior workers who are susceptible to physical stress, like those on assembly lines, thus need particular attention in the fourth industrial wave. Indeed, the digitalization of many industrial processes and systems, as well as the virtualization of teleoperations, could be surely beneficial for this population. For instance, operating in VR instead of operating on physical robots and machines keeps users from the risk of physical accidents and minimizes their physical effort. Furthermore, being informed on the user's level of workload throughout the work shift, as inferred via eye parameters, can help optimize working tasks and conditions (e.g., reducing working speed, suggesting breaks, etc.). However, it needs to be considered that aging not only brings cognitive and physical decay (Brough et al., 2011; Kenny et al., 2008), but it is also typically related to decreasing flexibility and adaptation to new technologies (Di Pasquale et al., 2020), which may play a counterposing role in the manufacturing virtual revolution. In this view, it is important to systematically assess: i) whether and how senior users can adapt to such new interactive technologies; ii) if they can perform well enough as compared to younger users; iii) whether their workload capacity can withstand repetitive VR teleoperation tasks.

4.1.1 Ageing, interactive technologies and VR

Nowadays, technology can be accessed more directly, naturally, and intuitively through novel devices (e.g., BoMIs), which allow users to interact with smart tools using joint movements and gestures. However, older adults may encounter difficulties adopting such interactive modalities due to their lack of technological literacy and mobility impairments (Ketcham et al., 2002; Verrel et al., 2012). In this view, a number of studies have assessed whether new interactivity methods are feasible for older users and if they can supplant the most traditional ones (i.e., mouse, keyboards, and controllers).

Research testing gesture recognition systems demonstrated that action-based technologies seem to accommodate individuals of all ages, including older ones with hand impairments such as tremors and arthritis (Teimourikia et al., 2014). However, when evaluating the performances, older adults usually show more difficulties than younger ones. For instance, Bobeth et al. (2014) compared tablet touch interface, hand gestures and remote controllers for smart TV applications in older and younger adults. All three input modalities, and particularly the touchscreen tablet, were rated as very easy to use by both age groups. The elderly, however, were slower and less accurate than the young users in all conditions, especially when using a remote control. A similar example was provided by Gerling et al. (2013), who compared sedentary and motion-based game controls (i.e., pointing, steering and tracking tasks) over older adults and younger participants. They did not observe any age-related differences in device comfort or enjoyment; nonetheless, older individuals performed worse than younger ones when using the motion-based game control. Besides, young outperformed older adults in all tasks except for the steering task, suggesting a task-related impact of age on performance.

Beyond the aforementioned technologies, immersive VR brings interaction technologies to another layer. It offers the opportunity to adopt different paradigms of interactivity (Martín-Barrio et al., 2020), which could make it adaptable to different categories of users, such as the elderly, and also make it appealing to industry and manufacturing sectors (Firu et al., 2021). Recent literature provides mixed results regarding the feasibility and effectiveness of implementing VR with the elderly. Furthermore, only a few studies directly investigated age-related differences in work contexts that leverage VR, like for example (Adami et al., 2021).

Some virtuous findings were reported by Syed-Abdul et al. (2019), who evaluated the domestic use of VR: they demonstrated that the elderly are willing to use VR if it is perceived as useful, easy to use, and comfortable while using it. In a spatial navigation memory assessment performed in immersive and non-immersive VR, older individuals demonstrated higher performance, a higher sense of presence, better assessment feasibility, and a lower stress level in immersive VR (Ijaz et al., 2019). Researchers emphasized the importance of lower stress levels, particularly for the elderly, as they can increase their confidence in task performance, thereby increasing their intention to use VR.

On the other hand, some age-related difficulties in VR usage were also shown. For instance, Chen and Or (2017) compared the use of an immersive virtual reality CAVE projection system, mouse, and touchscreen in pointing and drag-and-drop tasks on three groups of young, middle-aged, and older individuals. Results demonstrated that, regardless of the task, groups of all ages were slower and made more errors in the virtual environment. Furthermore, regardless of the interface, older adults performed significantly worse compared to the other groups. The authors suggested that the lower VR performance might be related to the reduced individual experience with VR compared to the other interaction technologies. Among the few studies that tested users' performance in work environments, instead, Adami et al. (2021) evaluated the effectiveness of VR-based compared to real-setting training for learning how to teleoperate robots in construction settings. They collected qualitative data and accounted for age as a moderator variable. Results showed that age reduced the effectiveness of VR training. In particular, older participants showed fewer benefits from VR training in terms of acquired knowledge and prevention of risk behavior.

4.1.2 Age effects on work and teleoperation performance

Working in industrial production often entails repetitive and prolonged activities, which usually challenge workers' attentional and physical abilities. This is particularly true for older workers, who are known to face mental and physical deterioration (Claudon et al., 2020). Nonetheless, while several reviews have examined aging and productivity at work, there is still a lack of strong empirical evidence to support the conclusion that age affects performance in productivity systems (Boenzi et al., 2015). Older workers showed an unaltered ability to perform a repetitive manual task consisting of different phases (i.e., moving, collecting, assembly) under varying time constraints. However, they needed to strive harder to adapt to this type of task, resulting in a greater risk of developing musculoskeletal disorders than younger workers (Gilles et al., 2017). Similarly, a kinesiology-based study tested differently aged workers during a simulated repetitive assembly task that elicitepd muscle fatigue in the shoulder region (Qin et al., 2014). Young and older individuals showed similar kinematic and kinetic responses to the task requirements and as a consequence comparable performance. However, older workers also experienced more effort to maintain the same motor strategies as younger workers. A similar trade-off between performance and effort also applies to laboratory-based vigilance tasks that recall workplace systems monitoring situations (Bunce and Sisa, 2002). Specifically, when participants responded to a target stimulus while ignoring distracting ones, the performance did not differ between age groups, but older individuals self-reported higher workloads compared to younger ones.

In the telerobotics sector, literature unfolding age effects on performance mainly focused on human care tele-applications (Pang et al., 2021; Pavon-Pulido et al., 2015), whereas, to the best of our knowledge, only one study examined aged users in robotics teleoperations also involving VR (Grabowski et al., 2021). In the latter, researchers compared different control interfaces (i.e., walking, controller buttons) for driving a mobile robot equipped with two robotic arms in a walking warehouse context. The task consisted in steering the robot into the physical space for picking and placing objects in specific positions. In the controller buttons condition, both young and old individuals were assessed, while the walking condition was performed only by the young group. Furthermore, half of the young participants underwent a training phase in VR, while the other half did not perform any training, with the intent of also evaluating the effectiveness of VR training on teleoperation performance. All older participants performed the VR training before the actual task, but none performed the task without the training; therefore, conclusions on the effects of VR training did not extend to the aged group. When comparing young and older participants, age significantly affected performance (i.e., time on task and correct trials rate) and task load (i.e., NASA-TLX scales, (Hart, 2006)). Specifically, older participants reported higher effort, frustration, performance and mental demand but not physical or temporal demand compared to young participants. Differently, the level of stress measured via the DSSQ questionnaire (Matthews et al., 1999) was not affected by age. Overall, older individuals demonstrated greater difficulty and fatigue when driving a robot via controller buttons compared to younger individuals. However, as the old group did not perform the task in the walking condition, it is unclear whether more natural control interfaces can benefit the older population.

4.1.3 Age, workload and dual-task

In applied research, it is common practice to use dual-task methodology for manipulating the imposed demand in a controlled way (Sweller et al., 2011). The dual-task, which is the simultaneous performance of two activities, increases mental workload, creates task interference and leads to performance degradation in one or both tasks (Leone et al., 2017). Given the age-related functional decay, literature studying dual-task differences between young and older agents is vast. The meta-analysis of (Verhaeghen and Salthouse, 1997) revealed that, when concurrently executing multiple mental tasks, older people tend to slow down their performance in the face of reasonable task accuracy compared to younger people. Indeed, the performance seems to worsen (i.e., time on task and accuracy) when moving from single- to dual-task for all ages, independently from the type of mental tasks and in an additive way. Put differently, no specific age deficits were associated with the observed dual-task accuracy cost when overlapping multiple mental tasks. Different results were generally observed when the dual-task involved a cognitive and a motor task. In this situation, cognitive-motor interference is likely to occur, which leads to performance deterioration in one or both tasks. As age-related cognitive decline is usually accompanied by motor deterioration, cognitive-motor interference is generally higher in older than younger agents (Leone et al., 2017).

Broadly speaking, this phenomenon attracts particular attention in workers' safety research as it potentially leads to falls or accidents (Habibnezhad et al., 2020). Indeed, all teleoperation activities usually involve a wide range of sub-tasks (e.g., paying attention to obstacles, remotely maneuvering the robot components, etc.) on top of the general monitoring of the task progression (e.g., following the task's sequence, checking for potential errors, etc.). In this view, it is crucial to understand to what extent senior workers' workload capacity can withstand the coordination of cognitive and motor aspects in teleoperation tasks, particularly when they involve physical activity (such as with BoMIs). This information can help comprehend how to adapt the new technologies proposed by Industry 4.0 and 5.0 to different users, therefore improving working conditions for all ages.

4.2 Gender

Literature on HRIs often suggests gender as an influential individual factor that can impact the interaction with the robot. Generally speaking, it is known how males and females seem to use different processing strategies: males use to segment a task discretely, while females tend to configure a task as a whole (Darley and Smith, 1995). Furthermore, many research works demonstrate how spatial abilities change with gender, with men usually outperforming women (Menchaca Brandan, 2007). Literature also evidenced how males tend to tinker or play more with tools, which usually facilitates scientific understanding (Lamers et al., 2013). As all these factors also play a role when considering humans interacting with robots, gender needs to be considered when evaluating HRI design and solutions.

A large proportion of research on gender and HRI specifically focused on gender (usually also the robot's gender) as a social factor affecting the human-robot interplay. While these researches are important for social robotics, they exceed the interests of the present thesis and, therefore, will be just quickly mentioned. For instance, the study of Nomura and Takagi (2011) demonstrated how participants' gender and educational background affect the perception of a humanoid-shaped robot's politeness, mildness, ambitiousness and assertiveness. Participants were placed in front of a robot and were briefly introduced to it. The robot was genderized by the experimenter, who called it with male or female names. Thereafter, they just observed the robot doing some tasks automatically, and then filled some questionnaires. Results demonstrated that, besides the effects of the robot's gender, female participants reported higher levels of politeness than male participants. Furthermore, males had higher positive effects of perceived femininity and assertiveness impressions for the robot compared to females. Mutlu et al. (2006), in their experimental study, underlined how adapting HRIs to use attributes (in particular gender) and task structure (especially a competitive vs. a cooperative structure) are key design elements. They used specifically a social humanoid robot and examined the impact of users' gender on social interactions with the robot. Their interactive experience was designed in a way that the participant and the robot could act as peers, in one case generating competitiveness, in another case cooperation. Men reported their positive affect significantly higher in the competitive compared to the cooperative situation only when interacting with the robot, while they demonstrated the opposite trend when interacting with another human. Differently, the positive affect of women did not differ across the task situations. Therefore, it seems that both people's gender and task structure affect their social experience with a humanoid robot. These are just a few examples of research focusing on gender, social behaviors, and social acceptance of the robot.

On the other hand, among the research evidencing whether and how the human gender impacts the interaction with the robot, Showkat and Grimm (2018) evaluated the influence of humans' gender in information processing strategy, self-efficacy and exploratory behavior when interacting with a physical humanoid robot via joystick. They highlighted how the information on how to use the robot was processed in a comprehensive way by the female participants, while males processed them more selectively. Furthermore, as compared to females, males were more confident when using the robot and tinkered more with it, and also demonstrated greater task success and lower task completion time. The latter result was likely due to the males' tendency to tinker more with the robot. Similar results were observed when teleoperating a commercially available six degrees-of-freedom robotic arm via desktop and mouse (Paperno et al., 2019). Participants were asked to perform a series of tasks with a robot, including a pick-and-place task. Researchers measured reaction times and the following abilities: taking egocentric or allocentric views, mentally transforming objects in space, interpreting the surrounding environment, perceiving the distance of an object, manual dexterity, visual acuity, and working memory abilities. Men and women significantly differed specifically in manual dexterity abilities. Furthermore, males performed the tasks quicker and with fewer moves than females. Menchaca Brandan (2007) tested participants performing space teleoperation and analyzed gender effects on their performance. Participants were instructed to pick up a payload and dock it in a specific position while avoiding collisions and remaining within the working area. Women required significantly longer times than men to complete both the pickup and the dock phases. They also had lower orientation docking accuracy compared to men. In this case, women also obtained lower scores on the spatial tests, which could be the reason for their poor teleoperation performance. In the study of Chan et al. (2019), a tracked vehicle was guided through a joystick. Participants had to guide it through straight-line paths, and also steer it through corners of different geometries. The number of collisions was related to the task difficulty only in females, who also tended to require more practice to feel comfortable that they could adequately control the vehicle. Furthermore, males, after the initial practice session, showed clear learning on straight-line tasks, whereas females did not. Females also showed small learning (as a function of six consequent trials) whereas males did not. Differently, Dybvik et al. (2021) conducted a principal component analysis within a predictive display framework for teleoperation. They explored whether explored if there were any interesting patterns or observations in the data without any previous hypothesis. They found interesting relations between gender and gaming experience, which may have impacted the teleoperation performance as a function of the subtle spatial abilities. This discussion will be further exploited in the next section 4.3.

4.3 Gaming experience

Another critical factor potentially affecting performance in HRIs is the experience with video games. Gamers are known to benefit from their gaming practice for their increased spatial abilities (Green and Bavelier, 2003; Dye et al., 2009), handiness with joysticks and visual interfaces and, in turn, teleoperations (Gomer and Pagano, 2011; Chuan et al., 2007; Brizzi et al., 2017). Human performance in robot teleoperation is indeed strongly related to both spatial skills and gaming experience. A good example in this regard is given by Gomer and Pagano (2011), who asked participants to guide one of four types of robots (namely, a stationary arm, a mobile 6-wheeled vehicle with a robotic arm and claw, a tracked vehicle, and a wheeled vehicle) through tasks at high and low difficulties via desktop. They measured time on task and the number of collisions during the teleoperation. People with higher spatial abilities demonstrated shorter time on task and fewer collisions. Interestingly, gaming experience seemed to help participants with lower spatial abilities in completing tasks nearly as fast as those with higher spatial abilities. Authors thus proposed that industry and defense sectors should probably look for gamers as future teleoperators. The advantage of gamers over non-gamers was also observed during teleoperations involving AR. Specifically, Brizzi et al. (2017) tested participants teleoperating a Baxter robot through a pick-and-place task by physically moving their upper limbs, which were tracked via wearable sensors. Different visual feedbacks were implemented in an Augmented Reality (AR) fashion, which were meant to help through the task execution. The authors differentiated their participants according to their level of expertise with AR/VR technologies and gaming. Both AR/VR and video game experts performed better than their non-expert counterparts. However, this difference was nullified when full feedback was provided, suggesting how the advantage of gaming experience on HRIs emerges as long as they demand higher effort or spatial abilities. Furthermore, the reduced gap obtained when adding information via AR is a good example of how these technologies can significantly help operators, particularly the less experienced ones.

There is also evidence of possible relations between gender and gaming experience in teleoperations. For instance, Chuan et al. (2007) let their participants control motion, speed, camera movement, and zoom, as well as arm and gripper movement of two different robots (namely a 4-wheel drive all-terrain pioneer robot, and a lightweight tracked vehicle) via different control modalities (namely keyboard, mouse and gamepad). They programmed two scenarios: an indoor disaster scene with numerous obstacles, and a different scenario in which participants had to control the arms of the tracked vehicle to insert a camera inside a wall aperture. Their performance was quantified based on time spent accomplishing their mission, and the number of times the robot rolled over. Results demonstrated that there was a significant difference between female gamers and non-gamers, but not between male gamers and non-gamers. Furthermore, male non-gamers were faster compared to female non-gamers. Similarly, the already mentioned study from Dybvik et al. (2021) demonstrated how women mostly self-reported having gamed yearly or never. Differently, men used to game most often, typically yearly, monthly, weekly and in a few cases even daily. It follows that, when measuring the effects of gaming experience on robot teleoperations, gender should always be considered too, and vice versa. In fact, it is possible that when finding performance differences between men and women, such differences might be mediated by their respective gaming experiences, and/or spatial abilities.

4.4 Other individual factors

Besides age, gender, and gaming experience, there are other individual factors or soft skills that are often understudied but still fundamental for the emerging occupations created by Industry 4.0. For instance, **learnability skills** embrace the willingness and ability to learn flexibly, which are essential in the ever-changing industry world (Ra et al., 2019). For instance, the telerobotic industry is increasingly adopting new digitalized solutions for operating robots, such as VR. In this view, considering that especially elderly workers might find these devices new, examining the impact that learnability skills would have on teleoperation performance when leveraging such new interactive technologies might open new questions and possible solutions for their implementation.

Moreover, **problem-solving** has been identified as one of the attributes deemed most positively by assembly-line workers (Welfare et al., 2019). However, people have different problem-solving attitudes and this might have an impact on work performance as well.

Finally, **trust in technology** is another individual factor that is known to influence human approaches to technology (Lankton et al., 2015). In the field of HRI, effects of trust have been investigated specifically in relation to social robots (Hancock et al., 2011), with the idea that a higher trust in the robot is likely to promote better human-robot partnerships, but less often as a general inclination to trust technological devices in everyday life. As VR devices as well as robots might be unknown to most of the population, a wider investigation of trust toward general technology might also be informative of how an individual would approach VR-based telerobotics. Overall, research systematically investigating relations between these individual factors and the actual user's performance in the HRI domain is scarce. Yet, they may have a significant impact on users' behavior when driving a robotic system.

Chapter 5

Human factors metrics in robotics

This chapter focuses on the metrics usually leveraged for assessing human factors in robotics research. Specifically, studies in the HRI sector that proposed performance (Section 5.1), and workload assessments (Section 5.2) leveraging self-reports (Subsection 5.2.1), eye-tracking (Subsection 5.2.2) and electroencephalogram (EEG, Subsection 5.3.1) are here reviewed. This literature reappraisal is relevant to the present thesis as it serves as a foundation for all experiments conducted on the interactions between humans and our industrial robotic arm.

A second aspect that cuts across all studies included in this thesis is related to principles of the mobile cognition approach. Indeed, real-world-based or realistic experimental settings were always favored in our works. Furthermore, we always promoted a multimodal and multidimensional assessment of human factors. The rationale behind these concepts is deployed in Section 5.4. More specifically, research approaches such as embodied cognition and mobile cognition are briefly introduced in Subsections 5.4.1 and 5.4.2. They provide a good theoretical background to mobile and multimodal methodologies for studying humans actively interacting with technological devices (e.g., robotic systems). Specifically, Mobile Brain/Body Imaging (MoBI) is tackled as one of the most valuable methodologies in the latest neuroergonomics field allowing concurrent recording of behavioral, physiological and neurophysiological metrics. Finally, in Section 5.4.3, VR is discussed as a powerful mobile research tool for neuroergonomics that can successfully be coupled to other neurophysiological measures, such as EEG.

5.1 Human performance

Performance evaluations are quite common in HRI. Interestingly, while most technical and engineering-based studies use the term "performance" to usually indicate *robot performance*, studies focusing on humans interacting with robotic machines use "performance" to indicate *human performance*. In this thesis, we are mostly interested in *human performance*.

Particularly in the industrial or telerobotics sector, investigations analyzing human performance when interacting with a robot have used video analysis (Rossato et al., 2021), motion capture (Hsieh and Lu, 2018), or datalog directly obtained from the robot system (Rossato et al., 2021). For instance, Rossato et al. (2021) studied the subjective experience of younger and older adults teaming up with a cobot. Besides questionnaires, they measured the user's performance by coding the task execution time from the video recordings of the tasks (video analysis). Chacón et al. (2021) assessed the usability of a human-robot collaboration workspace involving the collaborative robot from Universal Robots, model UR3. For user performance, they measured efficiency as the time to complete the task and effectiveness as the percentage of task fulfillment. The authors used both video analysis and a tool for obtaining the values as recorded directly from the robot through a communication protocol with an external desktop. Hsieh and Lu (2018) measured the task completion time of operators collaborating with either a physical interface or one of three different virtual interfaces. Specifically, the joint coordinates and angles of shoulders, elbows and wrists were captured by a motion capture system, which allowed precise measurement of each task phase duration. Similarly, Weistroffer et al. (2014) assessed the performance of users assembling a car door by recording the task duration and the number of completed operations and collision alarms activated, both in the physical and in a virtual environment. In the physical situation, performance data were computed off-line via video analysis. In the virtual situation, those data were automatically saved through the software.

5.2 Mental workload

Mental workload has been central in human factors research for its association with safety-critical performance (Dehais et al., 2020b). As shown in Table 1 of Young et al. (2015), references to 'mental workload' in literature have more than tripled since the 1980s. This indicates that this concept is still relevant in ergonomics and human factors research. A recent definition of workload was given by Longo et al. (2022), who described it as the combination of: (1) environmental factors, that are situational and whose influence is not under one's direct control; (2) individual factors, like age, personal skills and past experiences, which affect the amount of effort required to cope with the environmental factors and with the primary task over time; finally, (3) the finite pool of cognitive resources that limits every individual's cognitive processing capacity. Therefore, an adequate performance will be achieved if, on top of environmental and individual factors, the task demand does not exceed the individual's cognitive resources.

In human factors, studying mental workload is primarily important as it allows to *quantify* the transaction between humans and machines, tasks or protocols, and to *predict* the probability of a performance decrement in scenarios involving humans and machines (Dehais et al., 2020a). Furthermore, it is common practice to use *dual-task methodology* for manipulating the imposed demand in a controlled way (Sweller et al., 2011). The dual-task, which is the simultaneous performance of two activities, increases mental workload, creates task interference and leads to performance degradation in one or both tasks (Leone et al., 2017). Such protocol is well-established and widely used both in laboratory-based research and in evaluations in the field.

However, when measuring workload, two factors need to be taken into consideration: first, there is the need for **multimodal evaluations** of workload, as implicit and explicit measures of workload have often been found to be divergent (Dehais et al., 2020b; Matthews et al., 2015); second, the operators' behaviors and minds should be evaluated in **real-world environments** (Dehais et al., 2020a; Hoffmann et al., 2018; Ladouce et al., 2017).

Regarding the first point, mental workload can indeed be categorized as a result of subjective self-assessment, or associated with psychophysiological parameters. Each measurement has strengths and weaknesses (O'DONNELL, 1986) and their respective sensitivity can vary depending on the level of workload experienced by the operator (De Waard and Brookhuis, 1996). As pointed out by Dehais et al. (2020b), conducting multidimensional assessments of workload by leveraging different measures inevitably led to an inclusive framework for mental workload, at the price of a dissociation between the different measures, and psychometric gaps (Matthews et al., 2015). Three measures typically used for measuring workload, both explicitly (i.g., self-reports) and implicitly (i.e., physiological indexes that are out of the user's direct control) are tackled in the following Subsections 5.2.1, 5.2.2, 5.3.1. Differently, the second point on the importance of measuring workload in the wild is deployed in Section 5.4.

5.2.1 Self-reports

To date, the most widely used and extensively validated measures of workload are subjective self-reports. The NASA Task Load Index, usually called NASA-TLX questionnaire (Hart, 2006), is the most known one. It is composed by six items, namely mental demand, temporal demand, physical demand, performance, effort and frustration. To date, it is still implemented in a large number of studies that wants to investigate workload in human-computer or human-robot interaction. For instance, Rossato et al. (2021) used the NASA-TLX questionnaire to study the subjective experience of younger and older adults teaming up with a cobot. Similary, Kaufeld and Nickel (2019) evaluated human mental workload related to HRIs in VR. They collected performance data such as response times, error rates, and the number of missed trials and the perceived mental workload through the NASA-TLX scale.

While questionnaires are definitely useful to collect information on the experienced (and therefore explicit) workload, they entail diverse disadvantages. First, their administration inevitably disrupts the task flow. Second, it is prone to inter-subject variability, and to the individual's ability to self-assess. Given those limitations, the question arises whether and how workload can be assessed continuously, without task disruptions, and in a more objective way.

5.2.2 Eye-tracking parameters

The above-mentioned limitations of self-reports can be successfully overcome via the adoption of psychophysiological tools. For instance, eye-tracking systems are appealing in work contexts as they are increasingly portable and affordable and can capture workload-related eye behavior in the field (Novak et al., 2015) Indeed, they were often used to predict workload in human-machine interactions, particularly during driving or in air traffic control operations (Ahlstrom and Friedman-Berg, 2006; McIntire et al., 2014), and in fewer cases also in the robotics domain (Novak et al., 2015). Also, the most recent VR headsets are provided with an integrated eye-tracker that allows the measurement of eye-related physiological indexes within virtual environments, without disrupting users' actions. For example, the VIVE Pro Eye is equipped with a Tobii eye-tracker, has an output frequency 120Hz, and has a trackable field of view 110°. Furthermore, it is easily viable in wireless modality, allowing higher freedom of movement and greater flexibility. This is an exceptional opportunity for continuous workload monitoring during human-robot interactions even in VR. However, since virtual devices that couple eye-movement detection with the highly interactive nature of VR are fairly recent, VR eye data in experimental and particularly robotic scenarios are still unexplored (Souchet et al., 2022).

Pupil size. A number of studies found empirical evidence in favor of utilizing increased pupil diameter as an indicator of a higher mental workload (Beatty, 1982; Iqbal et al., 2004; Van Orden et al., 2001). Specifically, several studies demonstrated how pupil diameter increases with increasing cognitive workload (Kahneman, 1973; Marinescu et al., 2018; Pomplun and Sunkara, 2019). Pupil variation is linked to the parasympathetic nervous system, which is involved in arousal and wakefulness (Mathôt, 2018). Therefore, how pupil size variations are modulated by operations with one or the other interface can provide information about the level of implicit workload within a joint operation (Mingardi et al., 2020). It is worth noting that many variables other than the user's cognitive and emotional state (e.g., ambient lighting) can affect this metric (Kramer, 2020). This makes crucial crucial to keep constant lighting of the setting, preprocessing pupil data, and apply proper baseline correction in order to exclude pupil size variations that are possibly unrelated to the user's cognitive activity (Mathôt, 2018).

Despite the stable relationship between pupil size and workload, research works applying pupillometry to the industrial and work field are scarce. Evidence of the relation between pupil diameter and workload in robotics was repeatedly found in the surgical domain, where participants were asked either to perform a robotic surgical task under different difficulty levels (Wu et al., 2020) or to bring rubber objects over dishes with different target sizes and distances (Zheng et al., 2015). More broadly, eye movements were employed as indicators of mental workload in a desktop-based version of a combat management workstation aboard naval vessels (Greef et al., 2009);

the authors proposed eye parameters as a potential trigger for adaptive automation of the system. Savur et al. (2019) presented two case studies involving a robotic arm in a collaborative pick-and-place puts with the user's and the robot's behavior (pupil dilatation, EEG, GSR, PPG), the data collected about the user's cognitive state were not analyzed as a function of the task. Van Acker et al. (2020) tested the feasibility of deploying pupillometry in a work setting demanding operator mobility. They tested participants performing two manual assembly tasks with a different degree of complexity and found no significant differences in the implicit workload as suggested by pupil size variations, in contrast with substantial differences in the subjective workload. Therefore, they advocated further testing of pupillometry measures in real-life work settings to better understand their actual feasibility.

Perclos. Perclos is a robust measure of vigilance for humans interacting with machines, particularly in the automotive area (Du et al., 2022). It can be defined as the percentage of time that the eyelids cover the eye area by more than 80% and can be thus gained from continuous data on eye openness. Literature on this metric showed that higher levels of fatigue and lower vigilance are associated with a higher perclos (Marquart et al., 2015). However, to the best of our knowledge, there is no study demonstrating the reliability of perclos as a measure of vigilance or fatigue in the robotics sector. Indeed, in the previously citepd study of Wu et al. (2020), only pupil diameter and gaze entropy differed between different task difficulty levels, while perclos did not show any significant variation between the conditions.

Blinks. Blinks can also be informative of one's level of workload and/or fatigue (Fogarty and Stern, 1989; Kim et al., 2022; Marquart et al., 2015). For example, blink frequency was demonstrated to be inversely related to the level of mental load (Borghini et al., 2014; Holland and Tarlow, 1972; Zheng et al., 2012), and to the performance in a static simulated air traffic vigilance task (McIntire et al., 2014). Similarly, during air traffic control operations, a decrease in blink duration was observed in case of increased visual workload (Ahlstrom and Friedman-Berg, 2006), while increased blink duration was associated with a deterioration of performance in a vigilance task (McIntire et al., 2014). A possible explanation is that, under high mental demand, users tend to inhibit blinks to reduce the risk of missing incoming information (Fogarty and Stern, 1989). Evidence in this direction was also found when executing a simulated laparoscopic task: fewer and shorter blinks were associated with a higher workload as self-reported at the NASA-TLX score (Zheng et al., 2012). More recently, Guo et al. (2021) evaluated the mental workload during a space robot teleoperation: participants controlled a robotic arm via desktop and joystick under varied latency and time pressure, which is known to affect workload. When time pressure increased, blink frequency decreased and pupil size increased, while no substantial differences were observed across latency manipulations.

Integrating multiple eye-tracking parameters in robotics. Some studies also showed how eye-tracking measures could predict workload during HRIs. For instance, Novak et al. (2015) investigated the sensibility of different continuous eye-tracking indexes in estimating workload via machine learning, proposing a continuous inference rather than a classification into discrete classes of workload indexes. During the task, targets containing equations were presented on a screen for a few seconds. Using the ARMin robot, participants had to hit only the targets that contained the correct equations. In this context, eye-tracker metrics were able to identify progressive increases in workloads induced by a gradual increase in the task's difficulty. Specifically, pupil dilatation was the most sensitive index to workload, while blink and fixation frequencies were the most sensitive to effort. Similar results were obtained by Gao et al. (2013), who compared the ability of different workload-related eye measurements (including blink parameters and pupil dilatation) in predicting the overall mental workload as self-reported via the NASA-TLX questionnaire during digital nuclear power plant operations. Interestingly, none of the single measures was reliable in assessing the overall mental workload. However, when integrating all the measures within a predictive model, the overall mental workload was assessed accurately. Furthermore, blink rate demonstrated higher sensitivity to workload, while pupil size was more sensitive to error-related attention and arousal.

5.3 Embodiment and motor control

The ability to accurately and precisely control the movement of a device is essential for successful teleoperations, as it allows the human operator to effectively and efficiently perform tasks, navigate complex environments, and interact with the surrounding world. In this view, motor control is extremely important. In this thesis, and specifically in Chapter 10, we present a study on embodiment and motor control on our industrial robotic arm in VR, in which we conducted a multimodal analysis including self-reports, behavioral and EEG metrics. While behavioral aspects related to embodiment and their impact on teleoperation have already been addressed in Chapter 3, this section is meant to unfold brain mechanisms of embodiment and motor control, particularly via EEG.

5.3.1 EEG parameters

Neural markers of workload and motor control have been extensively studied in literature through EEG. Its strength builds upon its potential to continuously and unobtrusively estimate workload and motor control difficulty with high temporal resolution (Memar and Esfahani, 2018), which makes it a valuable research tool in many applied contexts. As will be addressed in subsection 5.4.2, EEG data can also be recorded from participants in free motion, significantly increasing the ecological validity of applied research. Literature proposing EEG assessments in telerobotics, and more in general in teleoperation contexts, is limited. Even more, to the best of our knowledge, neurophysiological research on the embodiment and/or motor control of industrial robotic arms, is lacking. Therefore, we here list the most influential research works that served as a base for our study, as they specifically analyzed EEG data by targeting the μ-desynchronization phenomenon when driving and/or embodying external objects/bodies.

μ-desynchronization. EEG μ-desynchronization is a phenomenon that typically occurs during voluntary movements. When it is observed in relation to an event, it is also called μ-ERD (mu event-related desynchronization). It is characterized by a decrease in the power of the EEG alpha (8-13 Hz) frequency band over the motor and sensorimotor cortex, which are involved in preparing and executing voluntary movements, as well as in the processing of sensory information and the coordination of movement. In electroencephalography, these areas can be monitored using electrodes placed over the scalp. The region corresponding to the primary motor cortex and sensorimotor cortex are usually covered by the electrodes "C3", "C4", "P3" and "P4" according to the international 10-20 system, and are also the most typically targeted in EEG studies on embodiment and motor control (Alchalabi et al., 2019; Ding et al., 2020; Evans and Blanke, 2013; González-Franco et al., 2014).

There has been some research on the relationship between μ-ERD and the embodiment of external arms, such as prosthetic arms or virtual arms. For instance, Evans and Blanke (2013) induced illusory hand ownership in participants passively wearing VR and experiencing congruent and incongruent stroking on their hand and the hand displayed in VR. Participants could see either stereoscopic virtual arms, or non-body objects projecting from their body on a virtual table. The authors found that illusory hand ownership is reflected by a body-specific modulation in the μ-band over fronto-parietal cortex. Specifically, synchronous visuotactile stimulations generated a stronger μ-ERD (event-related desynchronization, i.e., decrease of the signal power compared to baseline) compared to the condition with asynchronous stimulations particularly over the electrode C4, but only when participants were displayed with humanoid arms. In the study of González-Franco et al. (2014), participants observed an attack on their virtual hand while immersed in VR: in one case, they saw a knife attacking their hand, in the other case, the knife only struck the virtual table. When the virtual hand was threatened, μ-ERDs were clearly observed over the motor cortex (particularly over the electrode C3), and the participants retreated their hand to avoid harm. Remarkably, the higher was the perceived ownership, the stronger was the μ-ERD. Ding et al. (2020) integrated vibrotactile stimulation to the mirror visual feedback, which is typically used in neurorehabilitation. They instructed participants to open and close their non-dominant hand (active side) while keeping their dominant hand still (static side), and tap a foot pedal when they felt a sense of embodiment. By comparing conditions of no vibration, continuous vibration and intermittent vibration, they found that a stronger sense of embodiment was self-reported during either vibration conditions. Furthermore, in these conditions, a stronger μ-ERD was observed over central-frontal regions (C3 and F3) only in the active side of the brain. Alchalabi et al. (2019), instead, used EEG to measure embodiment in participants controlling a walking self-avatar in VR. Similarly to our work, they introduced visuomotor inconsistencies during the walking task (i.e., the avatar took a step with the contra-lateral limb or stopped walking before the participant stopped). They demonstrated that subjective levels of embodiment changed with the visuomotor inconsistencies, and strongly correlated with differences in μ-ERS (event-related synchronization, i.e., increase of the signal power compared to baseline) over the motor and pre-motor cortex. Specifically, when controlling the avatar's gait, they observed a strong central and parietal μ-ERS in the case of "no manipulation", and a stronger frontal μ-ERS in the case of visuomotor inconsistency.

It thus seems that EEG spectral power density, and specifically the μ-ERD, actually reflect brain mechanisms related to embodiment and motor control. However, as can be noticed, while all mentioned studies leveraged VR technologies to induce embodiment (González-Franco et al., 2014) or as a teleoperation mean (Evans and Blanke, 2013), none of them specifically focused on applied telerobotics, let alone on the industrial field. Further applied research leveraging neurophysiological tools such as EEG in this sector is needed to better understand mechanisms of embodiment and motor control, which are extremely relevant for ensuring better teleoperation performances.

5.4 Human factors in the wild: mobile perspectives for neuroergonomics

5.4.1 Embodied cognition

The inspiring paper of Ladouce et al. (2017) illustrates how our minds are used to i) perceive multiple stimuli from an ever-changing environment, ii) process and understand those inputs, and finally iii) produce flexible and adaptive reactions. Specifically, multimodal sensory and motor signals are continuously integrated, and a real-time feedback system allows to execute dynamic, flexible and coherent actions. Therefore, every behavior that a researcher may want to study and analyze usually generates from an interdependent relation between cognition, perception and movement. This is nothing but the main definition of *embodied cognition* (Coello and Fischer, 2015; Shapiro, 2010). According to this approach, none of these components (i.e., cognition, perception, movement) can be studied independently from each other. For instance, it was proposed that cognitive functions can only be understood when considering their relevance for actions (Wilson, 2002). Similarly, Hoffmann et al. (2018) thoroughly argue about the close interplay between *actions* and *cognition*: specifically, the neurophysiological mechanisms of actions, as well as the mechanisms generated by any interaction with the environment, may be the core of cognition. In this view, it is clear how applied research should allow participants to freely move in any experimental setting, in order to study cognition in its most natural form, that is *in the wild*.

However, **traditional research** has always tried to isolate behaviors to increase the experimental control. Consequently, many experiments took place in relatively static and often simulated laboratory settings, using artificial and stereotyped stimuli and constraining the participant to highly specific instructions (Ladouce et al., 2017). While these set-ups benefit from high experimental control, they lack real-world dimensionality and freedom of movement.

5.4.2 Mobile cognition

Recent technological development is producing increasingly portable, wireless, and mobile devices. Those devices include those intended for the end-user (e.g., smart glasses, VR headsets) and also experimental tools for the analysis of human brain and behavior (e.g., eye-tracking, fNIRS, EEG, motion capture). Researchers can thus record behavioral, neural and physiological metrics underlying cognitive processing while humans naturally interact with their environment (Gramann et al., 2014; Makeig et al., 2009). Consequently, participants benefit from more real-world-based experiments, more realistic scenarios and higher freedom of movement while proper experimental control is still ensured. The approach that is fostering such evolution is called *mobile cognition* (Ladouce et al., 2017). Remarkably, the mobile cognition approach does not aim at replacing traditional and laboratory-based research; on the contrary, it intends to complement this research in a parallel way. Indeed, while there are research questions that can only be answered in laboratory contexts, other research questions can be better addressed in real-world settings. Combining both research approaches will naturally lead to a more accurate understanding of human behaviors and brain functioning in each applied context.

Mobile Brain/Body Imaging (MoBI)

By following the wave of a mobile cognition approach, one specific method that revolutionized the cognitive and neuroscience field is the Mobile Brain/Body Imaging (MoBI) method (Gramann et al., 2011; Jungnickel et al., 2019). Such methodology allows to record and analyze brain and motor activities in naturalistic conditions, also in work contexts (Wascher et al., 2014). One of the first examples of successful implementation of MoBI comes from Gramann et al. (2010), who tested participants wearing a high-density EEG system while actively walking on a treadmill and performing a visual oddball response task. Unlike mobile EEG, MoBI examines the relationship between movement, cognition and brain dynamics while accounting for human motion as an informative input that entangles cognitive processes. However, the multimodal nature of this approach poses various methodological and technical challenges, which have been thoroughly and increasingly addressed in the latest years.

One of the technical challenges of MoBI regards the **synchronization of different data streams** often acquired via multiple sensors. In this regard, software such as Lab Streaming Layer (LSL) (Kothe et al., 2014) can provide a collection of data recorded from different interfaces (e.g., EEG, motion capture, eye-tracking data, experimental markers streams), unifying and synchronizing the different data streams. Multiple libraries exist for MATLAB, C++, C#, Python, and Java.

Furthermore, the contamination of the brain signal with **muscular, electrical and mechanical artifacts** is another important factor to consider when conducting MoBI research. Indeed, muscular artifacts can generate from the participant's physical movements during the experiment, while electrical or mechanical artifacts can generate from touching cables or the cap, or putting off and on equipment such as an eye-tracker system or VR headset. In this regard, Klug et al. (2022) developed the BeMoBIL Pipeline, which is a streamlined and well-documented pipeline in MATLAB for the time-synchronized handling of multimodal data, including EEG and motion data. It provides automated functions for EEG preprocessing, advanced artifact handling and source separation, as well as functions for motion data processing. It also allows extraction of event markers from different data modalities and especially from EEG (e.g., eye movements, gait-related events) using Independent Component Analysis (ICA). It is an open-source project which can be found at https://github.com/BeMoBIL/bemobil-pipeline and is open to community-driven adaptations. This contribution makes mobile cognition research more transparent and affordable for everyone, representing a unified methodological approach to mobile data.

5.4.3 VR as a mobile research tool for neuroergonomics

VR headsets are increasingly gaining ground also in mobile cognition research (Jungnickel et al., 2019). This is mainly due to their multisensorial nature, their flexibility, portability, and highly interactive features. Specifically, VR devices allow to control visual, auditory and to some extent also haptic inputs during movement, which makes them a more naturalistic solution compared to traditional desktop-based experiments. Furthermore, most VR devices are equipped with tracking systems intrinsic to the headset and controllers. Additional body trackers can be easily implemented too, making it possible to accurately track also chest, wrist, elbows and fit, with six degrees of freedom (i.e., position and orientation in the three axes). This feature allows an easy collection of motion capture data during any interaction with the virtual environment. The latest VR devices are even equipped with an integrated eye-tracker. For instance, the VIVE Pro Eye is provided with a Tobii eye-tracker system, allowing the measurement of eye movements in free motion while also being protected from external light. All motion, interaction and behavioral streams deriving from the VR headset can be implemented via game engines such as Unity, Vizard, or Unreal, which allow programming virtual environments. Furthermore, VR can also be coupled with other external physiological tools, such as EEG (Djebbara et al., 2019; Gehrke et al., 2019; Nenna et al., 2021; Singh et al., 2018). As previously mentioned, software frameworks such as LSL can then synchronize the large arrays of data collected from the VR system, with those generated via external physiological systems (e.g., EEG, eye-tracker).

In the end, VR can provide enriched MoBI environments including visual, auditory and tactile stimuli, as well as motion capture and eye-tracking options (e.g., VIVE Pro Eye), with the possibility of being integrated with external physiological devices. All these characteristics make VR a powerful tool for mobile cognition research. Literature repeatedly demonstrated the feasibility of coupling VR with EEG. Gehrke et al. (2018) used VR, EEG and motion capture to investigate human navigation during free ambulatory exploration of a virtual environment. Singh et al. (2018) investigated brain mechanisms of prediction error as a function of visual appearance of their hand in VR. On the same research line, Gehrke et al. (2019) used the prediction error negativity registered via EEG to detect visuo-haptic mismatches in a VR environment. Djebbara et al. (2019) leveraged MoBI and demonstrated how brain dynamics related to sensorimotor experiences in VR reflect architectural affordances. We successfully recorded attentional neural markers (i.e., P300) in a dual-task walking paradigm leveraging VR and a 128 channels mobile EEG (Nenna et al., 2021).

These exemplary research works show how multimodal evaluations of human behaviors, including motion, and the relative brain dynamics can be successfully conducted using mobile neurophysiological tools and VR. In such cases, the VR potential as a research tool extends not only to behavioral and cognitive science research but also steps inside neuroscientific questions, opening new perspectives in the **neuroergonomics** sector (Dehais et al., 2020a; Gramann et al., 2017). For instance, Carrieri et al. (2016) simulated an interaction with a real, remotely-driven, system placed in a critical environment in VR; in other words, a VR-based simulated teleoperation. They examined the neural correlates of a cognitive/motor task through Functional Near Infrared Spectroscopy (fNIRS) and pointed toward combining VR and fNIRS as a promising platform for neuroergonomic studies. Similarly, Zhu and Du (2020) used VR as a testbed and data collector for examining users' reactions to different human-robot interface designs. They also collected eye-tracking, hand, and body position data, and additionally used fNIRS to measure real-time hemodynamic responses. They argue that the neurobehavioral data collected from the VR device serve directly as a personalized model for human-robot interface optimization. Those are just a few examples of multimodal assessments of human factors leveraging VR in the neuroergonomics sector. Interestingly, while mobile EEG was deployed in numerous examples of Brain-Computer Interfaces for allowing communication between humans and robots, it is scarcely explored as a tool for assessing humans' brain dynamics when interacting with robots.

Part II

EXPERIMENTS

Chapter 6

Study 1 - The virtualization of human-robot interactions: a user-centric workload assessment

Interest in the virtualization of human–robot interactions is increasing, yet the impact that collaborating with either virtual or physical robots has on the human operator's mental state is still insufficiently studied. In the present work, we aimed to fill this gap by conducting a systematic assessment of a human–robot collaborative framework from a user-centric perspective. Mental workload and performance were measured in participants jointly performing a pick-and-place task with both a physical collaborative robot and its virtual equivalent. They performed the same task in conditions of either low or high cognitive load, which was induced through dual-tasking (i.e., concurrent execution of the pick-and-place task with a secondary arithmetic task). Therefore, the experiment followed a 2 × 2 repeated measures design over interface (physical and virtual) and task load (single task, dual task). Performance and self-reported data as well as eye-tracking data were collected and analyzed. Furthermore, the level of participants' expertise with VR and users' preferences for potentially working with virtual or physical robots collaboratively were assessed. The hypotheses and research questions are listed below.

6.1 Hypotheses and research questions

6.1.1 Task load manipulation

Engaging the user in a dual task is a well-known condition that causes an increase in the load on cognitive resources (Navon and Miller, 1987). Therefore, as a methodological control, we expected the task load manipulation to generate a higher explicit and implicit workload in the dual—compared to the single-task condition. More specifically, we predicted a greater pupil size increase and a higher perceived workload from the NASA-TLX in the dual-task compared to the single-task condition. Additionally, we predicted longer operation times for the pick-and-place task and more errors for the arithmetic task in the dual-task condition, which would indicate behavioral interference of the secondary task with the primary pick-and- place operation.

6.1.2 Operators' behavioral performance

The current literature on the impact of virtual vs. physical cobot manipulation on a user's performance is limited. However, the utility of virtual interfaces in industrial contexts can be corroborated only to the extent that the performance of users collaborating with a robot in VR does not decrease compared to that of users working in the physical space. In this respect, performance with the virtual interface was expected to be comparable to performance with the physical interface under both high and low workloads.

6.1.3 Operators' cognitive state

Previous studies did not systematically address how direct interactions with a physical or virtual cobot affect the user's cognitive state. Therefore, we intend to fill this gap by exploring whether collaborating with a physical or virtual cobot affects the user's workload under either the single- or dual-task condition. With this aim, we analyzed pupil size variations and responses to the NASA-TLX questionnaire as a function of implicit and explicit workload.

6.2 Methods

6.2.1 Sample

The experimental sample consisted of 26 participants, 8 women and 18 men (M_{age} = 26.65; SD_{age} = 5.29), who volunteered to take part in the study and signed the informed consent. None of the participants had current or past neurological or psychiatric problems. They all had normal or corrected-to-normal visual acuity and reported having normal color vision. The experimental protocol was approved by the local ethics committee and the study was conducted according to the principles of the Declaration of Helsinki. Three participants were excluded for technical issues related to the eye tracker device. Moreover, one participant was excluded for having an error rate of more than 50% at the arithmetic task and another for missing data in the arithmetic task. The final sample comprised 21 participants, 5 women and 16 men (M_{age} = 26.95; SD_{age} = 2.52).

6.2.2 Technical setup

Physical condition. In the physical condition, participants were provided with a pair of binocular eye-tracking glasses (Pupil Labs GmbH ©, Berlin, DE; weight 22.75 g) connected to an MSI laptop (model GT63 Titan 8RF, processor Intel Core i7-6700HQ, screen resolution 1920 × 1080, RAM 16 Gb).

The software Pupil Capture (Pupil Labs GmbH©, Berlin, DE) enabled system calibration and data recording (sampling frequency: 120 Hz; calibration: 5-point). The software Pupil Player (Pupil Labs GmbH©, Berlin, DE) was utilized to export the eye-tracking data. Besides the eye data, the eye-tracker device also enabled first-person video recording through the embedded scene camera (480p, field of view: 100° × 74°; sampling frequency: 120 Hz). The video recordings were then used for conducting a video analysis of the participants' arm behavior. The arithmetic task was managed through a program written and compiled in Visual Studio 2019 running on the same MSI laptop handling the eye-tracking recording. The pick-and-place task was performed jointly with an UR10e cobot 6.1, which was installed on a height-adjustable worktable and was programmed in Polyscope (version 5.11) through its teach pendant. All data were recorded and processed by the same laptop and were thus synchronized based on the same internal clock.

Virtual condition. In the virtual condition, participants were provided with an HTC Vive Pro Eye headset (resolution: 1440 × 1600 pixels per eye; refresh rate: 90 Hz; Field of view: 110°) and its controllers. The same headset also comprises an eye-tracking system (sampling frequency: 120 Hz; calibration: 5-point) which enables recording of eye parameters throughout the tasks. The virtual environment 6.1 was programmed in Unity (version 2019.4.18f1) and faithfully reproduced not only the cobot and its workstation but also the surrounding environment (that is, windows, furniture, door). Participants interacted with the virtual cobot by means of physical action and by responding through the HTC Vive controller. At the end of each experimental session, all data (behavioral and eye data) were automatically saved on an MSI laptop (Intel Core i7-6700HQ, screen resolution 1920 × 1080).

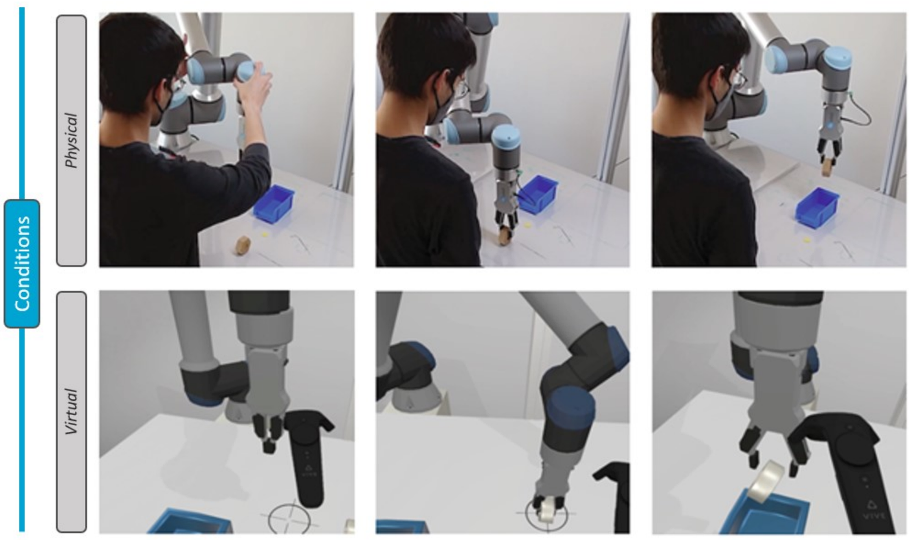

FIGURE 6.1: A participant driving the physical and the virtual robot

6.2.3 Procedure, experimental tasks and design

After signing informed consent, all participants filled out a demographic questionnaire and answered questions about their VR expertise and individual preference for virtual vs. physical cobots. Then, they undertook six tasks composed of 25 trials each. In particular, as shown in Figure 6.2, a pick-and-place task, an arithmetic task and a dual task were performed both in the virtual and physical environments (counterbalanced order). Half of the participants started with the virtual condition and the other half with the physical condition. At the beginning of each task condition (virtual and physical), all participants underwent a training session and performed a few trials of the same tasks administered in the subsequent experimental session. The task instructions were presented in paper format in the physical condition, and they were virtually delivered in text format in the virtual environment. The experiment started only when the participant understood all the task rules. In both contexts, a 5-point calibration of the eye-tracking systems was conducted before starting the experiment. After each task, participants filled in the NASA-TLX questionnaire and, only at the end of the virtual experimental session, the MEC-SPQ was also administered. Additionally, between each task, it was possible to take a break both in the virtual and physical environments, after which the eye-tracking system was recalibrated before starting the next task. At the end of both the virtual and physical experimental sessions, the final questionnaire on the individual preference for virtual vs. physical cobots was administered.

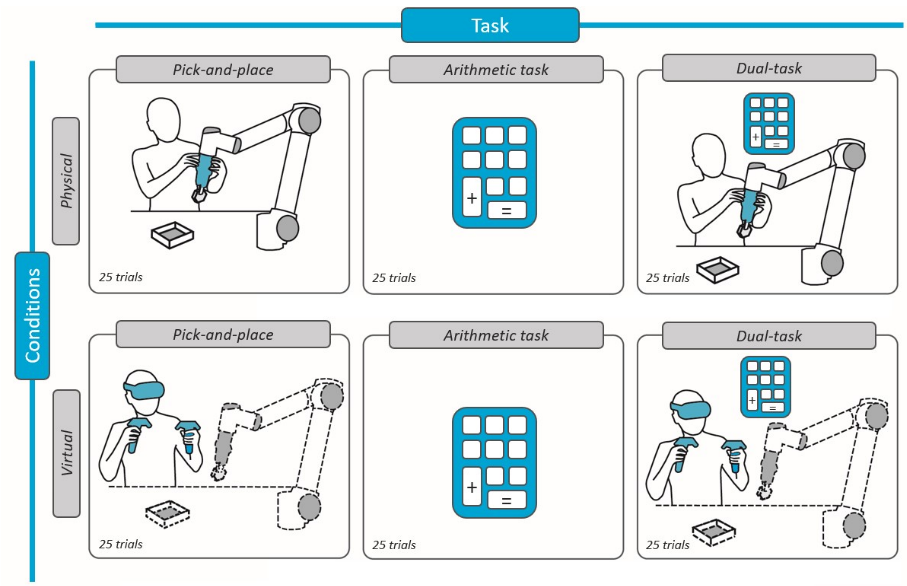

FIGURE 6.2: Experimental tasks and conditions

Pick-and-place task

For each trial of the pick-and-place task, a bolt and a box were always placed in random positions on the worktable, still keeping 50 cm of distance between them. Participants were instructed to pick the bolt up from the worktable and place it into the box by physically moving the robotic arm. The activity was designed to distinguish clearly between the pick and place phases. The pick phase required precise maneuvering of the robotic arm to align its effector with the bolt to be picked up. For the place phase, on the other hand, less precision was needed because the box in which to place the bolt was relatively large.

In the **physical condition**, participants first had to grasp the robotic arm with their hands and physically move it close to the bolt (*moving robot* of the pick phase, Figure 6.3). Once the robot's effector was in line with the bolt, participants initiated the grab bolt automation depicted in Figure 6.3 by gently hitting the worktable with their hand and the robotic arm automatically picked up the bolt (*bolt grabbed*, Figure 6.3). Afterward, they grasped the robotic arm, positioned it over the box (*moving robot* of the place phase depicted in Figure 6.3), and hit the worktable again to enable the cobot to automatically release the bolt in the chosen position (*release bolt*, depicted in Figure 6.2). We used the Wizard of Oz method Hsieh and Lu (2018) for initiating the grab bolt and release bolt automations: when participants touched the worktable, an experimenter standing behind the participants initiated the grab/release bolt command from the teach pendant. This mechanism was hidden from the participants, who were led to perceive the feature as related to their action of touching the table.

In the **virtual condition**, participants used the HTC VIVE controllers to perform the same VR task. Specifically, they were instructed to approach the cobot with their hand. When the controller physically collided with the virtual cobot, participants could grasp it by keeping the grip button pressed and move it to the desired position as in the physical condition. To initiate the grab bolt and release bolt automations, they pressed the pad button on the right controller.

Arithmetic task

A series of numbers randomly ranging between 1 and 10 were aurally presented to the participants, who were asked to mentally sum them and then report the result of the arithmetic operations. Between each number, a time interval of 2.5 s ± 0.3 s of jitter elapsed, and each series comprised 4 or 5 numbers to avoid possible learning effects. In the virtual condition, participants reported the result of each mental operation by interacting with a virtual numeric keyboard via controller. In the physical condition, they were asked to report the sum's result verbally. The response was systematically collected via the Visual Studio application described in subparagraph 6.2.2.

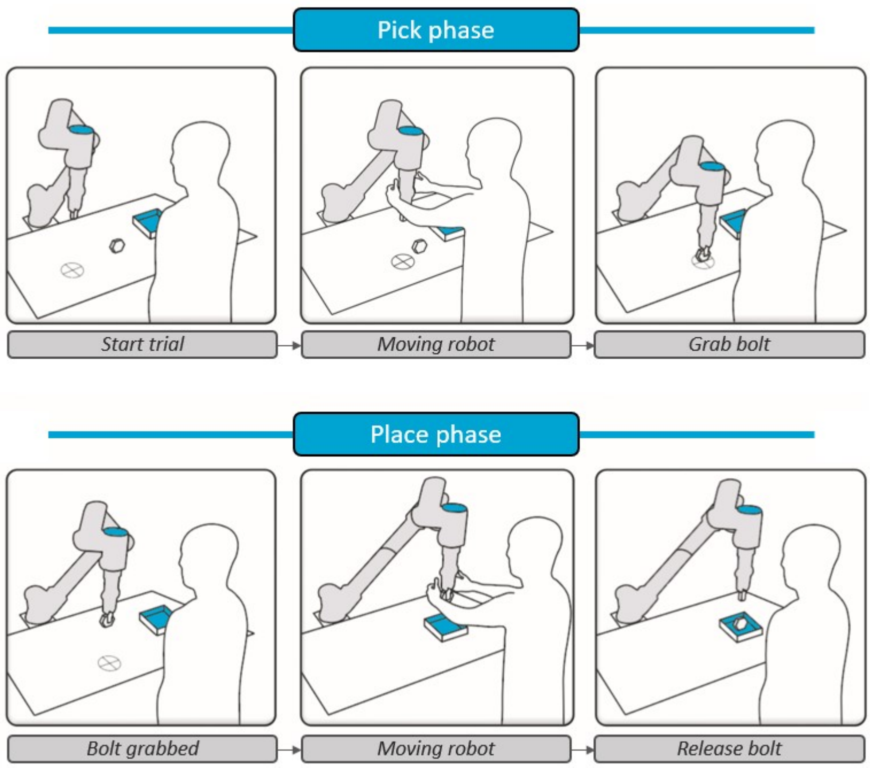

FIGURE 6.3: The pick-and-place task

Dual task

In the dual task, participants were instructed to perform the pick-and-place task and the arithmetic task concurrently. In each trial, the numbers of the arithmetic task were presented for the whole pick-and-place task, and the result was then reported only after the release bolt action 6.2.

6.2.4 Measurements

Behavioral performance

Pick-and-place task. In the pick-and-place task, the behavioral performance was measured as operation time, that is to say, the time required for the user to move the robotic arm to the desired location to either grab or release the bolt. More specifically, the operation time was computed from the time the user first touched the robotic arm (start) until the moment when the user released it (end) for both the pick and the place phases 6.2.

In the physical condition, the operation times were computed by coding the video recordings of the experimental trials1 with the software BORIS (version 7.10.5, Friard and Gamba (2016)). More specifically, the first frame showing the user's hands touching the physical cobot was coded as the beginning of the operation time, and the frame showing the user's hands releasing the cobot was coded as the end of the operation time. The obtained "start" and "end" timestamps were imported into the pupillometry data stream.

In the virtual condition, the first movement of the virtual robotic arm was automatically logged by the Unity software as the timestamp of the button press (grip button) that co-occurred with the contact between the controller and the virtual cobot. Likewise, the software logged the end of the operation time as the timestamp of the button press (pad button). Overall, for both the physical and virtual conditions, the operation times for picking up and placing the bolt were considered independently because of different levels of difficulty; the pick phase required higher precision for positioning the cobot's joints in a suitable and accurate way, but in the place phase less accuracy was required.

Arithmetic task. In the arithmetic task, we computed the percentage of wrong answers. This performance index provided information on the degree of cognitive interference that occurred in the dual—compared to the single-task condition both in the virtual and physical conditions.

Pupil size variation

The pupil size variation was computed only during the moving robot phase (Figure 6.3) and was considered as a proxy of the experienced workload (Beatty, 1982; Iqbal et al., 2004; Van Acker et al., 2020). In this study, we followed the preprocessing methodology of Kret and Sjak-Shie (2019) and accommodated the precautions of Mathôt (2018) for the baseline correction. First, we selected time windows within the moving robot parts of both the pick and the place phases (depicted in Figure 6.3) and handled them independently from each other. Considering that different operation times resulted in different lengths of the selected time series, we used dynamic time warping (Berndt and Clifford, 1994; Keogh and Pazzani, 2001) to standardize the length of each time series. Thus, all pupil samples were constrained to fall within a warping window of 30 data points. The average length of the selected windows was about 1.5 s; therefore, each data point of the warped window corresponded to 50 ms on average. After averaging over the left and right eye, we computed the percentage of missing data in each trial and participant and removed those for which more than 35% of the data were missing (1 trial and 0 participants were removed). Data were then filtered through a median filter, and the first 4 data points of each trial—which correspond to 200 ms on average—were used to apply a subtractive baseline correction (Mathôt, 2018). By addressing the difference in pupil size compared to a baseline period, we marginalized absolute differences caused by external variables other than those due to changes in the cognitive state. Unlike for the processing of pupil response during the pick-and-place task, in the arithmetic task we selected four time windows cor- responding to each number presentation. Because their duration varied between 2.3 and 2.7 s, dynamic time warping was applied in each of the windows to standardize their length (Berndt and Clifford, 1994; Keogh and Pazzani, 2001). Then the same procedure was followed for the processing of pupil data in the arithmetic task.

Self-reported workload

After each task, participants were asked to fill in the NASA-TLX questionnaire as a measure of perceived workload. This scale has been used extensively in many areas, with the industrial context being just one of them (Kaufeld and Nickel, 2019).

Individual factors

Participants were asked to self-report their level of previous experience with VR technology by rating the frequency with which they had used VR devices on a 5-point scale. The aim was to control for the level of VR experience within the sample. Additionally, we explored participants' expectations of working with the cobot compared to their experience with it by asking their preferences before and after the experiment. More specifically, before the experiment, we asked: "If you had to collaboratively work with a cobot, which of the following interfaces would you prefer?" and, after the experiment, we asked, "With reference to the experience you have just concluded, which of the following interfaces did you prefer?". The possible answers were "Virtual cobot" and "Physical cobot".

6.2.5 Statistical analysis

Behavioral performance

Performance data were analyzed using generalized linear models (GLMs) from the lme4 package (Bates et al., 2014) in RStudio (Team, 2021). To analyze performance at the arithmetic task and at the pick-and-place task, we computed a GLM that included the factors task load (single task, dual task) and interface (virtual, physical) with participant as a random effect. Specifically, for the operation times at the pick-and-place task, the pick and place phases were analyzed independently. The Bonferroni correction was always applied when interpreting the post hoc contrasts within the significant interactions (Bonferroni, 1936).

Pupil size variation

To analyze pupil size variations, we used linear mixed-effects models (LMERs) (Bates et al., 2014). Specifically, we ran a chunk analysis over six windows (each corresponding to 250 ms on average) to determine significant differences in the time course. The models involved task load (single task, dual task), interface (virtual, physical), window (1, 2, 3, 4, 5, 6), and their interactions with the participant as a random effect. As for the operation time, the pick and the place phases were analyzed independently. In the single arithmetic task, we analyzed whether there were significant differences in pupil size between the beginning of the arithmetic task (start) and the following arithmetic sums (first, second, and third arithmetic operations). As three or four arithmetic operations occurred randomly, only the first three arithmetic operations were considered to prevent the learning effect. We computed one LMER for each interface condition (virtual and physical) with arithmetic operation as a fixed factor (start, first sum, second sum, third sum) and participant as a random effect. The Bonferroni method was consistently applied in the post hoc contrasts analysis (Bonferroni, 1936).

NASA-TLX questionnaire

The analysis of the NASA-TLX questionnaire score was conducted through a GLM over task load (single task, dual task), interface (virtual, physical), and items (mental demand, physical demand, temporal demand, performance, effort, frustration), with participant as random effect. Post hoc contrasts were performed on each of the significant interactions with the application of the Bonferroni correction for multiple comparisons (Bonferroni, 1936).

Individual factors

For individual VR experience, we first standardized the participants' responses and then created two levels of VR experience: participants with a scaled score below 0.5 were assigned to the low VR experience level, and those with a scaled score higher than 0.5 were assigned to the high VR experience level. With regard to the individual preference for a virtual or physical cobot as expressed before and after the experiment, we reported the percentage of answers in favor of the virtual or physical cobot.

6.3 Results

6.3.1 Behavioral performance

Pick-and-place task. Results yielded significant main effects only for interface, both in the pick phase (X^2 (1, N = 21) = 1057.5, p < 0.0001) and in the place phase (X^2 (1, N = 21) = 1252.4, p < 0.0001), with a faster operation time for the virtual interface compared to the physical interface, both in the pick and in the place phases (Figure 6.4). The task load manipulation, however, did not yield any significant differences in operation times in the pick or the place phase. Descriptive statistics relative to the operation time are found in Table 6.1.

Arithmetic task. When analyzing the effects of task load and interface on the arithmetic task error, none of the factors reached the significance threshold. Nonetheless, the arithmetic task error is depicted in Figure 6.5.

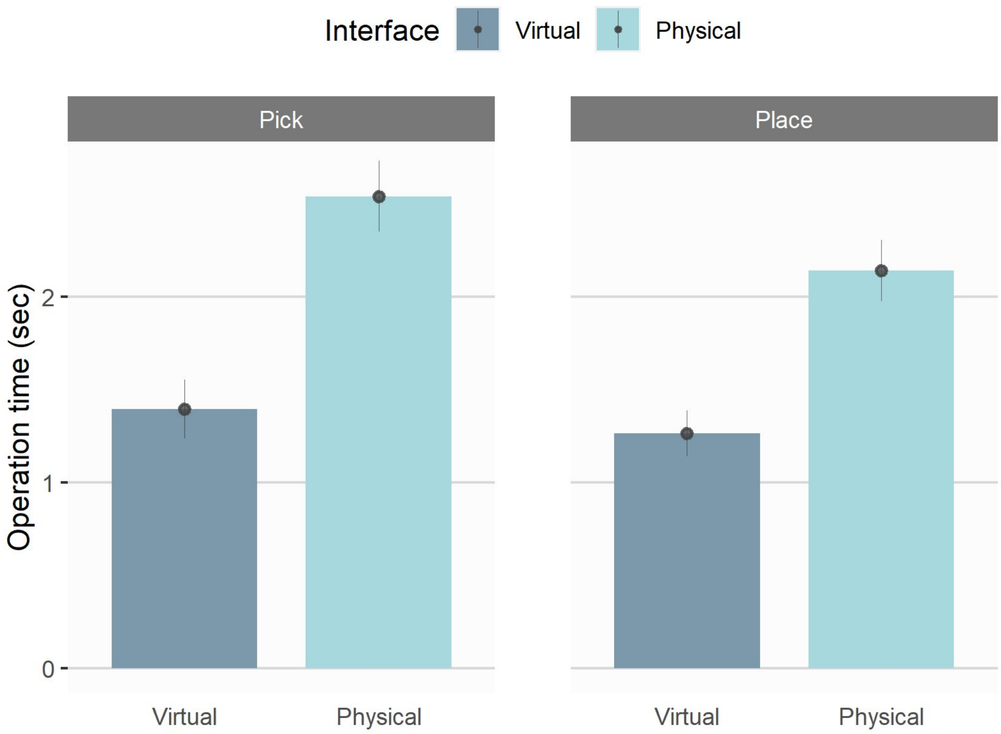

FIGURE 6.4: Operation times in the pick-and-place task

TABLE 6.1: Descriptive statistics of operation time at the pick-and-place task

		Pick (ms)		Place (ms)	
		mean	*SD*	*mean*	*SD*
Task load	Single-task	1.89	0.96	1.67	0.77
	Dual-task	1.91	1.00	1.69	0.81
Interface	Virtual	1.40	0.73	1.26	0.56
	Physical	2.54	0.87	2.14	0.75

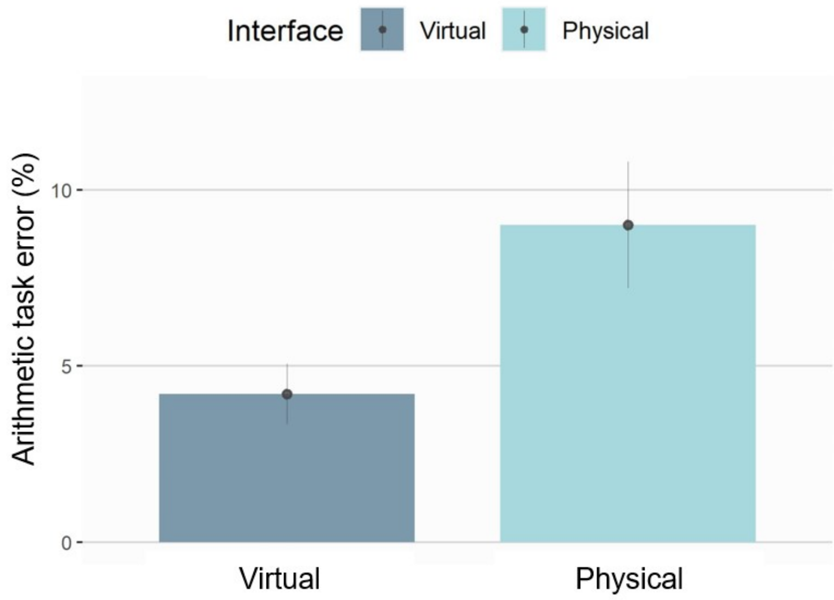

FIGURE 6.5: Error rate at the arithmetic task

TABLE 6.2: Descriptive statistics of the absolute pupil size (mm) in the arithmetic task

	Start		First sum		Second sum		Third sum	
	mean	*SD*	*mean*	*SD*	*mean*	*SD*	*mean*	*SD*
Virtual	3.06	0.37	3.06	0.37	3.10	0.38	3.12	0.38
Physical	2.82	0.94	2.83	0.92	2.86	0.93	2.88	0.93

6.3.2 Pupil size variation

Pick-and-place task. To analyze whether the effects changed in the time course, we specifically focused on interactions involving the window factor. Significant interactions were observed between: task load and window only in the pick phase (X^2 (5, N = 21) = 148.38, p < 0.0001), interface and window (pick: X^2 (5, N = 21) = 442.4, p < 0.0001; place: X^2 (5, N = 21) = 23.72, p < 0.001), and task load, interface and window (pick: X^2 (5, N = 21) = 80.51, p < 0.0001; place: X^2 (5, N = 21) = 34.88, p < 0.0001). Post hoc contrasts that are of interest for the present study are shown in Figures 6.6 and 6.7.

Arithmetic task. The results of pupil size variations in the virtual condition highlighted a significant effect of the arithmetic task (X^2 (3, N = 21) = 893.96, p < 0.0001). Similar results were observed in the physical condition, where there was a significant effect of the arithmetic task (X^2 (3, N = 21) = 97.05, p < 0.0001). Post hoc contrasts run with Bonferroni correction are shown in Figure 6.8, and descriptive statistics are shown in Table 6.2.

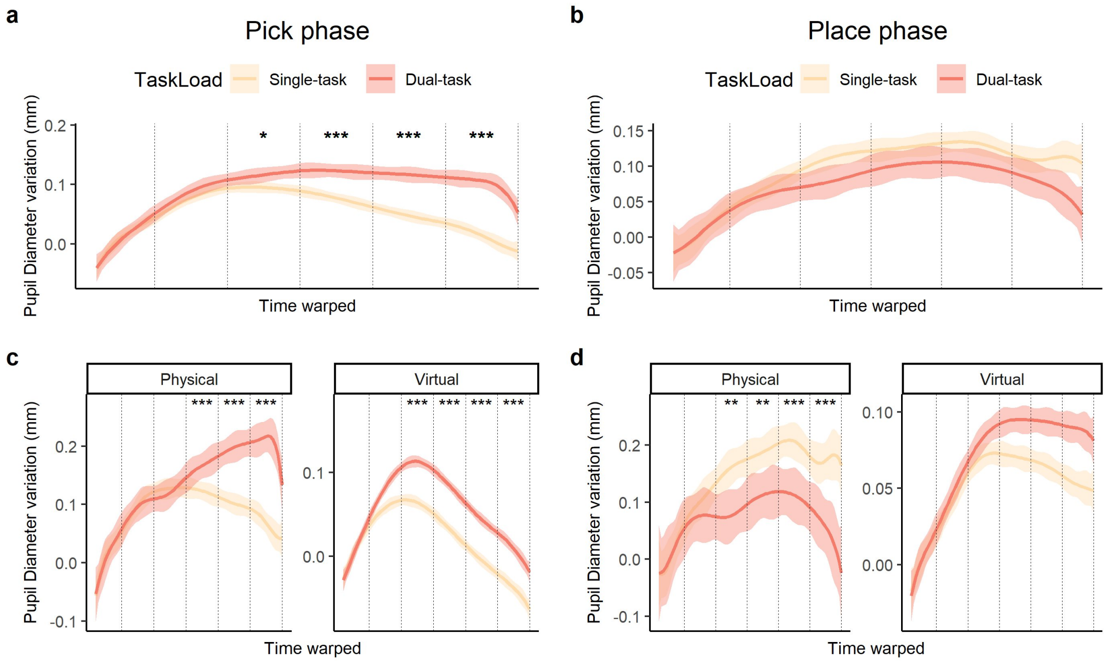

FIGURE 6.6: Pupil size variations relative to the task load conditions in the pick (**a**, **c**) and place (**b**, **d**) phases. Plots **a** and **b** depict the main effect of task load. Plots **c** and **d** display the effects of task load by interface. All the plots are complemented by stars indicating the significance level of the statistical test ($^{*}p \leq .05$;$^{**}p \leq .01$;$^{***}p \leq .0001$)

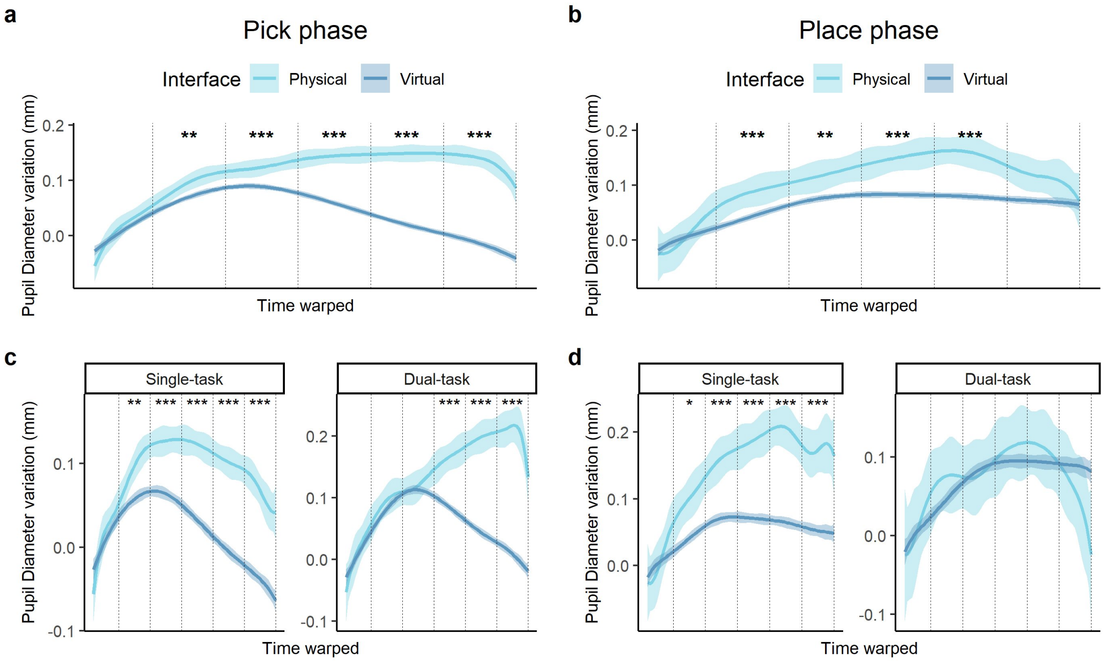

FIGURE 6.7: Pupil size variations relative to the interface conditions in the pick (**a**, **c**) and place (**b**, **d**) phases. Plots **a** and **b** depict the main effect of interface. Plots **c** and **d** display the effects of interface by task load in the pick phase and place phase. All the plots are complemented by stars indicating the significance level of the statistical test ($^*p \leq .05;^{**}p \leq .01;^{***}p \leq .0001$)

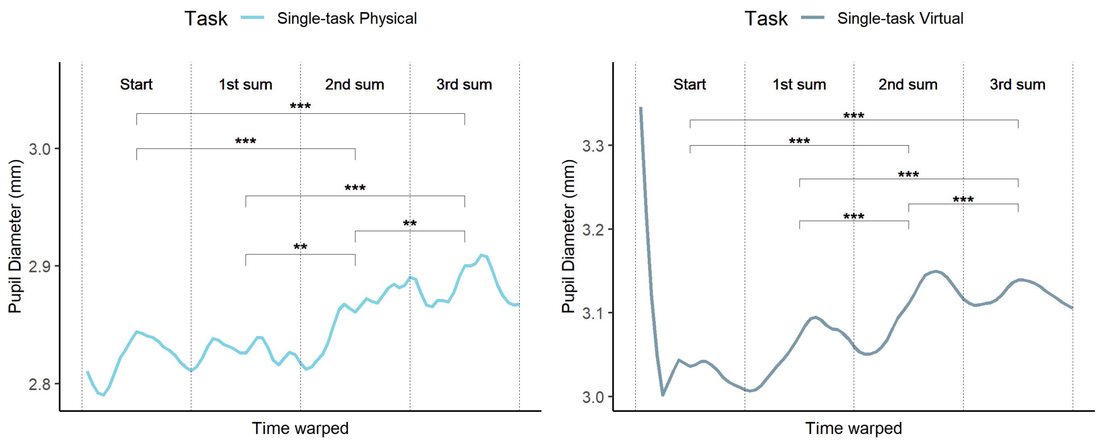

FIGURE 6.8: Pupil diameter in the physical and virtual conditions. The vertical dashed lines divide the plots by windows (start, first sum, second sum, third sum). Plots are complemented by stars indicating the significance level of the statistical test between windows ($^*p \leq .05;^{**}p \leq .01;^{***}p \leq .0001$)

6.3.3 NASA-TLX questionnaire

The results of the linear mixed model (LMM) demonstrated significant effects of task load (X^2 (1, N = 21) = 45.6, p < 0.0001) and item (X^2 (5, N = 21) = 311.79, p < 0.0001) and interactions between task load and item (X^2 (5, N = 21) = 42.04, p < 0.0001) and between interface and item (X^2 (5, N = 21) = 32.3, p < 0.0001). Specifically, a higher NASA-TLX score was reported in the dual-task condition (M = 10.9; SD = 5.09) than in the single-task condition (M = 8.79; SD = 5.82). The post hoc contrasts on the interaction between task load and item revealed a higher NASA-TLX score in the dual-task condition than in the single-task condition in the following items: mental demand (p = < 0.0001; ST: M = 7.85, SD = 5.44; DT: M = 13.00, SD = 4.80), physical demand (p < 0.05; ST: M = 5.85, SD = 4.54; DT: M = 7.67, SD = 4.38), and effort (p < 0.001; ST: M = 9.57, SD = 5.51; DT: M = 13.6, SD = 4.32). Moreover, post hoc contrasts over the interaction between interface and items yield a higher NASA-TLX score in the physical condition than in the virtual condition for the item performance (p < 0.01; virtual: M = 14.5, SD = 4.3; physical: M = 15.8, SD = 4.15) and a higher score in the virtual condition than in the physical condition for the item frustration (p < 0.05; virtual: M = 6.70, SD = 4.24; physical: M = 5.65, SD = 4.35). NASA-TLX score differences for both task load and interface are depicted in Figure 6.9.

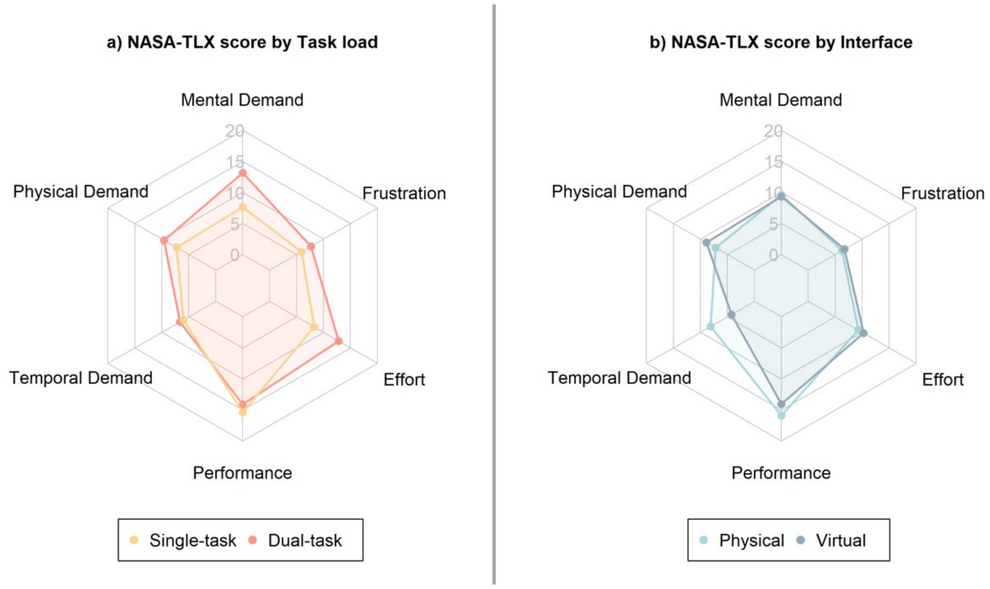

FIGURE 6.9: Averaged NASA-TLX score in each NASA-TLX item for task load (**a**) and interface (**b**)

6.3.4 Individual factors

VR experience. On a scale from 1 to 5, the median VR experience was 2, with a standard deviation of 0.84. In our sample, 3 participants were considered to have high VR experience, as their scaled rating was higher than 0.5, and 18 participants were considered to have low VR experience, as their rating was below 0.5.

Individual preferences for a virtual or physical robot. Finally, individual preferences for virtual or physical robot expressed before and after the experiment are shown in Fig. 10. Before the experiment, 19.05% of participants expressed a preference for virtual robot, but after the experiment, the percentage increased to 61.9%.

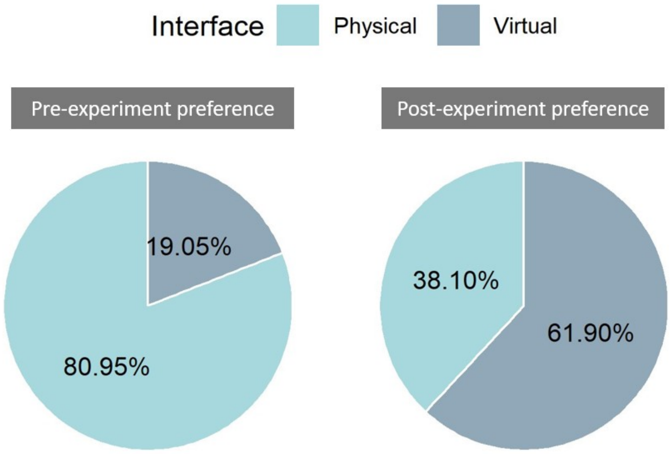

FIGURE 6.10: Pre- and post-experiment preferences for working in collaboration with a physical or virtual robot expressed by participants before and after the experimental session

6.4 Discussion

In this study, we conducted a systematic assessment of human performance and workload in users executing a pick-and-place task with the physical and a virtual reproduction of the collaborative robot UR10e, under low (single-task) and high (dual-task, concurrently with an arithmetic task) mental demands. Specifically, human performance was measured in terms of operation times at the pick-and-place task, and error rate at the arithmetic task, while mental workload was implicitly detected from eye-tracking parameters, and also explicitly referred via self-reports scales. Individual factors such as the participants' experience with VR technology, and the self-reported preferences for collaborating either with a virtual or physical robot were additionally considered. Our research questions are summarized and discussed based on our results below.

6.4.1 Task load manipulation

We expected our task load manipulation to be effective in creating different levels of task demand. Our expectations were confirmed by the pupil size data, and the NASA-TLX questionnaire, even though the dual-tasking did not affect the participants' performance.

Specifically, we observed a gradual increase in pupil size during the arithmetic task, particularly when moving from the first to the last sum. This trend is clearly visible in Figure 6.8, in both the physical and the virtual environments, and suggests that our arithmetic task induced a gradual increase in the implicit workload. Our participants also self-reported higher mental demands, physical demands, and efforts when executing the dual-task compared to the single-task (Figure 6.9). Such an observation supports our hypothesis that our task load manipulation (i.e., dual-tasking) affected the explicit workload as well. On the same interpretation line, a significantly higher pupil size variation was evident in the dual-task compared to the single-task condition particularly in the pick phase (Figure 6.6a). This effect was evident both when operating the virtual and the physical robot and indicates that performing the pick action concurrently with the arithmetic task successfully increased the participants' mental workload implicitly. The absence of such an effect in the place phase might be due to the extreme simplicity of that action.

Furthermore, and unexpectedly, a higher pupil size variation was observed in the single-task compared to the dual-task only in the physical condition (Figure 6.6d), which is in countertendency compared to the other conditions. This effect might be related to different levels of precision required by the two maneuvering actions (pick and place) and/or to the temporality of the same actions. Indeed, it is possible that in the dual-task condition, users employed higher cognitive resources at the beginning of the task for concurrently handling the arithmetic task and initiating the pick action (Figure 6.6c), and they relieved their mental efforts during the subsequent and more rough place action (Figure 6.6d). When participants were performing the same pick-and-place task as a single task, their pupil sizes instead just gradually increased throughout the task. Still, it is interesting to notice how this reverse effect was visible only in the physical condition but not in the virtual condition, where the dual tasking affected the pupil size variation without any influence of the temporality of the actions.

6.4.2 Operators' behavioral performance

We asked whether VR-based operations are faster compared to operations performed with the physical robot, either under low and high task demands (respectively, single- and dual-task). Our findings demonstrated that users saved about 1 s on average in each of the pick-and-place phases when cooperating with the virtual robot (Figure 6.4), which is a considerable time reduction that could significantly optimize manufacturing processes. In this regard, the absence of inertia forces in the VR-based operations led users to perform the operations as freely as if they were unbounded from the robotic arm.

6.4.3 Operators' cognitive state

Finally, we explored whether collaborating with a physical or virtual robot affects the user's mental workload under either low or high task demands (respectively, single- and dual-task). Interestingly, we found that the VR-based operations entail lower implicit workload (as reflected by the pupil size variation) compared to performing the same operations with the physical robot; however, such an advantage was not reflected in the explicit workload (which was self-reported in the NASA-TLX questionnaire).

Specifically, a significantly lower pupil size variation was observed in the virtual compared to the physical environment. This effect was observed throughout the whole pick-and-place task, but it was particularly evident for highly accurate movements (namely, the pick phase; Figure 6.7a). Indeed, the higher the task complexity (pick vs. place), the more the virtual simulation revealed to be preferable, because it allowed the user to save mental resources. This would translate in a greater benefit of VR-based system for highly complex or demanding work environments, where a higher risk of accidents is also involved. In this view, the lower implicit workload related to virtualization will enhance users' safety and well-being. Besides complex and potentially hazardous telerobotics contexts, such an advantage can actually be beneficial in all human-robot collaborative frameworks, whether dangerous or not. Indeed, it is always essential to ensure that the user is able to maintain high vigilance and awareness at a minimum workload for avoiding mental and physical safety issues (Matsas et al., 2018).

Interestingly, the advantage of the VR-based interface over the physical robot, which clearly emerged from the pupil size variation data and from the performance data, only partially emerged from the NASA-TLX questionnaire. Participants reported a generally higher NASA-TLX score when guiding the physical compared to the virtual robot. However, regarding the single questionnaire's dimensions, users self-rated their own performances as better, and also reported lower frustration, when using the physical robot compared to the virtual one. This might be due to the participants' limited VR expertise: performing operations in VR for the first time might generate uncertainties, likely leading participants to question the quality of their performances and possibly feel frustrated.

Another argument was proposed by Kaufeld and Nickel (2019), who explained the mismatch between performance and workload data according to the compensatory control model: users likely adjusted their task performance strategies by shifting to simpler or less precise procedures, which consequently lowered their mental workload Hockey (1997). Similar dynamics probably occurred in our scenario, where the performance worsening observed when participants were working with the physical robot might have reduced the human mental workload and thus mitigated the level of perceived mental demand. One important implication of such a finding is that questionnaire-based cognitive evaluations might not be sufficient because what the human being consciously perceives does not always accurately reflect actual activation. This speaks in favor of multidimensional evaluations involving implicit indexes of mental workload combined with performance and subjective ratings.

6.4.4 Individual factors

Before the experiment, only 19.05% of participants expressed a preference for driving the robot in VR rather than using the physical one. Notably, such a preference increased to 61.9% after participants used both interfaces for the duration of the experimental session. This is promising data, suggesting a likely positive acceptance of VR-based telerobotics. Furthermore, our results apply to VR users with relatively low experience who did not have any knowledge of or experience with robots. This thus suggests that even non-experts can benefit from the advantages of employing virtual simulations in human-robot frameworks.

6.5 Conclusions

6.5.1 Limitations

First, two different eye-tracking systems were employed in the two conditions: the PupilLabs device was deployed in the physical condition, while the Tobii eye-tracking system integrated into the HTC Vive Pro Eye was deployed in the virtual condition. Therefore, even though we applied a proper baseline correction (Mathôt, 2018), there might still be slight differences in the proprietary algorithms used by different systems to acquire eye-tracking data.

Second, on the applicability of our findings in the field, it is important to mention that the effectiveness of a virtual simulation depends not only on the realism of the virtual environment but also on the quality of the computerized tools employed. For instance, technical features of the head mounted display such as the visual field of view, resolution, and latency of the graphical interface might influence the user's performance with the virtual robot. Therefore, in view of a large-scale implementation of virtual devices into the industrial domain, highly immersive and advanced virtual device adoption is fundamental.

Third, even though we tried to leave out any differences between the virtual and the physical pick and place actions, some procedural differences between the two task flows were still present. For instance, the arithmetic sums were reported verbally in the physical condition, whereas they were reported on a virtual keyboard in the virtual condition. Moreover, the pick and place actions were initiated via controller buttons in VR, whereas they were initiated by the user physically touching the worktable in the physical condition.

Finally, our conclusions were gleaned from a pool of young users with relatively little experience. Although this choice was motivated by the desire to maintain a homogeneous sample, it comes at the cost of possible low generalizability.

6.5.2 Future directions

Considering that active industrial operators usually fall within a much greater age range, further investigations might increase the experimental sample and include both young and senior users. In this way, it would be possible to understand better whether the advantages of the virtual simulation also extend to older people. Additionally, based on the finding that even a scant knowledge of VR devices is sufficient for revealing the advantages of virtual simulations, it would be interesting to test whether these advantages increase as the VR experience also grows. Another crucial point arises when addressing the task complexity. The choice of such an easy task, such as the pick-and-place one, was intentional to allow a highly naturalistic investigation without constraining the users' actions and at the same time ensure good experimental control. A systematic assessment of increasingly demanding tasks would also provide relevant knowledge on the applicability of pupil size variation as an implicit workload index in different levels of task complexity. If future research proves pupillometry to be a reliable and flexible index of implicit workload—in either virtual or physical environments or both—it would become feasible for systems to auto-adjust the robot's behavior based on human pupil responses. Furthermore, by specifically focusing on VR-based solutions for telerobotics, different control systems might be implemented (e.g., via buttons, via physical interactions). However, which control system might be preferable for workers in terms of human perforance and workload is still unknown. Future research might dig into these aspects and clarify pros and cons of the diverse interaction and control systems offered by the VR devices for telerobotics.

Overall, this research has just started to shed light on the potential of virtual simulations within human-robot collaborative frameworks. With the introduction of VR devices in the industry, the design, validation, training, and even active operations on robotic systems can definitely take a turn for the better, with the humans' mental and physical health being the cradle of faster and safer interactions between humans and robots.

Chapter 7

Study 2 - Human-centric telerobotics: investigating users' performance and workload via VR-based eye-tracking measures

VR is gaining ground in the robotics and teleoperation industry, opening new prospects as a novel computerized methodology to make humans interact with robots. Compared to traditional teleoperation means (e.g., desktop, mouse, keyboard), immersive VR devices allow users to physically act in the virtual surroundings and, potentially, to use their own body movements to guide robots and machines (Martín-Barrio et al., 2020). Such feature would basically reproduce the manual guidance feature of collaborative robots, which allows humans to use their own hands to drive a robotic system over desired position. In this way, the amount of time typically spent on the physical robot programming is significantly streamlined and reduced (Faccio et al., 2022). However, little is known about how an action rather than button-based control system of a virtual industrial robot impact users' performance and workload. This is due to the reasonable novelty of VR-based robotic teleoperations, but also to the lack of multidisciplinary and human factor-based investigations in telerobotics. Furthermore, the latest VR devices are also equipped with integrated eye-tracking, which constitutes an exceptional opportunity for monitoring users' workload online. Nonetheless, such devices are fairly recent, and the above-mentioned lack of human factors studies in VR-based telerobotics has left a knowledge gap in the understanding of their sensitivity to workload fluctuations.

We thus covered these aspects by analyzing extensive behavioral data generated by 24 participants driving a simulated industrial robot in VR through a pick-and-place task. Users drove the robot via button-based and action-based controls and under low (single-task) and high (dual-task) mental demands. We collected self-reports, performance and eye-tracking data. Specifically, we asked how the interactive features of VR, which entails higher degree of physical motion, affect users' performance and workload. Additionally, we tested the sensibility of diverse eye parameters in monitoring users' vigilance and workload throughout the task. Such research questions are further deployed in the following section.

7.1 Hypotheses and research questions

7.1.1 Task load manipulation

The dual-task paradigm employed in this work was previously tested with a physical and a virtual replica of the industrial robot UR10e unfolded in the Study 1 discussed in the previous Chapter 6 (Nenna et al., 2022a). As opposed to the previous paradigm, we here transposed the secondary arithmetic task in the visual domain to increase the task difficulty in the dual-task condition: according to resource theories, dual-task costs are particularly evident when multiple concurrent tasks share the same resources (Pashler, 1984). Therefore, participants needed to visually pay attention to the numbers to sum at the arithmetic task while also being engaged in precise visuomotor coordination for executing the pick-and-place task. As a methodological control, we thus expect the dual-task to affect participants' performance (slower operation times and higher error rates) and workload (higher NASA-TLX score, higher pupil size variation, lower perclos, and shorter and fewer blinks in the dual-task compared to the single-task).

7.1.2 The impact of diverse control systems on the user

The core of our work revolves around the following research question: how the degree of interactivity of a control system affects users' performance and workload in virtual teleoperations? In contrast with the widespread button-based controls, action-based controls fully emphasize the opportunities offered by an interactive technology such as VR, and basically reproduce the same manual manipulation feature of collaborative robots in a virtual workspace. Therefore, by leveraging the use of natural and embodied controls, we hypothesize that an action-based system potentially represents a more efficient, intuitive, and less demanding solution for guiding robots compared to a button-based one, even in VR. Literature that served as a base to this hypothesis can be found in Chapter 2, Section 2.2.

7.1.3 VR-embedded eye-tracker sensitivity to workload

Eye-trackers-equipped VR headsets have the potential of capturing workload-related fluctuations in eye parameters while freely acting in the virtual environment. In this view, a relevant question is which of the eye parameters collected via the VR headset is the most sensitive to workload changes, also considering the higher degree of motion during VR interactions. To address this question, we inspected the relations between the self-reported workload and each of the investigated eye parameters (pupil size, PERCLOS, blink frequency and duration) both when using the action-based and button-based controls. We assume higher correlations for those eye parameters that are more sensitive to workload.

7.2 Methods

7.2.1 Sample

As suggested by the power analysis conducted on Gpower (Erdfelder et al., 1996), for our within-subjects experimental design, a total sample of 21 participants was needed to detect a medium effect size (d = 0.5) with 80% power. 24 participants, 11 females and 13 males (M_{age}= 26.15; SD_{age} = 1.85), voluntarily took part in the experiment after signing informed consent. They all reported having normal or corrected-to-normal visual acuity (via contact lenses), normal color vision, and no current or past neurological or psychiatric problems. The experimental protocol was approved by the local ethics committee, and the study was conducted following the principles of the Declaration of Helsinki. One participant was excluded for technical issues of the eye tracker, while one more participant withdrew from the experiment because was reporting visual difficulties in VR. Finally, three participants were excluded from the analysis for having committed more than 50% of errors in the arithmetic task. The final sample comprised 18 participants, 9 females and 9 males (M_{age} = 26.33; SD_{age} = 2.02).

7.2.2 Technical setup

An HTC Vive Pro Eye headset (resolution: 1440 x 1600 pixels per eye; refresh rate: 90 Hz; Field of view: 110°) was connected to an MSI laptop (model GT63 Titan 8RF, processor Intel Core i7-6700HQ, RAM 16Gb). This model of head-mounted display is the same employed in the Study 1 (Chapter 6) and has an eye-tracking system embedded (sampling frequency: 120 Hz; calibration: 5-points) which allows continuous recording of eye parameters. Based on the present research questions, the virtual environment (Figure 7.1) has been re-arranged from the one developed for Study 1 (Chapter 6), which was programmed in Unity (version 2019.4.18f1). At the end of each experimental session, a large datalog was automatically saved on the MSI laptop, which included time series data of position and rotation of the VR headset and controllers, of all interactions between the user and the robot, and of the accuracy of the performed pick-and-place task.

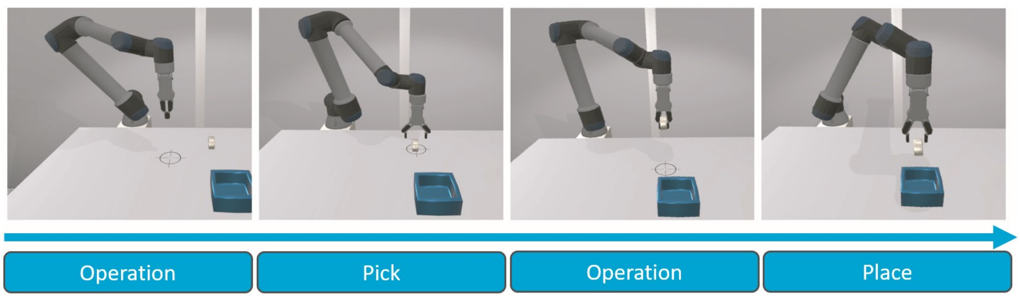

FIGURE 7.1: Overview of the pick-and-place task

7.2.3 Procedure, experimental tasks and design

All participants signed informed consent before starting the experiment. Thereafter, they filled out questionnaires about their demographics, VR expertise and general preferences for virtual robot control systems. Specifically, we asked: "If you had to guide a robotic arm in VR, which of the following control modalities would you prefer?". The possible answers were "Controller buttons" and "Physical actions". Afterward, all participants underwent a training session to familiarize with the tasks, in which they performed a few trials of each of them. All instructions were presented in text format in the virtual environment. Once the participant reported having understood all the tasks, a 5-point calibration of the eye-tracking system was conducted, and the experiment started.

During the experiment, participants performed 5 tasks composed of 40 trials each that were re-adapted from previous research (Study 1, Chapter 6): (1) an arithmetic task, (2) a pick-and-place task executed both via controller buttons (button-based control systems), (3) and physical actions (action-based control system), (4) and a dual-task performed via controller buttons (button-based control systems), (5) and physical actions (action-based control system) as well. These tasks were presented in a random fashion, and a NASA-TLX questionnaire (Hart, 2006) was administered at the end of each task. Participants could also claim a break after each NASA-TLX questionnaire; in that case, the eye-tracking system was re-calibrated before starting the next task. Once participants finalized all the tasks, the final question on the general preferences for virtual robot control systems was administered again ("With reference to the experience you have just concluded, which of the following control modalities did you prefer?") and the experiment ended.

In the **arithmetic task** (1), participants mentally summed a series of four numbers presented in text format in the virtual environment. The numbers appeared on a virtual panel that always followed the participant's head movements and was placed in the upper part of his/her view, in a way that it could not cover the work table nor the robot's effector and was always inside the participant's functional field of view. Thereafter, they reported the result of the arithmetic operation in a virtual keyboard by using the controller buttons. The presented numbers randomly ranged between 1 and 10 and a time interval of 2.5sec ± 0.3sec intercurred between them.

For the **pick-and-place task**, we used the same paradigm employed in the previous study (Chapter 6), in which participants were asked to guide the robotic arm to pick a bolt from the workstation and place it into a box (Figures 6.3 and 7.1). While the task design remained unchanged, we here introduced two different control systems: in the button-based pick-and-place task (2), participants used the pad button of the right controller to move the robot on the left-rightfarward-backward directions over the work table; in the action-based pick-and-place task (3), instead, they were allowed to reach the virtual robot with their right hand, grasp it by pressing the grip button on the right controller and then move it to the desired position by simply moving their own arm. The latter condition thus reproduced the direct manipulation feature of cobots. In both conditions, once the robot was placed in the right location, participants pressed the pad button on the left controller for picking or placing the bolt. Therefore, they used the VR controllers with both control systems, but only in the action-based condition they were allowed to interact with the virtual robot physically.

Finally, in the **dual-task**, the pick-and-place task and the arithmetic task were concurrently performed, once using the button-based (4) and once using the action-based (5) control systems. The series of numbers presented for the arithmetic task covered the whole pick-and-place task duration, and the result was reported only after the bolt was placed into the box. All task conditions are depicted in Figure 7.2.

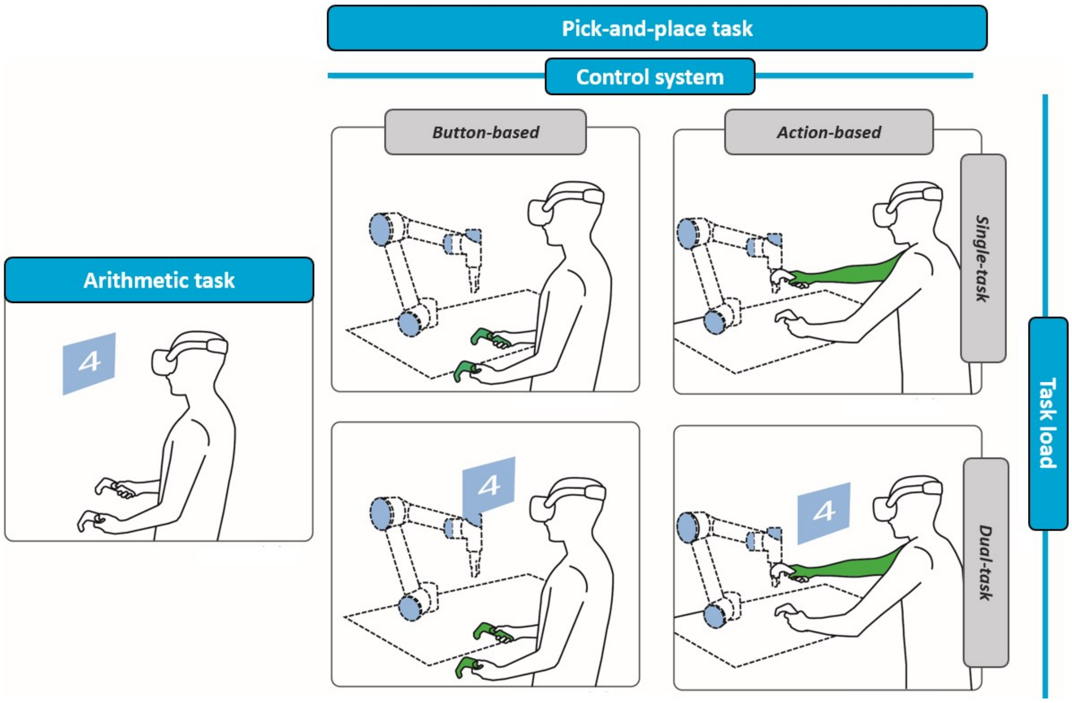

FIGURE 7.2: Experimental design

7.2.4 Measurements

Performance measures

Pick-and-place task. We measured the operation times as the time elapsed during the robot's movements (start: first movement of the robot; end: last movement before the pick/place action). Considering that the pick action required higher precision for aligning the robot effector with the bolt to pick as compared to the place action, users' performance was analyzed independently in the pick and in place phases. Trials whose duration exceeded 4 SD from the averaged duration were removed as they represented very unrealistic operation times (pick phase: 1.43% removed, range 13.45sec-46.99sec;

place phase: 0.08% removed over 2.827, range 14.91sec-42.94sec). The same trials were not considered for the analysis of the other independent measures either. Additionally, the error rate at the pick-and-place was measured independently for the pick and the place phases. Particularly, the pick and place automation were executed only if the left pad button was pressed while the robot was perfectly positioned above the bolt in the pick phase, and above the box in the place phase ("correct" event). If at the first attempt of button press the robot was not in line with the bolt/box, the event was registered as "incorrect". The participant thus had to relocate the robotic arm in the right position to initiate the automation. The percentage of "incorrect" events registered for each action informed on the error rate at the pick and at the place actions.

Arithmetic task. In the arithmetic task, we measured the arithmetic input time as the time elapsed from the end of the last number presentation to the moment the participant sent the result of his/her arithmetic calculation through the controller. This measure can be informative of the mental effort deployed for finalizing the mental calculations in each condition. When computing it, we only considered trials in which the correct sum was sent. The reason behind this choice is that in case participants lost count of the sum during the arithmetic task, they often quickly insert a casual number, thus resulting in casual input time. Differently, trials in which participants sent the correct result are more likely to be the product of meaningful cognitive processes. Finally, the error rate at the arithmetic task was computed as the percentage of wrong sums reported, allowing an understanding of the dual-task-induced interference with the main task (pick-and-place task).

Eye-tracking measures

Based on previous studies (Guo et al., 2021; Nenna et al., 2022a; Novak et al., 2015; Wu et al., 2020), we computed pupil size variation, perclos, blink frequency and duration as workload-related eye-tracking indexes. Notably, we do not intend to directly compare the control system-related eye parameters as they might be strongly influenced by the different movement magnitude involved in the action and button-based conditions. In this respect, systematically observing how the eye parameters are affected by dual-tasking when using action- rather then button-based controls can provide important insights on the resources deployed dependently from the control system involved.

Pupil data preprocessing followed the same procedure used in the previous study (Nenna et al., 2022a). For more details, please refer to Chapter 6, Section 6.2.4. For the pick-and-place task, variations in pupil size were analyzed within the pick and the place actions independently, whose length was standardized via dynamic time warping to fit 30 data points. For the arithmetic task, instead, we selected four time windows, one for each number presented, and applied dynamic time warping to standardize their length (which ranged between 2.3 and 2.7sec). Additionally, we used the eye openness data stream outputted from the HTC Vive headset for computing perclos and blinks. Perclos was calculated as the percentage of time during which the eyelids covered pupils by more than 80% of their area (Wu et al., 2020) in four time windows, each including 10 trials. Blinks were detected as eye closures lasting a minimum of 70ms and a maximum of 500ms (Benedetto et al., 2011). If eyes were closed for less than 70ms, it was considered a technical issue of the eye-tracker that likely lost pupil tracking for some frames (Faure et al., 2016). Blink frequency was operationally defined as the blink rate per minute.

Self-reports measures

The NASA-TLX questionnaire (Hart, 2006) was administered after each task as a measure of self-reported workload. Once before starting the experiment and once in the end, we additionally administered a question asking to express the individual preference for guiding a robot in VR either via controller buttons or physical actions. With these questions, we intended to measure whether the individual preferences for one or the other control system would change after having tested both the button-based and the action-based control systems.

7.2.5 Statistical analysis

Performance measures

For analyzing performance data, we used Generalized Linear Models (GLMs) from lme4 package (Bates et al., 2014) in RStudio (Team, 2021). Data were first fitted through the function descdist() of the package fitdistrplus (Delignette-Muller and Dutang, 2015), then the appropriate models were chosen accordingly to data distribution. For each of the performance measures at the pick-and-place task, we computed a GLM including the factors Task load (single-task, dual-task) and Control System (button-based, action-based) with Participant as a random effect. For analyzing performance measures at the arithmetic task, instead, we run a GLM over the factor Task (single-task, button-based dual-task, action-based dual-task). The Bonferroni correction (Bonferroni, 1936) was always applied when interpreting the post hoc contrasts within the significant interactions.

Eye-tracking measures

GLMs from lme4 package (Bates et al., 2014) were also used for the analysis of eye-tracking data in RStudio (Team, 2021). Specifically, each model was chosen based on data distribution (Delignette-Muller and Dutang, 2015). Models analyzing pupil size variation during the pick-and-place included the factors Task Load (single-task, dual-task), Control System (button-based, action-based) and Window (1, 2, 3, 4, 5, 6) with Participant as a random effect. The factor Window allowed to consider pupil size variation changes in the time course on trial level. When analyzing the pupil size variation throughout the arithmetic task, instead, we ran a model including the factors Task (arithmetic task, button-based dual-task, action-based dual-task) and Arithmetic operation (start, 1st sum, 2nd sum, 3rd sum), with Participant as a random effect. For this analysis, we only considered the first 3 arithmetic operations in order to compare the Single-task with the Dual-tasks. Therefore, the analysis of pupil size variation at the arithmetic task was conducted at trial level too. Differently, the statistical models analyzing perclos, blink frequency and duration included the factors Task Load (single-task, dual-task), Control System (button-based, action-based) and Window (1, 2, 3, 4) with Participant as a random effect. Each window included 10 trials (window 1: trials 1-10; window 2: trials 11-20, etc.) and allowed to look into eye parameters' changes in the time course on task level.

Self-reports measures

The analysis of the NASA-TLX questionnaire was conducted through a GLM that included the following factors: Task Load (single-task, dual-task), Control system (button-based, action-based), and Items (mental demand, physical demand, temporal demand, performance, effort, frustration), with Participant as a random effect. Post hoc contrasts were performed specifically between the levels of Task load and Control system in each of the questionnaire's items, with the application of the Bonferroni correction (Bonferroni, 1936).

Correlations between self-reported workload and eye-tracking metrics

Relations between the NASA-TLX score (both the overall score and the score at the individual NASA-TLX items) and each of the eye-tracking measures were also assessed via Pearson's linear correlation tests. Furthermore, we reported the response rate of the individual preference for action- vs. button-based control systems expressed before and after the experiment.

7.3 Results

7.3.1 Performance measures

Pick-and-place task. The GLM conducted on the operation times demonstrated significant main effects of Task Load only in the pick phase (X^2 = 20.02, p<.0001), but not in the place phase (X^2 = 20.02, p=.053), while a main effect of Control System was observed both in the pick (X^2 = 1462, p<.0001) and in the place (X^2 = 1976.7, p<.0001) phases. Additionally, interaction effects between Task Load and Control System were observed only for the place (X^2 = 10.64, p<.01) but not for the pick action (X^2 = 0.56, p=.45). Post-hoc contrasts revealed significant differences between the button- and action-based control systems both under single- (p<.0001) and dual-task (p<.0001). After applying the Bonferroni correction, differences between the single-and dual-task were not significant in any of the control system modalities. When analyzing the pick-and-place error rate, results demonstrated significant main effects of Task Load (X^2 = 5.91, p<.05) and Control System (X^2 = 22.27, p<.0001) only in the pick but not in the place phase. No interaction effects were observed. All performance results are depicted in Figure 7.3 and summarized in Table 7.1.

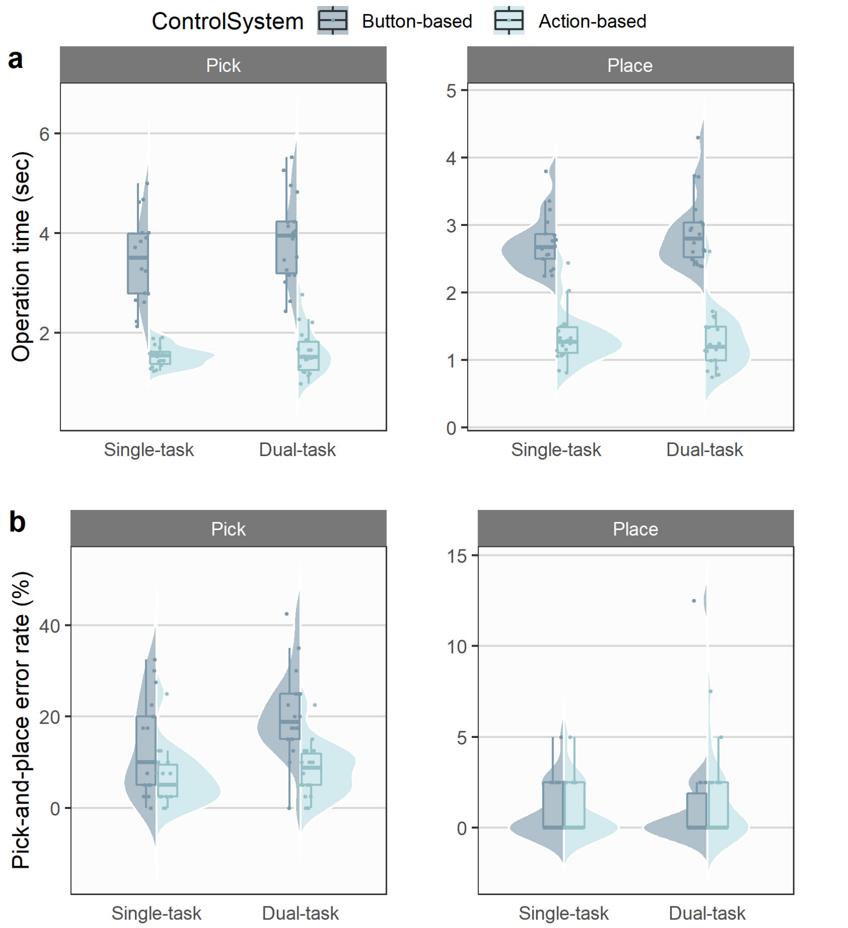

FIGURE 7.3: Performance results at the pick-and-place task. Figure **a** shows the averaged operation times, Figure **b** depicts the error rate at the task.

Arithmetic task. The GLM on the arithmetic error rate indicated a significant main effect of the factor Task (X^2 = 14.58, p <.001). Post-hoc contrasts revealed that, compared to the single arithmetic task, the error rate was significantly higher only while executing the pick-and-place task via button-based (p<.01) but not via action-based control system (p=.16). Moreover, the error rate at the arithmetic task did not differ significantly between the two dual-tasks (p=.39). For the analysis of the arithmetic input time, instead, the main factor Task resulted in being statistically significant (X^2 = 146.04, p<.0001). Post hoc contrasts revealed that the arithmetic input time was significantly lower in the Single-task compared to both the Dual-task conditions

TABLE 7.1: Descriptive statistics of the performance at the pick-and-place task

		Operation time (sec)				Error rate (%)			
		Pick		Place		Pick		Place	
		mean	*SD*	*mean*	*SD*	*mean*	*SD*	*mean*	*SD*
Task load	Single-task	2.51	1.88	2.04	1.18	9.93	9.03	0.97	1.50
	Dual-task	2.73	2.00	2.11	1.28	14.2	9.91	1.32	2.57
Control system	Button-based	3.67	2.07	2.83	1.07	16.8	10.4	1.11	2.35
	Action-based	1.57	1.57	1.31	0.85	7.36	5.91	1.18	1.84

(ps<.0001). Moreover, a significantly higher arithmetic input time was observed in the button-based dual-task condition compared to the action-based dual-task condition (p<.0001). Performance results at the arithmetic task are depicted in Figure 7.4 and summarized in Table 7.2.

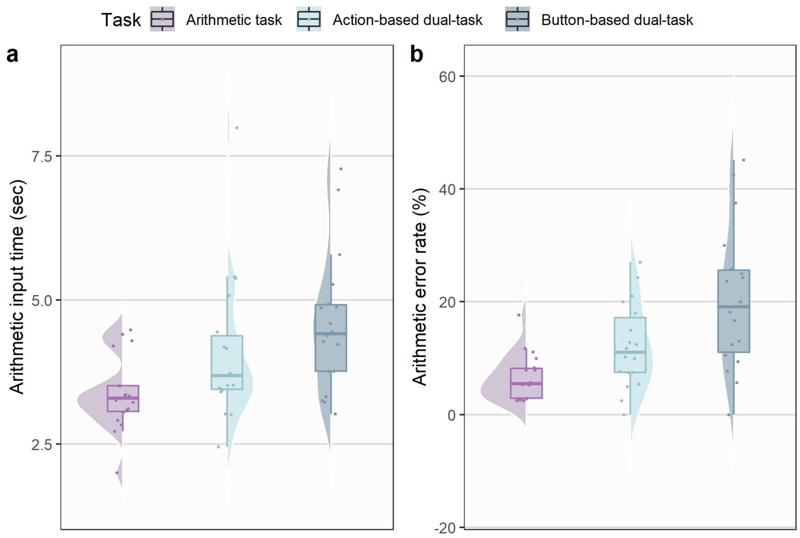

FIGURE 7.4: Performance results at the arithmetic task. Figure **a** shows the averaged input times, Figure **b** depicts the error rate at the task.

7.3.2 Eye-tracking measures

Pupil size variation

Arithmetic task. As depicted in Figure 7.5, the analysis of pupil size at the arithmetic task yielded significant main effects of both Task (X^2 = 1553.6, p <.0001) and Arithmetic operation (X^2 = 1017.3, p <.0001). Furthermore, two factors interacted significantly (X^2 = 1346.3, p <.0001). Post-hoc contrasts Bonferroni-corrected showed a significant increase in pupil size when moving from Start to 2nd sum only for the button-based dual-task (p < .001) and from Start to 3rd sum for all task conditions (all ps <.001). Similarly, significant pupil size increases were observed when moving from the 1st sum to 2nd for the arithmetic task (p<.01) and the button-based dual-task (p<0001),

TABLE 7.2: Descriptive statistics of the performance at the arithmetic task

		Error rate (%)		Input time (sec)	
		mean	*SD*	*mean*	*SD*
Task	Arithmetic single-task	6.55	4.10	3.37	1.49
	Action-based dual-task	12.1	7.52	3.95	2.16
	Button-based dual-task	20.4	12.6	4.38	2.28

and from 1st sum to 3rd sum for all tasks (all ps <.0001). When moving from 2nd to 3rd sum, only the dual-task conditions yielded significant contrasts (all ps <.0001).

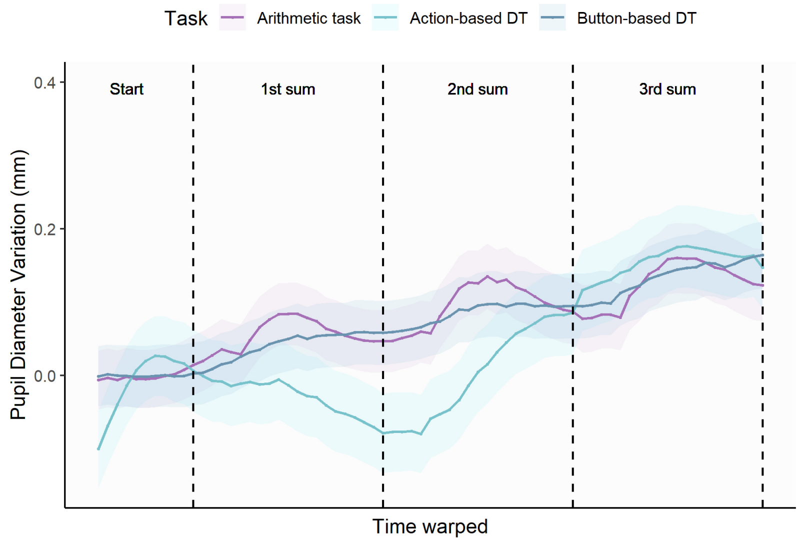

FIGURE 7.5: Pupil size variation throughout the arithmetic task (trial level)

Pick-and-place task. The statistical model yielded significant interactions between Window and both Task Load (pick: X^2 = 981.76, p<.0001; place: X^2 = 382.18, p<.0001) and Control System (pick: X^2 = 1884.75, p<.0001; place: X^2 = 819.96, p<.0001), and between Window, Task Load and Control System (pick: X^2 = 52.20, p<.0001; place: X^2 = 66.23, p<.0001). Post hoc of interest included the comparison between Single- and Dual-task in each Control System condition and within each Window. Specifically, pupil size variation was significantly higher in the dual-task compared to the single-task in both Control System conditions and from window 2 to 6 specifically in the pick phase (all ps < .0001) and in the button-based condition of the place phase (all ps < .0001). Differently, in the action-based condition of the place phase, pupil size variation was higher in the dual-task compared to the single-task in windows 4 (p<.01), 5 and 6 (ps < .0001). Results of pupil size variation are depicted in Figure 7.6.

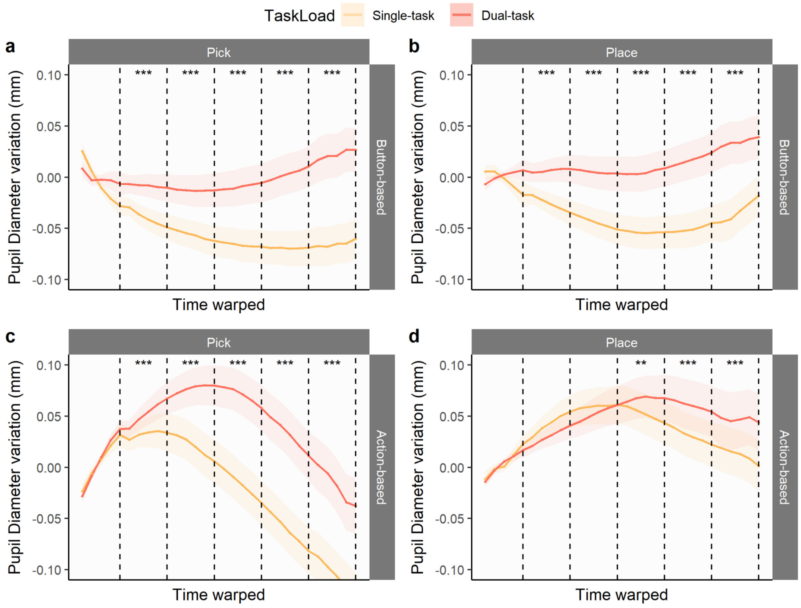

FIGURE 7.6: Pupil size variation throughout the pick-and-place task (trial level). The first row depicts the pick (**a**) and place (**b**) phases of the button-based condition. The second row depicts the pick (**c**) and place (**d**) phases of the action-based condition. All the plots are complemented by stars indicating the significance level of the statistical test ($^{*}p \leq .05$;$^{**}p \leq .01$;$^{***}p \leq .0001$)

Perclos

A main effect was observed only for Task Load (X^2 = 16.55, p<.0001). Interactions with the factor Window did not reach the significance threshold not for Task Load (X^2 = 0.27, p=.96) nor for Control System (X^2 = 2.06, p=.55). Only the interaction between Task Load and Control System was statistically significant (X^2 = 4.71, p<.05), with post hoc contrasts revealing significant differences between single- and dual-task only in the action-based condition (p<.0001). Results on the perclos are depicted in Figure 7.7.

Blink parameters

The analysis of blink duration yielded a significant main effect only for the factor Window (X^2 = 14.2, p<.01), while Task load (X^2 = 1.29, p=.25) and Control system (X^2 = 3.05, p=.08) were not significant. A significant interaction effect was observed only between Task load and Control System (X^2 = 6.69, p<.01); however, after applying the Bonferroni correction, none of the contrasts reached the significance threshold. Differently, when analyzing blink frequency, the factors Task load (X^2 = 12.32, p<.001) and Window (X^2 = 31.79, p<.0001) were demonstrated to be statistically significant. Specifically, higher blink frequency was observed in the single task (M = 3.99, SD = 3.7) compared to the dual-task (M = 2.41, SD = 2.43). Furthermore, a significant interaction effect was observed between Task load and Control system (X^2 =

10.52, p<.01), with higher blink frequency observed in the single- compared to the dual-task only in the action-based (p<.001) but not in the button-based condition (p=0.54). Results on blink parameters are shown in Figure 7.7.

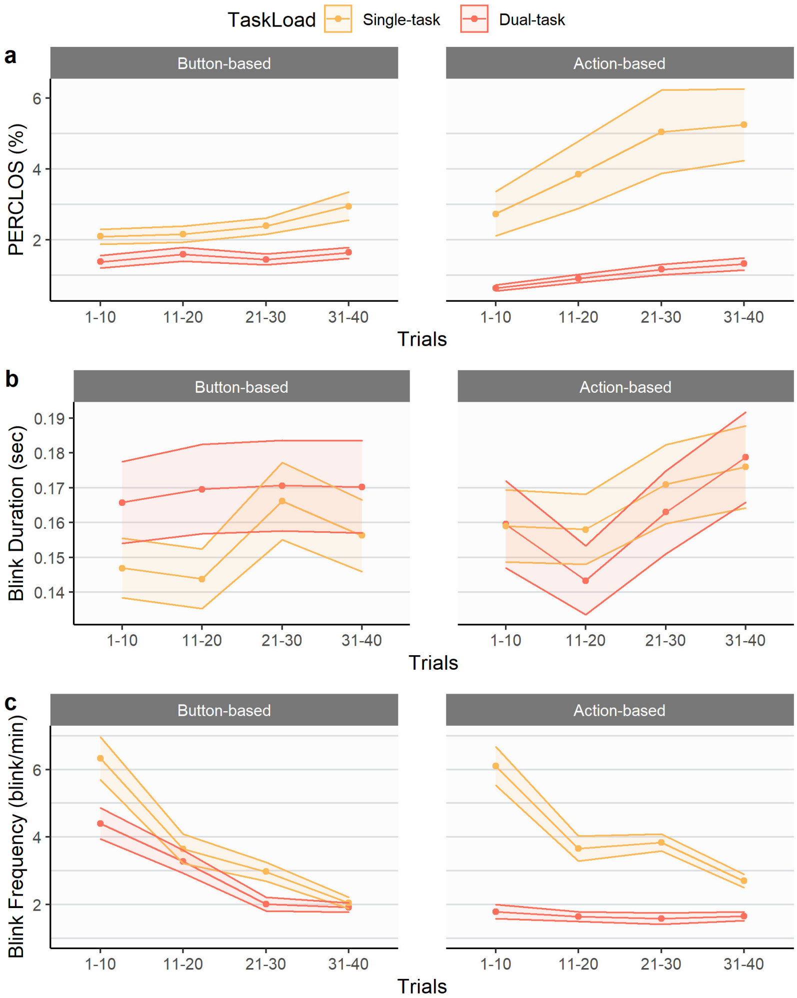

FIGURE 7.7: Perclos (**a**), blink duration (**b**) and blink frequency (**c**) throughout each task condition (task level)

7.3.3 Self-report measures

NASA-TLX questionnaire. Results on the NASA-TLX questionnaire, as depicted in Figure 7.8, yielded a significant main effect of both Task Load (X2 = 212.99, p<.0001) and Control system (X2 = 12.79, p<.001). Significant interaction effects were also observed between Item and Task Load (X2 = 47.32, p<.0001) and Item and Control system (X2 = 11.10, p<.0001). Post hoc contrasts revealed significant differences between Single- and Dual-task for the following items: Mental Demand (p<.0001), Temporal Demand (p<.0001), Performance (p<.05), Effort (p<.0001), and Frustration (p<.0001). Differently, significant differences between the button-based and action-based conditions were only observed for the item Frustration (p<.01).

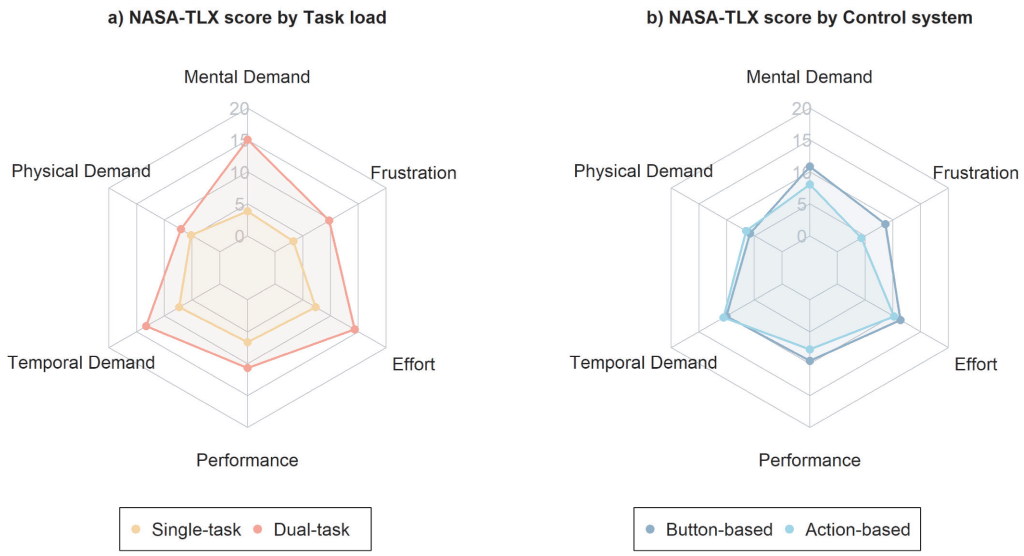

FIGURE 7.8: Averaged NASA-TLX score in each item according to the Task load (**a**) and Control system (**b**)

Individual preferences for button- vs. action-based control systems. Finally, individual preferences for button- or action-based control systems expressed before and after the experiment are shown in Figure 7.9. Before the experiment, 73.68% of participants expressed a preference for guiding the robot via physical actions, and after the experiment, their percentage increased to 89.47%.

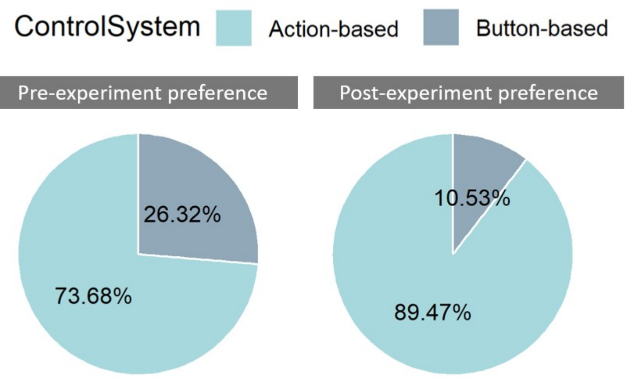

FIGURE 7.9: Pre- and post-experiment preferences for operating through action- and button-based Control systems expressed by participants before and after the experimental session

7.3.4 Correlations between self-reported workload and eye-tracking metrics

Results of the correlation matrix are reported in Figure 7.10. Independently from the level of interactivity, we observed that blink duration had only a significant inverse relation with the self-reported mental demand (R = -0.28, p = .026). Similarly, perclos showed an inverse correlation with the overall NASA-TLX score (R = -0.31, p = .013), and also with mental demand (R = -0.25, p = .042) and effort (R = -0.31, p = .011). Pupil size additionally demonstrated a positive correlation with the overall score at the NASA-TLX (R = 0.28, p = .022), and the single following dimensions: mental demand (R = -0.36, p = .003), temporal demand (R = 0.29, p = .021) and effort (R = 0.31, p = .011). It is to notice that all the latter were weak relations.

Differently, when differentiating for the two control systems, we observed stronger relations between self-reported workload and eye parameters. Specifically, in the button-based condition, pupil size was strongly correlated both with the overall workload (R = 0.6, p < .001) and the self-reported mental demand (R = 0.66, p < .0001). Furthermore, moderate positive relations were observed between pupil size and effort (R = 0.58, p < .001) and frustration (R = 0.46, p < .01). Differently, perclos demonstrated only weak negative relations with performance (R = -0.39, p = .027) and temporal demand (R = -0.39, p = .029), while blink frequency was moderately correlated with performance (R = -0.44, p = .012). Blink duration did not show any significant correlation with any of the self-reported workload dimensions.

When looking at the same relations in the action-based condition, we notice significant strong relation between pupil size and mental demand (R = 0.66, p < .0001), and moderate relations between pupil size and frustration (R = 0.55, p = .0016), effort (R = 0.41, p = .018) and even the overall NASA-TLX score (R = 0.43, p = .014). Perclos was negatively related to the overall workload (R = -0.38, p = .031) and with the self-reported effort (R = -0.39, p = .026). Finally, blink frequency and duration demonstrated moderate negative relations with the mental demand (blink frequency: R = -0.48, p = .005; blink duration: R = -0.40, p = .023).

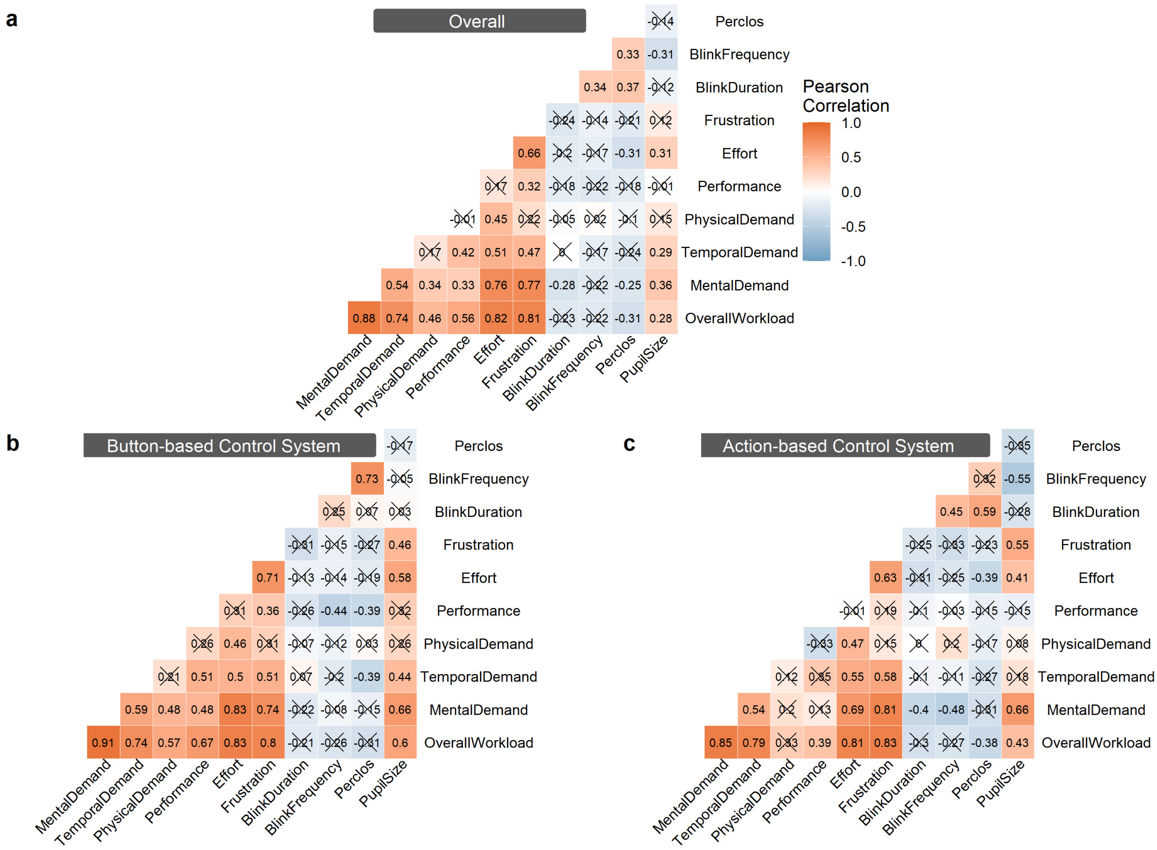

FIGURE 7.10: Correlation matrices between NASA-TLX scores (Overall workload, Mental demand, Temporal demand, Physical demand, Performance, Effort, Frustration) and eye-tracking parameters (Blink duration, Blink frequency, Perclos, Pupil size). Figure **a** depicts correlations on the overall dataset, independently from the control system used to teleoperate the robot; Figure **b** shows correlations within the button-based condition; Figure **c** shows correlations for the action-based control system. The numbers in each cell indicate the Pearson's R coefficients. Not significant correlations are marked with an X (p>.05). The color of the cells reflects the values of the Pearson's R: stronger correlations are characterized by more saturated colors, blue if negative, red if positive.

7.4 Discussion

We here proposed a systematic user-centered investigation of performance and workload during simulated robotic teleoperation in VR. The robot UR10e was faithfully virtualized and our participants guided it through a pick-and-place task via button-based and action-based control systems. The same task was also performed under low (single-task) and high (dual-task) mental demand. We leveraged all interaction data deriving from the tested human-robot collaborative framework, as well as eye-tracking data directly gained from the VR device. In this way, it was possible to monitor users' mental workload without interfering with the task. Furthermore, we collected self-reported workload and preferences for either control systems before and after the experiment. All in all, we draw a broad overview of the human psychophysical state by coupling workload and performance measurements throughout the tasks. Our main questions are deeply discussed in the following subsections.

7.4.1 Task load manipulation

Our results are in line with our hypotheses and support the idea of a cognitive interference between the arithmetic task and the pick-and-place task. Specifically, performance and self-reported workload demonstrated the higher demand of the dual-tasking. Interestingly, despite the action-based control system entailed a higher degree of physical motion compared to the button-based control, participants did not perceive a higher physical demand in that condition.

Furthermore, eye-parameters partially reflected task load-related differences. First, pupil size increased throughout the arithmetic task when moving from the start to the following arithmetic sums, no matter the control system involved. This indicated that our arithmetic task always imposed an increasing workload on the user. Furthermore, pupil size variation during the pick-and-place was significantly higher in the dual-task compared to the single-task, independently of the control system deployed. This finding is thus fairly robust and is in line with literature on teleoperation and/or robotics reporting higher pupil size for higher task load (Nenna et al., 2022a; Wu et al., 2020; Zheng et al., 2015).

Differently, perclos and blink parameters did not capture task load differences with the same accuracy as pupil size variations. Specifically, perclos and blink frequency were affected by the task load manipulation only when participants guided the robot through the action-based control system (Figure 7.7. One of the reasons behind this result, as deeply discussed in subsection 7.4.2, might be related to the different difficulties of use of the two control systems (i.e., button- and action-based). Specifically for the action-based condition, on a macro-level, a lower perclos and blink frequency were observed in the dual-task, likely suggesting higher levels of vigilance compared to the single-task (Marquart et al., 2015). Similar results on workload-related blink variations were also observed in previous teleoperation research (Zheng et al., 2012; Guo et al., 2021). A common assumption is that users are likely to inhibit eye closures to reduce the risk of missing salient information (Fogarty and Stern, 1989), and such interpretation seems to apply to the present task too.

On a micro-level, instead, perclos values gradually increased throughout the single-task from 2.74% on average in the first trials, to 5.26% in the last trials, while they only increased from 0.63% to 1.32% in the dual-task. Similarly, blink frequency also decreased from 4.65 blink/min on average in the first trials, to 2.11 blink/min in the last trials of the task. Even though it was not supported by a statistical significance, this trend might reflect changes in the level of fatigue (Marquart et al., 2015): as time passed, users got tired, and their eye closures decreased. Another possible explanation is that performing the same monotonous task for some minutes can be tiring, thus affecting the level of vigilance in the task course (Körber et al., 2015).

7.4.2 The impact of diverse control systems on the user

Our findings confirmed the hypothesis of better performance and lower workload when using the action-based compared to the button-based control system in VR. Participants also demonstrated higher levels of vigilance throughout the whole pick-and-place executed via the button-based compared to the action-based control system. Performance, self-reports and eye-tracking measures differences between the two conditions were prominent, and they are deeply discussed in the following paragraphs. Overall, this clear advantage of action-based controls might be related to embodied mechanisms involved in physical and direct operations, that can lead to more intuitive control of the virtual robot's movements in the 3D space. The hand-eye coordination is indeed a primal embodied behavior that makes every operation more affordable and natural. When guiding the robot via buttons, instead, the spatial intentions of the user need to be transposed from a 3-dimensional view to 4 static directions that run over two axes (forward-backward, left-right), which increases the operation complexity.

Performance measures. Participants were significantly faster when executing the teleoperations via physical action in both pick and place phases: they saved about 2 sec on average in each pick, and 1.5 sec in each place action compared to when using controller buttons. Furthermore, the error rate at the pick action decreased from about 17% to 7% when switching from controller buttons to physical actions, while the same advantage was not observed in the place action. Again, this is possibly due to the friendliness of the place operation, in which the diameter of the box where to place the bolt was much larger than the bolt itself. Furthermore, by looking at participants' performance at the arithmetic task, as compared to solely summing the presented numbers, the averaged error rate was almost doubled when additionally performing the pick-and-place task via physical actions, and even tripled when driving the robot via controller buttons. However, only the difference between error rates at the single arithmetic task and button-based dual-task was statistically significant. It thus seems that driving the robot via physical actions did not impose a degree of cognitive effort as high as when driving the same robot via controller buttons. In general, there is strong evidence in favor of using action-based controls for guiding the robotic arm. These findings are consistent with previous research that found better performance when using highly interactive control systems during robotic teleoperations (Franzluebbers and Johnson, 2019; Gliesche et al., 2020; Martín-Barrio et al., 2020; Vozar, 2013).

Eye-tracking measures. As indicated in the methods (subsection 7.2.4), we did not intend to directly compare the control system-related eye parameters as they might be strongly influenced by the different movement magnitude involved in the action- and button-based conditions. This precaution was corroborated, for example, by the findings on pupil size variation: while in the button-based condition pupil size gradually increased from the start (window 1) to the end (window 6) of each action, in the action-based condition we observed a more rapid pupil size increase that reached its peak in windows 3 or 4, and then decreased (Figure 7.6). This could be related either to the larger physical motion involved in the action- compared to the button-based condition, which may have elicited higher arousal and activation, or to a constantly higher level of vigilance throughout the whole task session in the button-based condition, which may have flattened the pupil size variation. The latter interpretation seems to be further supported by the higher self-reported workload (Figure 7.8) and by the constantly lower level of perclos observed in the button-based compared to action-based condition (Figure 7.7), which is known to be related to higher level of vigilance (Marquart et al., 2015).

On the same line of interpretation, the perclos difference between single- and dual-task was more evident in the action- than in the button-based condition, likely reflecting that executing the pick-and-place single-task via physical actions was so easy that it required very low vigilance compared to executing the same task via controller buttons. Furthermore, we noticed how perclos and blink duration were affected by our task load manipulation only in the action-based, but not in the button-based condition. Additionally, when using the button-based control system, the task load manipulation similarly affected both pick and place actions, while when using the action-based control system, it only affected the pick action. Taken together, it seems that, no matter the task difficulty, participants were always investing more mental resources when operating via controller buttons rather than physical actions. This might have prevented the emergence of different blink and perclos trends between single- and dual-task, as well as a different pupil size trend between the pick and the place actions specifically in the button-based condition.

Self-reports measures. A generally higher perceived workload was observed in the button-based compared to the action-based condition. Furthermore, participants perceived higher frustration when driving the robot via buttons compared to physical actions. A tendency for higher mental demand was also observed in the button-based compared to the action-based condition, which however did not reach the significance threshold. These results are in accordance with the preference for either control system as rated after the whole experiment. Even before testing the teleoperation modalities, there was a clear tendency to prefer action-based control systems. This preference further increased after the experiment, whereby 89.47% of the tested sample reported personally preferring guiding the robot via physical actions, which was perceived as the less frustrating control system.

7.4.3 VR-embedded eye-tracker sensitivity to workload

As discussed in the subsection 7.4.1, our task load manipulation was effective in producing two distinct levels of workload. From this observation, we further assessed the sensitivity of each eye-tracking metric to workload variations by correlating the self-reported workload and each eye-tracking parameter. From a first glimpse at Figure 7.10, independently from the control system involved, pupil size and perclos are particularly sensitive to changes in mental demand and effort, while blink duration responds specifically to mental demand. However, the latter relation was not supported by the results shown in Figure 7.7, as blink duration did not differ significantly between single- and dual-tasks. Differently, blink frequency did not show significant correlations with any workload dimensions on the overall task, suggesting that it might not be the best indicator of workload in VR. Furthermore, all observed relations in the overall task did not exceed the R=0.36, thus being quite weak.

When addressing the two control systems independently, the positive but weak relations between workload and pupil size observed in the overall task became even stronger. Relations between pupil size and mental demand reached a correlation of R=0.66 in each control system condition, almost doubling the R coefficient observed in the overall task. A positive relation between frustration and pupil size, which was not observed in the overall task, additionally stood in both control system conditions. This could be explained by the strong positive relations between self-reported mental demand and frustration (R = 0.77): the higher the mental demand, the higher the frustration, and the larger the pupil variation. Furthermore, it is worth briefly commenting on Figure 7.5, which shows trends in pupil size variations during the arithmetic task. Specifically, when the dual-task was performed via physical actions, there was higher variability in the pupil size trend compared to the button-based and to the arithmetic task conditions, in which the pupil increase was more linear throughout the task. Yet, a pupil size increase was captured across all conditions. Again, this is indicative of the resilience of such a metric in measuring workload under either higher (action-based) and lower (button-based) degrees of physical motion in VR. These findings on workload and pupil size are in line with robotics (Wu et al., 2020; Zheng et al., 2015) and with our Study 1 addressed in previous Chapter 6 (Nenna et al., 2022a).

While relations between pupil size and workload persist to the varied control system conditions, perclos and blink frequency better respond to workload fluctuations when using the action-based rather than the button-based control system. When considering button-based actions, only some weak relations were observed between perclos and temporal demand and performance. Differently, perclos showed stronger relations with overall workload and effort during action-based operations. Generally speaking, this is in line with the literature (Marquart et al., 2015) demonstrating how perclos is responsive to levels of vigilance and fatigue (which might be reflected in the 'effort' dimension in the NASA-TLX), but it is also in contrast with previous research in robotics that did not demonstrate significant relations between perclos and workload (Wu et al., 2020). Furthermore, if blink frequency did not yield any significant relation with workload in the overall task, it showed a moderate relation with mental demand exclusively in the action-based condition. This finding aligns with literature showing inverted relations between blink frequency and mental demand (Borghini et al., 2014; Zheng et al., 2012). Overall, the sensitivity of pupil size to workload stands out compared to the other eye metrics, which is consistent with previous research (Novak et al., 2015).

7.5 Conclusions

7.5.1 Limitations

As a first limitation, we recognize that real-world teleoperation tasks require more complex and varied activities. However, as in all conducted studies for the present thesis, the choice of a simple experimental task such as the pick-and-place was intentional to guarantee appropriate experimental control yet allowing a natural behavior with the least possible constraints. Second, due to the health emergency spread during the data collection, our results were gathered from a sample of young users (mainly students) with no prior experience with robot teleoperations. Whether such result would also apply to an older population, which is also more representative of an actual working population, is still unknown. Finally, we underline that the virtual robot used in the experimental study was not directly linked to the UR10e, and as such, was just a simulated teleoperation. This simulation was sufficient for fulfilling our aim of focusing on human performance and workload while driving an industrial robot in VR. However, we acknowledge that an actual link between the physical and the virtual model of the robot would constitute a more practical instance of telerobotics, also allowing to test technical aspects such as data streaming and effects of data transmission latency or disruptions.

7.5.2 Future directions

As also underlined in the end of the previous Chapter (Section 6.5), future research might include participants that have more familiarity with teleoperation tasks for better understanding whether these findings transpose to teleoperation experts as well. Furthermore, it is to mention that the medium age of the European labor force in 2019 was about 40 (Statista, 2019), while the averaged age of our sample was about 26. On this point, it would be interesting to assess whether the same performance and workload trends observed in our young sample also apply to an older population, which is more likely representative of the eventual final users of such a technology. This would also help in understanding to what extent older users, who might be particularly unfamiliar with unconventional technologies such as VR, could be willing to accept such devices in their work life.

At a glance, this study paves the way for new perspectives in the telerobotics sector, which see eye-tracking-equipped VR as a valued resource in the ongoing 4.0 and 5.0 industrial revolutions. Such devices allow natural and embodied control of robotic systems, embracing the advantages of collaborative robotics in virtual spaces. We believe that such research line, which complement the innovations of telerobotics with knowledge coming from human factors and cognitive ergonomics sectors, will streamline and improve interactions between humans and robots, thus bringing substantial contribution to industry and society.

Chapter 8

Study 3 - Gaming experience, gender and other individual factors in VR-based telerobotics

A valid interface for telerobotics should be effective for the majority of the population. However, gender (Chan et al., 2019; Paperno et al., 2019; Showkat and Grimm, 2018), gaming experience (Brizzi et al., 2017; Chuan et al., 2007; Gomer and Pagano, 2011), or other individual factors are often likely to affect users' performance when guiding or interacting with a robot. In the present study, we leveraged the same experimental set-up of Study 2 (deployed in the previous Chapter 7) and analyzed data from participants executing the same pick-and-place task in VR via the two control systems, namely controller buttons and physical actions. In this investigation, the dual-task was not considered as it exceeded the main focus of the work. The following data were thoroughly analyzed and discussed: operation times at the pick-and-place task, responses at the NASA-TLX questionnaire, and self-reports on the individual gaming habit, skills and attitudes towards technology. Our research questions, methodologies and results are unfolded and discussed as follows.

8.1 Hypotheses and research questions

In this research, we aimed to understand which are the individual factors affecting human performance and perceived workload in VR-based simulated telerobotics and whether their effects cut across different control systems. We thus explored possible effects of gender, gaming experience, learnability skills, problem solving and trust in technology across the two control system modalities.

8.2 Methods

8.2.1 Sample

The experimental sample consisted of 23 young adults, 11 females and 12 males (M_{age}= 27.23, SD_{age}= 5.2), who volunteered to participate in the study without compensation by signing informed consent. None of the participants had current or past neurological or psychiatric problems. They were all right-handed and had normal or corrected-to-normal visual acuity. The local ethics committee approved the experimental protocol and the study was conducted following the principles of the Declaration of Helsinki.

8.2.2 Technical setup

The technical set-up is the same as the one deployed for Study 2. Please refer to Chapter 7, Subsection 7.2.2 for detailed description.

8.2.3 Procedure, experimental tasks and design

Before starting the tasks, participants filled a questionnaire asking their demographics, gaming habits and attitude towards technology. For the latter, they expressed their agreement on some 5-points likert items about Trust in technology (e.g., "I think that technological devices can help solving daily issues"), Learnability skills (e.g., "When I approach a new technological device, I autonomously learn how to use it") and Problem Solving (e.g., "In case I encounter issues with a device, program or application, I always try to solve it on my own"). Particularly about the Trust in technology, we opted for assessing the individual inclination to trust general technological devices rather than the specific attitude relative to the experience with the robot (Hancock et al., 2011). Differently, for the gaming habits, participants were first asked the question "Have you ever played video games?" with the following multiple-choices: "No, never", "Yes, but I currently don't play video games anymore" and "Yes, I've always played video games and I still do". Then, a second question was administered only to those who selected the second or third choice: "For how long you have played video games?". Thus, by following previous research, participants were assigned to a gaming experience group based on the frequency of play (Brizzi et al., 2017; Chuan et al., 2007).

After the questionnaires, all participants underwent a training session to familiarize themselves with the virtual environment. They then executed a pick-and-place task via both button-based and action-based control systems in VR. Please, refer to Subsection 7.2.3 for details about the task. After each task condition, the NASA-TLX questionnaire (Hart, 2006) was administered to measure the perceived workload. This questionnaire comprises six dimensions on a scale from 1 to 20: Mental Demand (MD), Physical Demand (PD), Temporal Demand (TD), Performance (P), Effort (E), Frustration (F).

8.2.4 Measurements

Three groups of participants were identified based on the initial questionnaire on the gaming experience. Particularly, those who reported having never played videogames were identified as Non-gamers (NG, n = 8), who have played only in the past and for a period between the 2 and 10 years as Past gamers (PG, n = 6), and who have at least 2 years of gaming experience and still play videogames were considered as Actual gamers (AG, n = 9). For all groups, participants' performance was monitored through operation time, computed within each HRI from the first movement of the robot to the last before participants enabled the grabbing/releasing of the bolt. Considering that the robot needed to be perfectly in line with the bolt to grasp it, the Pick phase required high precision to be performed. Differently, the Place task required less accuracy as the box where to release the bolt was bigger. For these reasons, as in previous contributions (*Study 1* presented in Chapter 6 and *Study 2* in Chapter 7), the operation time was analyzed independently in the Pick and Place phases. Moreover, the perceived mental workload measured through the NASA-TLX was analyzed independently in each of its dimensions (MD, PD, TD, P, E, F). The scores at Trust in technology, Learnability Skills and Problem Solving obtained through the initial questionnaire were scaled (0-1) for better visualizing their effects.

8.2.5 Statistical analysis

Statistical tests were conducted using RStudio (Team, 2021). Differences in performance and perceived workload between groups (Control system, Gender, Gaming experience) were analyzed through linear mixed effect models (LMMs) from lme4 package (Bates et al., 2014) with Participant as a random effect. Post hoc contrasts were performed on each significant interaction by applying the Bonferroni correction (Bonferroni, 1936). Model comparison was performed for determining the best fitting models according to the Akaike Information Criterion (Akaike, 1974). The statistical analysis was conducted on the models demonstrating a minimum reduction threshold of 2 AIC units. The model including Control system (button-based, action-based) * Gender (M, F) and Control system * Gaming habits (EGs, MGs, NGs) was considered as the full model (m1). This model was compared with m2 (Control system + Gaming habits + Gender), m3 (Control system * Gaming habits), m4 (Control system * Gender), m5 (Control system) and the null model m6 only including the intercept. Moreover, for both dependent variables, possible associations with Learnability skills, Problem Solving and Trust in technology were assessed through the Pearson's linear correlation test.

8.3 Results

8.3.1 Operation times

Both for the Pick and Place phases, the model m1 better fitted the data according to the AIC. Results demonstrated significant main effects of Control system for both the Pick (X^2 = 897.9, p <.0001) and Place phases (X^2 = 2108.3, p <.0001), while Gaming habits reached the significance threshold only in the Pick phase (X^2 = 6.006, p = .049). Significant interactions were observed between Control system and Gaming experience (Pick: X^2 = 35.49, p <.0001;

TABLE 8.1: Descriptive statistics of operation time (sec) by gaming experience, gender and control system

		Pick		Place	
		Button-based	**Action-based**	**Button-based**	**Action-based**
		mean (SD)	*mean (SD)*	*mean (SD)*	*mean (SD)*
Gaming experience	NG (n=8)	4.68 (3.22)	1.55 (0.97)	3.67 (2.10)	1.35 (0.70)
	PG (n=6)	4.33 (2.36)	1.90 (1.61)	3.16 (1.33)	1.51 (0.89)
	AG (n=9)	3.31 (2.21)	1.55 (1.20)	2.84 (1.39)	1.26 (0.98)
Gender	F (n=11)	4.67 (3.01)	1.69 (1.28)	3.43 (1.83)	1.34 (0.75)
	M (n=13)	3.47 (2.25)	1.90 (1.25)	3.01 (1.53)	1.36 (1.00)

Place: X^2 = 46.07, p <.0001) and Control system and Gender (Pick: X^2 = 23.3, p <.0001; Place: X^2 = 31.37, p <.0001). Significant post hoc contrasts are depicted in Figures 8.1 and 8.2. Descriptive statistics are reported in Table 8.1.

FIGURE 8.1: Effects of Gaming experience on the pick-and-place operation time

Additionally, on the correlation tests, moderate negative relationships between operation time and Learnability Skills (R = -0.53, p < .05), Problem Solving (R = -0.46, p < .05), and Trust in technology (R = -0.43, p < .05) were observed in the Pick phase. Differently, in the Place phase, only the relation between operation time and Learnability skills was just upon the significance threshold (R = -0.43, p < .05). The tests' results relative to the same relations within each Control system condition are reported in Figure 8.3.

8.3.2 NASA-TLX questionnaire

The model comparison revealed m3 to best fit data on P, m4 on TD and F, and m5 on MD, PD and E according to the AIC. Results demonstrated significant main effects of Control system for MD (X^2 = 5.34, p <.05), PD (X^2 = 14.69, p <.001), TD (X^2 = 5.68, p <.05) and F (X^2 = 11.92, p <.001). Gaming experience and Gender never reached the significance threshold for any of

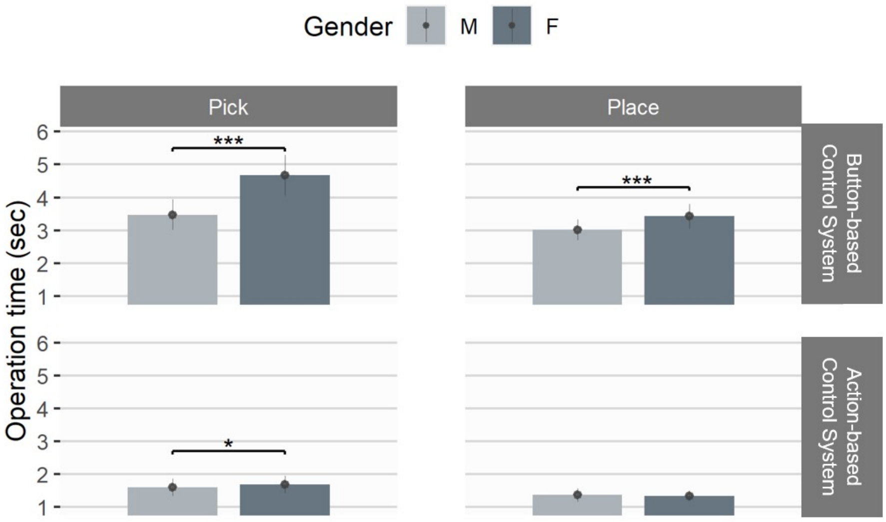

FIGURE 8.2: Effects of Gender on the pick-and-place operation time

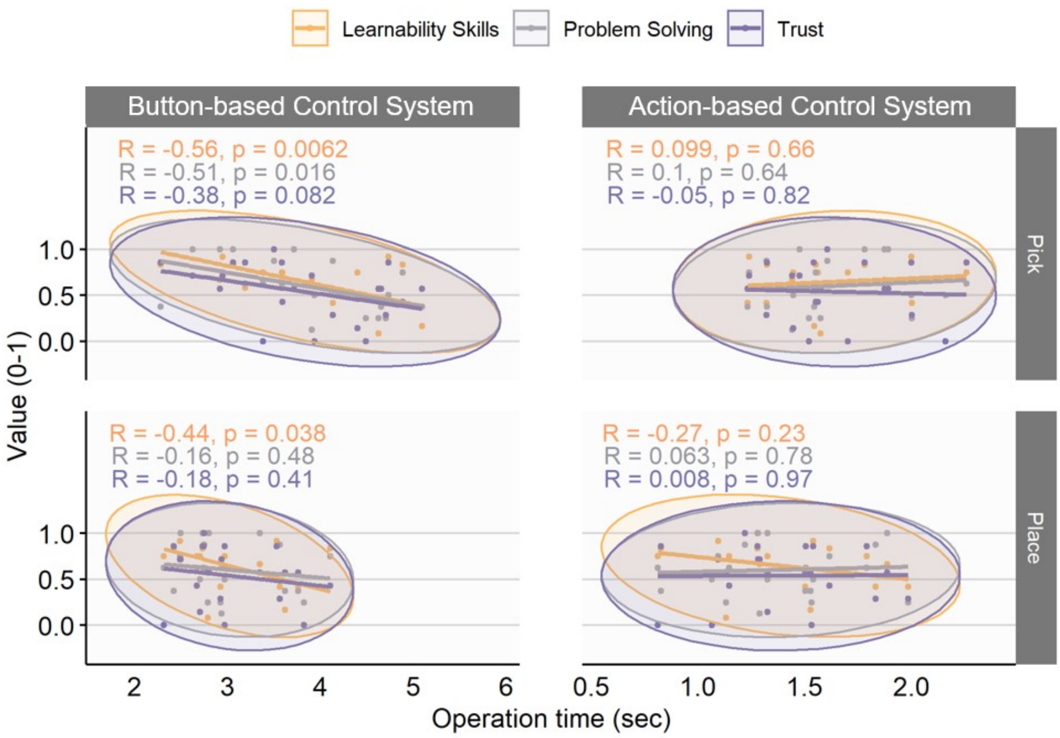

FIGURE 8.3: Correlations between operation time and Learnability Skills, Problem Solving and Trust in technology

the NASA-TLX item. A significant interaction between HRI modalities and Gaming experience was observed for P (X^2 = 8.15, p <.05), and between HRI modalities and Gender for F (X^2 = 5.06, p <.05). However, post-hoc contrasts failed in reaching significance after applying the Bonferroni correction.

On the correlation tests, moderate negative relationships between the NASA-TLX score and Trust in technology were observed for the dimensions MD and E and only in the CB condition. Differently, in the PA condition, no significant relations resulted. Results on correlation are reported in Figure 8.4.

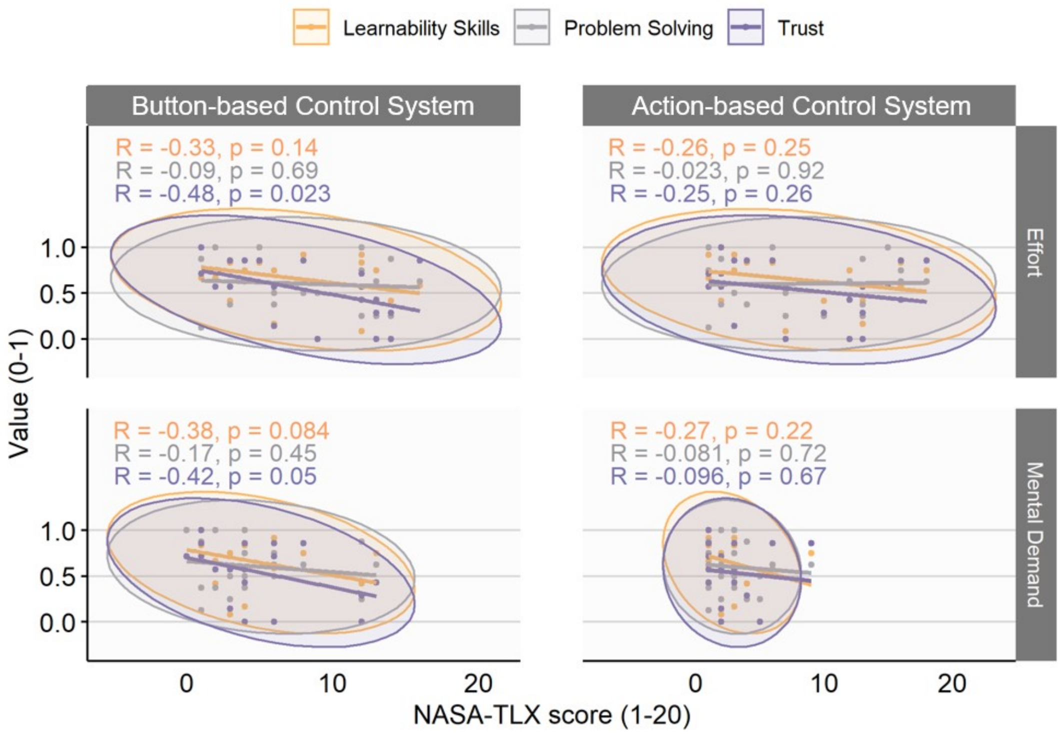

FIGURE 8.4: Correlations between NASA-TLX score and Learnability Skills, Problem Solving and Trust in technology

8.4 Discussion

Results demonstrated that all the individual factors considered in the present investigation (namely, Gaming experience, Gender, Learnability Skills, Problem Solving and Trust in technology) impacted either performance or perceived workload when guiding the robot via VR, but only in the button-based control modality. Specifically, operation time was influenced by Gaming experience, Gender, Learnability skills and Problem solving. Particularly for Gaming experience and Gender, males outperformed females, and gamers executed the pick-and-place task faster than non-gamers. In this regard, the previously observed Gender effect on performance when teleoperating a robot via desktop and mouse (Paperno et al., 2019) or joystick (Showkat and Grimm, 2018) was here replicated in a virtual environment.

However, it is to notice that Gender differences only arose in the button-based condition, while for easier tasks (i.e., the Place phase) allowing participants to guide the robot via actions, Gender differences were leveled out.

On the other hand, our findings on the influence of Gaming experience on teleoperations are only partially in line with the literature (Brizzi et al., 2017; Chuan et al., 2007; Gomer and Pagano, 2011). Notably, a previous study conducted in VR demonstrated how gamers outperformed non-gamers even when interacting with a robot by physical movements (Brizzi et al., 2017). In contrast, and even in this instance, we observed how differences between NGs and AGs were mainly evident when using controller buttons to guide the robot (button-based control system) but were levelled when operating by physical actions (action-based control system), particularly in the Pick phase. Our consideration in this regard is that gamers may be more accustomed to the use of joysticks and button- controlled visual scenes than non-gamers. Therefore, their manual skills likely transposed into an operation time advantage for gamers when teleoperating the robot through button-based controls. Additionally, and exclusively in the Pick phase executed in the action-based condition, PGs were slower than both NGs and AGs, while in all other conditions and tasks a clear incremental advantage was observed from NGs, to PGs and finally AGs. However, this difference was in the order of a few milliseconds (about 350ms).

Operation time was also influenced by the ability to learn how to use a new technological device (namely, Learnability Skills) and the capability of Problem Solving with technology. Indeed, the higher were the Learnability skills, the faster the operations were during the whole pick-and-place task. This finding supports the idea that skills of learnability are fundamental for operators that find occupations in Industry 4.0 (Ra et al., 2019). Interestingly, the influence of Problem Solving on the same operation time was only relative to the Pick task, which required higher precision. In this respect, our results suggest that problem solving skills bring an advantage only for particularly accurate teleoperations.

Furthermore, the perceived workload was observed to be influenced only by the Trust in technology and specifically in the button-based condition: lower effort and mental demand were experienced during the robot teleoperation by those reporting higher Trust in technology. This is significant evidence of how personal beliefs about technology can affect the self-perception of effort and demand when using an advanced industrial technological device. Also, the evidence that effects on perceived workload were explicitly observed for conditions requiring to operate the robot via buttons, but not for those involving physical actions, foster the superiority of physical human-robot interactions over button-controlled ones.

The take-home message of this research resides in the finding that all the analyzed individual factors affected performance or perceived workload, mainly when the virtual robot was guided through controller buttons, but less via physical actions. Thus, the physical action condition resulted in being more affordable for everyone, independently from gender, gaming experience, learnability skills, problem solving, or trust in technology. This result is also supported by evidence of lower self-reported mental demand and frustration – but also higher physical and temporal demand – when operating by physical actions as compared to using the buttons on the controller. One interpretation is that using physical movements to control the robot may have involved a much more natural and spontaneous behavior than using controller buttons. In this sense, the physical HRI modality likely triggered an embodied mechanism that is known to positively influence interactions with hardware or software in gaming experiences (Gregersen and Grodal, 2008), and here applied to one of the scenarios of Industry 4.0.

8.5 Conclusions

8.5.1 Limitations

Some limitations of the present work need to be accounted for. Our sample was composed of individuals that are not experts of telerobotic or VR. Therefore, our results may not generalize to actual industrial operators but still give an insight into the possible implication of introducing VR for teleoperating robots in a context of novices. Moreover, we can not state whether the advantage of experienced gamers is due to their trained abilities with controller/joystick hardware or their increased visuospatial abilities (Dye et al., 2009; Green and Bavelier, 2003). Finally, performance in teleoperation scenarios may also depend on the quality of the virtual interface used for teleoperation.

8.5.2 Future directions

Future research might surely consider a systematic assessment of spatial skills as well, because, as above-mentioned among the limitations of this study, the advantage of experienced gamers in guiding the robotic arm in VR might be due both to their train abilities with gaming hardware (e.g., controllers, joysticks) and to their potentially increased visuospatial abilities.

Overall, important design recommendations for virtual teleoperation of robotic systems can be gleaned from the present work. First, on the need to wisely select the proper tools and interaction modalities for teleoperators. Second, on the importance of conducting human-centered testing for tailoring machines to the needs of humans. Third, and more generally, we demonstrated how performance and self-reported individual differences in robotic teleoperation exist but can be leveled out by implementing proper human-robot interactions.

Chapter 9

Study 4 - Can senior workers benefit from VR as well? Insights on age, performance and workload in telerobotics

The increasingly aging workforce is bringing particular attention to senior workers in industrial and manufacturing sectors. Literature demonstrated how fatigue, mental workload, and occupational stress could alter the older worker's biomechanical control strategies for the upper extremities, causing muscle pain and even occupational injuries (Lagomarsino et al., 2022). Furthermore, the age-related decreased flexibility in the usage of new interactive technologies (Di Pasquale et al., 2020) could prevent senior users from adapting to VR-based teleoperations. Such aspects are thoroughly addressed in Section 4.1. Furthermore, our previous work demonstrated how action-based compared to button-based operations are largely advantageous in young individuals teleoperating an industrial robot via VR (Study 2 deployed in Chapter 7, and Study 3 in Chapter 8). However, whether such trend cut across different ages is still to be clarified. Moreover, to the best of our knowledge, a multimodal assessment of workload in telerobotics, involving both explicit (i.e., self-reports) and implicit (i.e., eye-tracking) metrics in older individuals is missing.

By treasuring the above-mentioned open questions, we thus built on the previously conducted user-centric assessment of performance and workload in virtual robotic teleoperations (Study 2 in Chapter 7) and reproduced it on a more representative population, that is senior individuals. We tested a pool of senior (>50 years old) participants to be compared with the participants previously recruited for Study 2. They thus performed the same identical tasks as in our previous work, which consisted in a simulated robotic teleoperation in VR. All senior participants drove the VR replica of the industrial robot e-Series URE5e through a pick-and-place task under different levels of mental demand (single-task, dual-task) and via button- and action-based control systems. We collected behavioral (i.e., operation times, error rate), eye-tracking (i.e., pupil size variation, perclos) and self-report data (i.e., NASA-TLX questionnaire, individual attitudes towards technology) and discussed the impact of age on VR-based teleoperations.

9.1 Hypotheses and research questions

9.1.1 Age-related effects on behavioral performance

First, we explored age-related effects on teleoperation performance. Specifically, we asked whether young and senior users can be comparably fast (through the analysis of operation time) and accurate (through the analysis of error rate) when driving an industrial robot via VR, or whether the age-related mental and physical decay affect their performance in VR-based telerobotics tasks. The latter case was confirmed for example by Adami et al. (2021) and Grabowski et al. (2021). Differently, similar performances between young and senior individuals were found in repetitive manual and assembly tasks, even though they were not performed in VR (Gilles et al., 2017; Qin et al., 2014).

9.1.2 Age-related effects on mental workload

Similarly, we were interested in understanding whether senior users show a different workload sensitivity, likely showing earlier and/or higher workload and fatigue signs compared to young users during the task execution. This was the case of the above-mentioned studies (Gilles et al., 2017; Qin et al., 2014), in which senior users had to strive harder to obtain similar performance as young participants, at the cost of higher mental workload. Such research question was here examined through the analysis of both implicit (e.g., eye-tracking measures, at both trial and task levels) and explicit measures of workload (i.e., the NASA-TLX scale).

9.1.3 Age, button and action-based VR controls

Furthermore, by following the findings of Study 2 (Chapter 7), we asked if senior individuals show the same tendency as young users for benefitting from action-based rather than button-based control systems in VR-based robotic teleoperations (Nenna et al., 2022a). For answering such question, we leveraged evidence coming from performance (e.g., operation time and task accuracy) and workload measures (e.g., pupil size, perclos and responses at the NASA-TLX questionnaire), as well as self-reports on the preference for either button- or action-based control systems as expressed before and after the experiments (similarly to Study 2, that can be found in Chapter 7).

9.1.4 Age-depended differences in individual factors

As the age-related decreased flexibility in the usage of new interactive technologies (Di Pasquale et al., 2020) could prevent senior users from adapting to VR-based teleoperations, we here further explored possible differences in individual factors according to age, namely frequency of use of VR, knowledge about VR, sense of presence, trust in technology, learnability and problem solving. Notably, the latter three factors can be referred to as *attitudes toward technology*. They were also explored in young participants in our previous contribution (Nenna and Gamberini, 2022), that is addressed in Chapter 8. Differently, possible differences in the frequency of use of VR and knowledge about VR were analyzed as a control, to better specify the characteristics of our samples. Finally, the sense of presence was also assessed as it is an influential factors in VR-based telerobotics platforms. Indeed, literature presents different examples of increased teleoperation performance for increased sense of presence (Ma and Kaber, 2006; Toet et al., 2020).

9.1.5 Age-depended relations between performance, workload and individual factors

Furthermore, considering that individual factors play a crucial role in the equation that constitutes workload (Longo et al., 2022), and that performance and workload are strictly related (Bruggen, 2015), we additionally explored relations between i) performance and workload, ii) individual factors and performance, and between ii) individual factors and workload. This analysis allows to better understanding the relation between performance and workload in our young and senior participants, and whether certain individual factors are more influential then others in promoting a better performance and/or a lower mental workload during teleoperation tasks.

9.2 Methods

9.2.1 Sample

An a priori power analysis conducted on Gpower (Erdfelder et al., 1996) indicated that a total sample of 26 participants was needed for our within-between subjects design to detect a medium effect size (d = 0.5) with 90% power. The experimental sample thus consisted of 30 participants: specifically, we recruited 15 participants who reported being more than 50 years old, which we refer to as the Senior group (M_{age} = 57; SD_{age} = 5.07), while the other half is considered as the Young group (M_{age} = 25.85; SD_{age} = 1.95). For the Young group, we randomly selected 15 participants from the sample deployed in our previous work (Study 2, Chapter 7). Specifically, the following participants were randomly selected: *p04, p05, p07, p08, p09, p10, p11, p13, p14, p17, p18, p20, p22, p23, p25*. No modifications were applied to the experimental setup nor to the tools used for the data collection. Thus, the Young group comprised 8 females and 7 males, while the Senior group comprised 9 females and 6 males. All participants volunteered to participate in the study and signed informed consent forms. None of the participants had current or past neurological or psychiatric problems. They all reported having normal or corrected-to-normal visual acuity only through contact lenses and normal color vision. The experimental protocol was approved by the local ethics committee and the study was conducted following the principles of the Declaration of Helsinki. One participant in the Senior group had very low technological skills and spent more than 30 minutes only completing the training phase; therefore, she was excluded from further analysis.

9.2.2 Technical setup

The technical set-up was identical to the one employed in our previous contribution (Study 2, Chapter 7). Please, refer to Section 7.2.2 for more details. A picture of the VR environment can be found in Figure 9.1

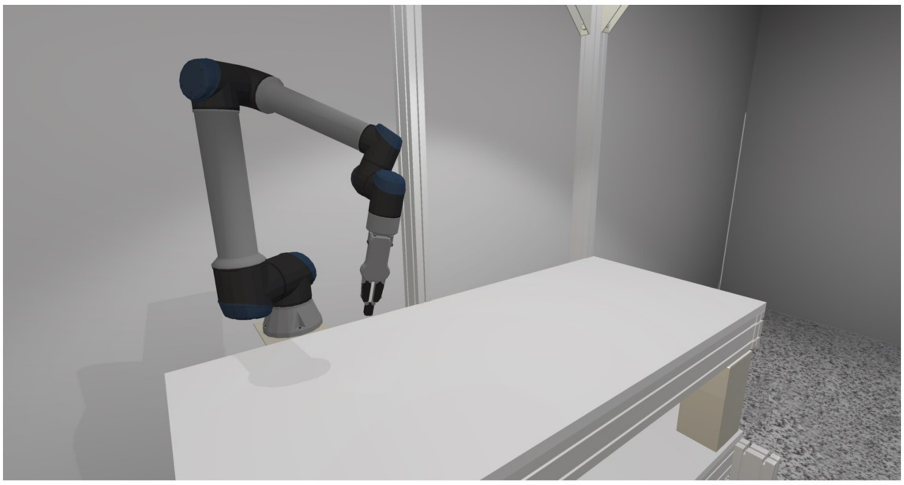

FIGURE 9.1: The virtual environment with the robotic arm UR10e faithfully reproduced via Unity

9.2.3 Procedure, experimental tasks and design

The task and procedure were identical to those employed in our previous contribution (Study 2, Chapter 7). Therefore, we here describe them briefly; for more detailed information, please refer to Section 7.2.3. After signing informed consent, all participants filled out questionnaires about their demographics, VR and robotics expertise and individual preference for guiding a robot via either a controller-based or an action-based control system. Thereafter, a training session was presented, after which the eye-tracking was calibrated and then the experiment started. Participants thus randomly performed 5 tasks: an arithmetic task as a baseline, and a pick-and-place task executed via controller buttons (button-based condition) and physical actions (action-based condition), both under single- and dual-task (concurrently with the arithmetic task). Breaks were suggested after each task. Afterward, the questionnaire on the individual preferences for either control systems was administered again and the experiment ended. Task conditions are depicted in Figure 9.2. Differently from previous research, as the senior group needed longer familiarization time for learning the commands to use the VR headset, these participants executed 20 trials of each task. In this way, the average experimental time (training and questionnaires included) did not exceed 1 hour and a half. To allow proper comparisons between the two aged groups, only the first 20 trials were included in the analysis of performance and workload of younger participants as well.

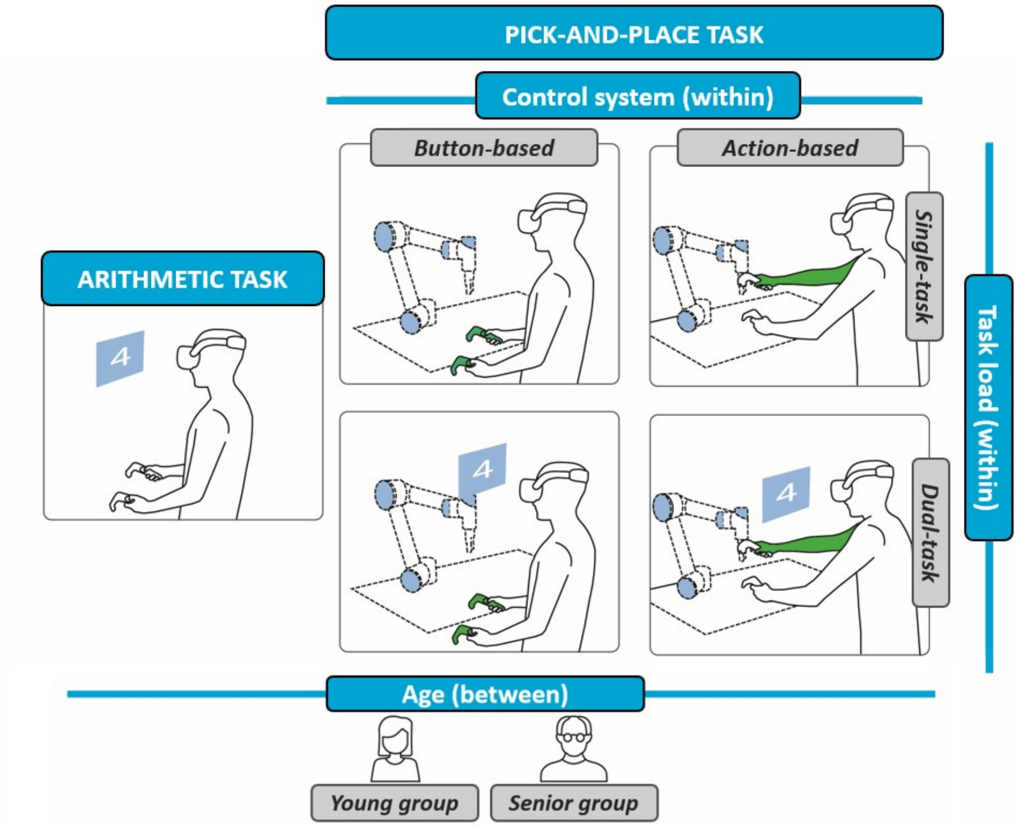

FIGURE 9.2: Experimental tasks and design

9.2.4 Measurements

Performance measures

Pick-and-place task. Following the same methodology as our previous contribution (Study 2, Chapter 7), we measured the pick-and-place operation time and error rate, independently in the pick and place phases. Trials with durations greater than 4SD were removed as they likely reflected unrealistic operations. The same trials were not considered for the analysis of the other independent measures either. Pick and place trials were registered as 'incorrect' in those cases in which the participant pressed the left pad for picking or placing the bolt, but the robot was not aligned with either the bolt or the box.

Arithmetic task. Furthermore, together with the performance at the pick-and-place task, the analysis of the error rate at the arithmetic task allowed to quantify the dual-task interference between the primary and secondary tasks.

Eye-tracking measures

Pupil size variation. As in Study 2, the participant's implicit workload was inferred via continuous pupil size variations, which are known to accurately reflect changes in workload (Nenna et al., 2022a). All eye data on the pupil size variation were preprocessed following the same methods adopted in Study 1 and and Study 2. For more details about the pupil size preprocessing please refer to Section 7.2.4. Specifically, we inferred workload in each pick and place phases by averaging the pupil size values over the trials of each aged group and in each task condition (trial level). Furthermore, and differently from the previous Study 2, we also assessed pupil size variations over longer periods, that is at task level. With this intent, we looked at pupil size trends in 4 windows including 5 trials each. This analysis allowed us to investigate whether senior users demonstrate earlier and/or higher workloads compared to young users during the prolonged execution of repetitive tasks.

Perclos. Furthermore, similarly to Study 2, the users' level of vigilance was inferred by their level of perclos, that is the percentage of time during which users' eyelids covered more than 80% of the eye (Marquart et al., 2015). For more details about the perclos data computation and preprocessing, please refer to Section 7.2.4. For the analysis, we looked at perclos trends in 4 windows that included 5 trials each (task level) to explore if senior users demonstrate different trends in vigilance compared to young users during the prolonged execution of repetitive tasks.

Self-reports measures

NASA-TLX questionnaire. Besides using eye-tracking measures for inferring implicit workload, the explicit workload was inferred through self-reports (namely, the NASA-TLX questionnaire (Hart, 2006), which was administered after each experimental condition).

Individual factors. Furthermore, we assessed a pool of individual prior experience, knowledge, attitudes and skills that might influence users' performance when working with virtual devices and/or robots. Particularly, we asked about participants' **knowledge about VR technology**, their **frequency of use of VR technologies**, and their **attitude towards technology**. The knowledge about VR was assessed by expressing agreement with the following statement on a 5-point scale: "I know what VR is". The frequency of use of VR devices was assessed by asking participants to rate the frequency with which they had used VR devices on a 5-point scale. Differently, for the attitude towards technology, we leveraged the same questionnaire administered in Study 3 (Nenna and Gamberini, 2022), which is thoroughly discussed in Chapter 8. Specifically, we asked participants to express their agreement on the following constructs: trust in technology (e.g., "I think that technological devices can help solve daily issues"), learnability skills (e.g., "When I approach a new technological device, I autonomously learn how to use it") and problem-solving (e.g., "In case I encounter issues with a device, program or application, I always try to solve it on my own").

Additionally, we also asked whether each participant had already **used a robot at least once** (possible answer: yes/no). Finally, as in Study 2, participants had to express their **individual preference** for driving the robotic system via buttons- or action-based control systems, both before and after the experiment. This can help understand whether a practical use of such interfaces can modify individual preferences toward either control system.

Sense of presence. The sense of presence was assessed through the short version of the MEC-SPQ scale (Vorderer et al., 2004), which was administered after the whole experiment and assessed the following constructs: attention allocation, spatial situation model, self-location, possible actions, higher cognitive involvement, suspensions of disbelief, visual spatial imaginery, and domain specific interest.

9.2.5 Statistical analysis

Before analyzing users' performance and pupil variations data, we first fitted the data through the function descdist() of the package fitdistrplus (Delignette-Muller and Dutang, 2015) and chose the appropriate model setting based on data distribution. Moreover, for the interpretation of all post hoc contrasts, we always applied the Bonferroni correction for multiple comparisons (Bonferroni, 1936).

Performance measures

We here used Generalized Linear Mixed-Effects Models (GLMM) for analyzing both pick-and-place operation time and error rate, which included the following factors: Age (Young, Senior), Task load (Single-task, Dual-task), Control system (Button-based, Action-based). Participant was set as a random effect.

Eye-tracking measures

After the pre-analysis process, pupil size variations were analyzed via GLMMs over the factors Age (Young, Senior), Task load (Single-task, Dual-task), Control system (Button-based, Action-based) and Window (1, 2, 3, 4, 5, 6). While at trial-level the factor Window comprised 6 levels for measuring workload throughout the trial, at task-level it comprised 4 levels, each including 5 trials. Similarly, perclos was analyzed over the same factors. Participant was always specified as a random variable.

Self-reports measures

NASA-TLX questionnaire. The NASA-TLX score was analyzed through a GLMM including the factors Age (Young, Senior), Task load (Single-task, Dual-task), Control system (Button-based, Action-based) and Item (mental demand, temporal demand, physical demand, performance, effort, frustration).

Individual factors. We ran Wilcoxon tests to compare the scores of young and senior participants in the following self-report scales: the knowledge about VR technology, the frequency of use of such devices, the problem solving, learnability and trust in technology. Finally, we reported the descriptive statistics of the individual preference for driving the robotic system via button- or action-based control systems as expressed before and after the experiment, and the number of participants who reported to already have used a robot at least once.

Sense of presence. Additionally, we investigated age-related differences in the sense of presence as self-reported in the MEC-SPQ scale. With this aim, participants' responses at the MEC-SPQ were analyzed via GLMM over the factors Age (Young, Senior) and Item (attention allocation, spatial situation model, self-location, possible actions, higher cognitive involvement, suspension of disbelief, visual spatial imaginery, domain specific interest).

Correlation measures

Finally, we explored Spearman correlations between individual factors (i.e., frequency of use of VR, knowledge about VR, sense of presence, learnability, problem-solving, trust in technology), performance (i.e., operation time, error rate) and workload (i.e., all items of the NASA-TLX questionnaire, perclos, pupil size).

9.3 Results

9.3.1 Performance measures

Pick-and-place task. When analyzing the operation times at the pick-and-place task, the GLMM demonstrated significant main effects of Task Load (pick: X^2 = 109.82, p < .0001; place: X^2 =150.13, p < .0001), Control System (pick: X^2 = 1162.73, p < .0001; place: X^2 = 1121.90, p < .0001) and Age (pick: X^2 = 21.12, p < .0001; place: X^2 = 26.03, p < .0001). Additionally, all interaction effects in the pick phase reached the significance threshold; therefore, we here report only the result of the full interaction Task load, Control System and Age (X^2 = 30.05, p<.0001). Similarly, the full interaction of Task load, Control System and Age demonstrated to be significant also in the place phase (X^2 = 70.23, p < .0001). For what concerns the effect of Age, senior users showed slower operation times as compared to young ones only in the action-based condition, both under single-task (pick: p<.001; place: p<.0001) and dual-task (pick: p<.0001; place: p<.0001). On the effects of Task load, post-hoc contrasts revealed shorter operation times in the single- compared to the dual-task for senior participants executing the pick action with both control systems (all ps < .0001), and for senior participants executing the place action specifically via physical actions (p<.0001). Young participants performing the pick action, instead, showed shorter operation times in single- compared to the dual-task in the action-based condition (p<.001) and during the place phase in the button-based condition (p<.05). Differently, regarding the Control System, operation times were always significantly higher in the button- compared to the action-based control system condition, both in young and senior users, under single- and dual-task and both in the pick and place phases (all ps < .001).

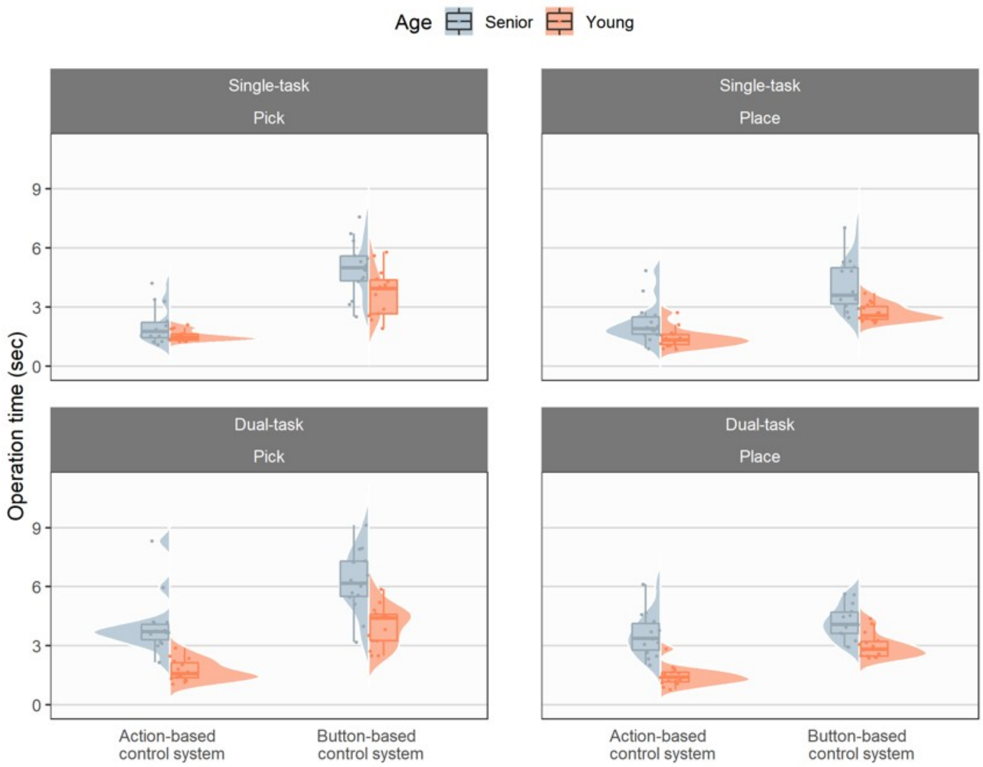

FIGURE 9.3: Operation time at the pick-and-place task for young (orange) and senior participants (grey)

The analysis of error rate at the pick-and-place task (Figure 9.4, instead, yielded significant main effects of Task Load (X^2 = 12.24, p < .0001) and Control System (X^2 = 29.51, p < .0001) only for the pick action. Differently, Age was not significant, not in the pick (X^2 = 0.06, p = .81) not in the place phase (X^2 = 2.52, p = 0.11). Additionally, the interactions between Task Load and Age (X^2 = 4.32, p < .05) and Task load and Control System (X^2 = 6.20, p < .05) were significant only for the pick action. Post hoc on the interaction between Task load and Age during pick action revealed a significant difference between single- and dual-task in the senior participants only (p<.01). Furthermore, in the pick phase, post hoc on the interaction between Task load e

Control System showed significant difference between dual- and single-task only in button-based condition (p<.001). When looking at post hoc contrasts on the interaction between Control System and Age, instead, none of the contrasts was significant in both pick and place actions.

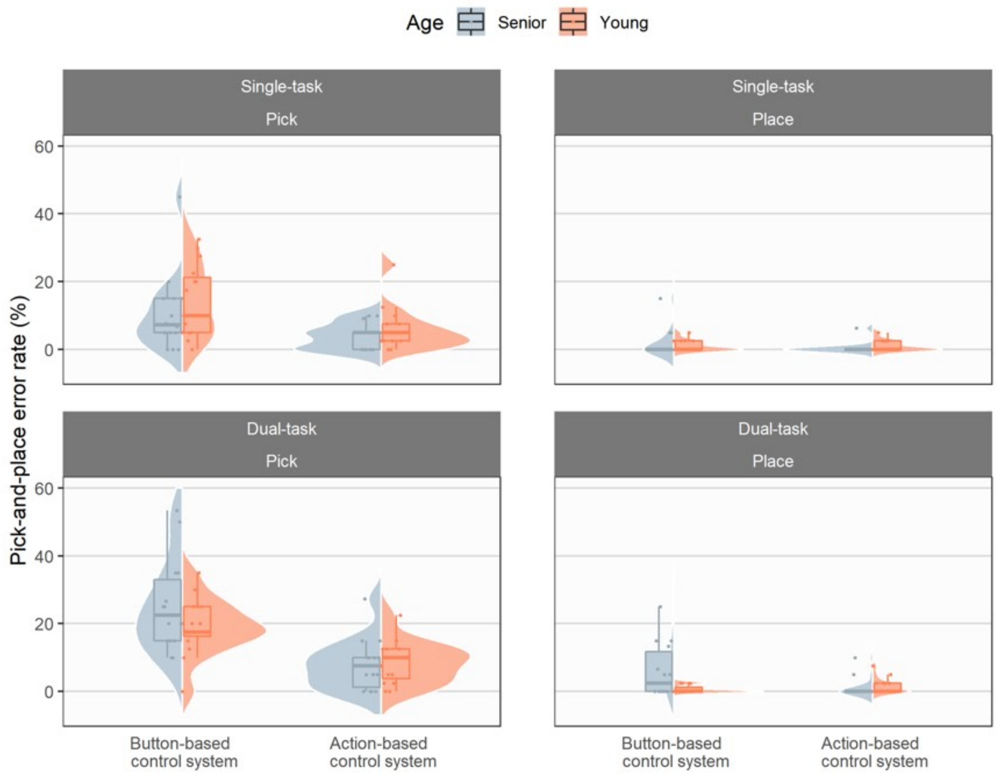

FIGURE 9.4: Error rate at the pick-and-place task in young (orange) and senior participants (grey)

Arithmetic task. Finally, the analysis of error rate at the arithmetic task (Figure 9.5 revealed significant main effects for both Task (X^2 = 36.70, p < .0001) and Age (X^2 = 12.51, p < .0001) and a significant interaction between them (X^2 =49.05, p < .0001). Post hoc on the factor Task showed significant differences between the Single Arithmetic task and both the Dual-tasks (ps < .0001), but not between the dual-tasks performed via action- and button-based controls (p=.52). When looking at the contrasts between Task and Age, a similar trend was only observed for senior participants, who demonstrated higher error rate in both dual-tasks compared to the single-task (ps < .0001), but no significant difference between the two dual-tasks (p = .99). Differently, for the young group, no significant differences in Task and Age contrast were found (all ps > .05). Furthermore, significant differences between young and senior participants were observed when comparing the error rate at the dual-tasks, both when using the action-based (p<.0001) and the button-based control systems (p<.01), but not when comparing the error rate at the single arithmetic task (p = .25).

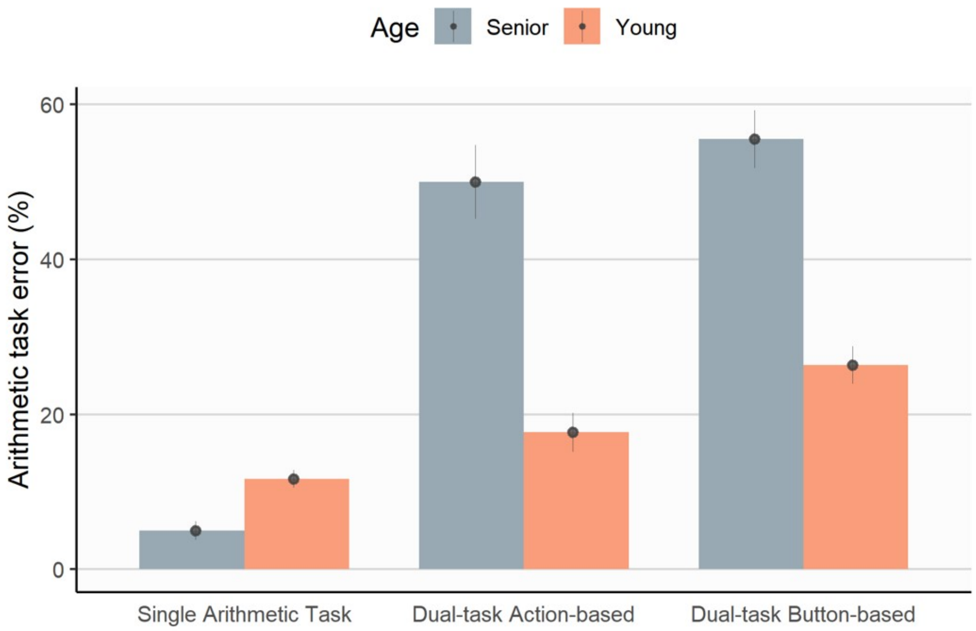

FIGURE 9.5: Averaged error rate at the arithmetic task in young (orange) and senior participants (grey). Error bars depict the standard errors.

9.3.2 Eye-tracking measures

Pupil size variation - trial level. At trial-level, significant main effects were observed for Task load only in the pick action (X^2 = 129.80, p < .0001), and for Control System both in the pick (X^2 = 271.23, p < .0001) and place action (X^2 =682.98, p < .0001), but not for Age (pick: X^2 = .91, p = .34; place: X^2 = 0.87, p = .35). When looking at the contrasts between the factor Window and each of the independent variables under manipulations (i.e., Age, Task load, Control System), we observed significant interactions between Window and Task Load (Pick: X^2 = 154.88, p < .0001; Place: X^2 = 36.73, p < .001), Window and Control System (Pick: X^2 = 608.19, p < .0001; Place: X^2 = 283.48, p < .0001), Window and Age (Pick: X^2 = 178.17, p < .0001; Place: X^2 = 31.49, p < .0001). Besides, we also observed significant interactions between all the factors Window, Task Load, Control System and Age (Pick: X^2 = 67.66, p < .0001; Place: X^2 = 23.91, p < .001). Post-hoc tests of interest covered the comparison between young and senior users within each window, and both for each control system and task load condition (Figure 9.7. Specifically, pupil size variation was observed to be significantly higher in the senior compared to the young group only in the single place action executed via button-based controls (window five: p<.01; window six: p<.05). In all other conditions, the pupil size variation of senior users did not differ significantly from young users (all ps > .05). Furthermore, the Task Load significantly affected pupil size in young participants executing the pick action both with the action-based control (windows three to six: p<.001), and button-based control (windows four to six: p<.0001), and also when executing the place action via action-based control (window five: p<.05; window six: p<.01), and button-based control (windows three to six: p<.0001). Similarly, for senior participants, Task Load significantly affected pupil size when executing the pick action both via action-based control (windows two, three, four, six: p<.0001) and button-based control (window five: p<.05), and also the place action via action-based control (windows three and four: ps<.0001) and button-based control (window three: p<.05; windows four to six: ps<.0001).

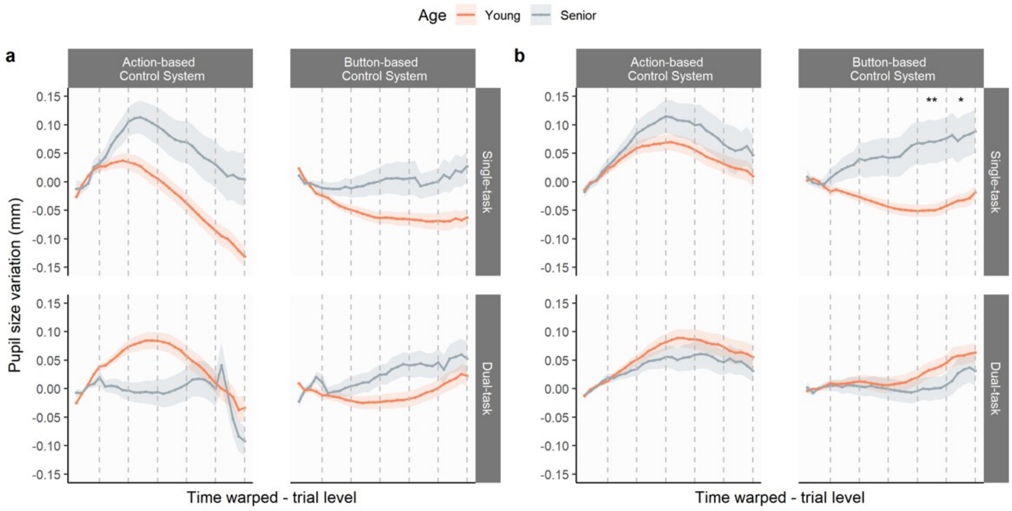

FIGURE 9.6: Pupil size variations at trial level in young (orange) and senior users (grey) according to the task load (single-task, dual-task) and control system (action-based, button-based). Panel **a**) refers to the pick phase, and panel **b**) to the place phase. The shaded area around each line represents the standard deviation.

Pupil size variation - task level. At task level, instead, significant main effects were yielded for Task Load (X^2 = 377.39, p < .0001), Control System (X^2 = 19.21, p < .0001) and Window (X^2 = 121.65, p < .0001). Differently, no main effect was found for Age (p = 0.052). The full interaction between Task Load, Control System, Age and Window also reached the significance threshold (X^2 = 12.97, p < .01). As shown in Figure 9.8, post hoc contrasts revealed how the pupil size variation of senior users was significantly different from young participants when performing the single-task via action-based controls in window four (p<.01), and when performing the dual-task via action-based controls in window three (p < .05) and four (p < .01). About the Task Load effect, post hoc contrasts revealed significantly higher pupil size for senior participants engaged in dual- compared to the single-task in button-based (windows one to four: ps<.0001) and action-based condition (window three: p<.001). Younger participants, instead, obtained higher pupil size in the dual- compared to single-task only in the execution of button-based condition (windows one to four: ps<.0001).

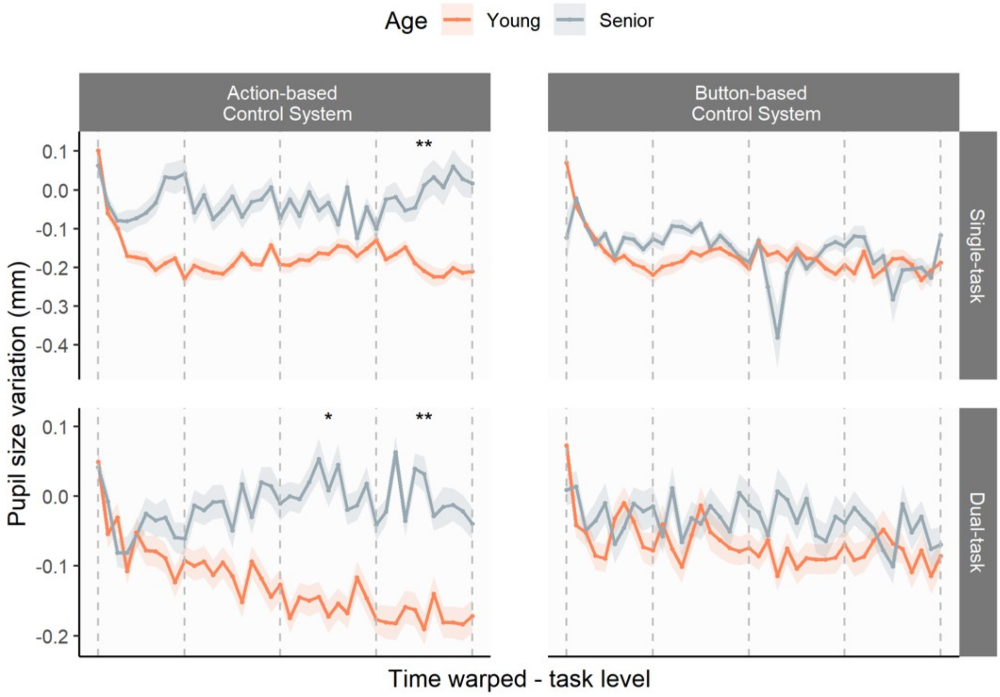

FIGURE 9.7: Pupil size variation at task level in young (orange) and senior (grey) users, as a function of Task load (single-task, dual-task) and Control system (action-based, button-based).

Perclos - task level. The GLMM showed only a significant main effect of Task Load (X^2 = 19.61, p < .0001), while the Control System, Age and Window were not significant. Additionally, interactions between Task Load and Control System (X^2 = 8.17, p < .01), Control System and Age (X^2 = 18.12, p < .0001) and Task Load, Control System and Age (X^2 = 10.84, p < .001) reached the significance threshold. Post hoc tests on the interaction between Task load, Control System and Age demonstrated how PERCLOS of Senior and Young users did not differ significantly in any of the conditions (all ps > .05). Task load affected the PERCLOS only in young users executing the pick-and-place task via action-based control, with significantly lower PERCLOS under dual-task compared to the single-task (p<.0001). Furthermore, Control System affected the PERCLOS in young participants engaged in single-tasks when using action- compared to button-based Control System, with a significantly lower PERCLOS registered in button-based condition (p<.0001). Results on perclos are depicted in Figure 9.6.

9.3.3 Self-report measures

NASA-TLX questionnaire. As depicted in Figure 9.9, results of the LMM demonstrated significant effects of Task Load (X^2 = 188.30, p < .0001), but not Control System (X^2 = 1.07, p = 0.3) and Age (X^2 =0.04, p = 0.8). Significant interaction effects were also observed between Item and Age (X^2 = 37.07, p < .0001), Item and Task load (X^2 = 50.11, p < .0001), and Item and Control System (X^2 = 11.53, p < .05). Post hoc tests on the interaction between Task

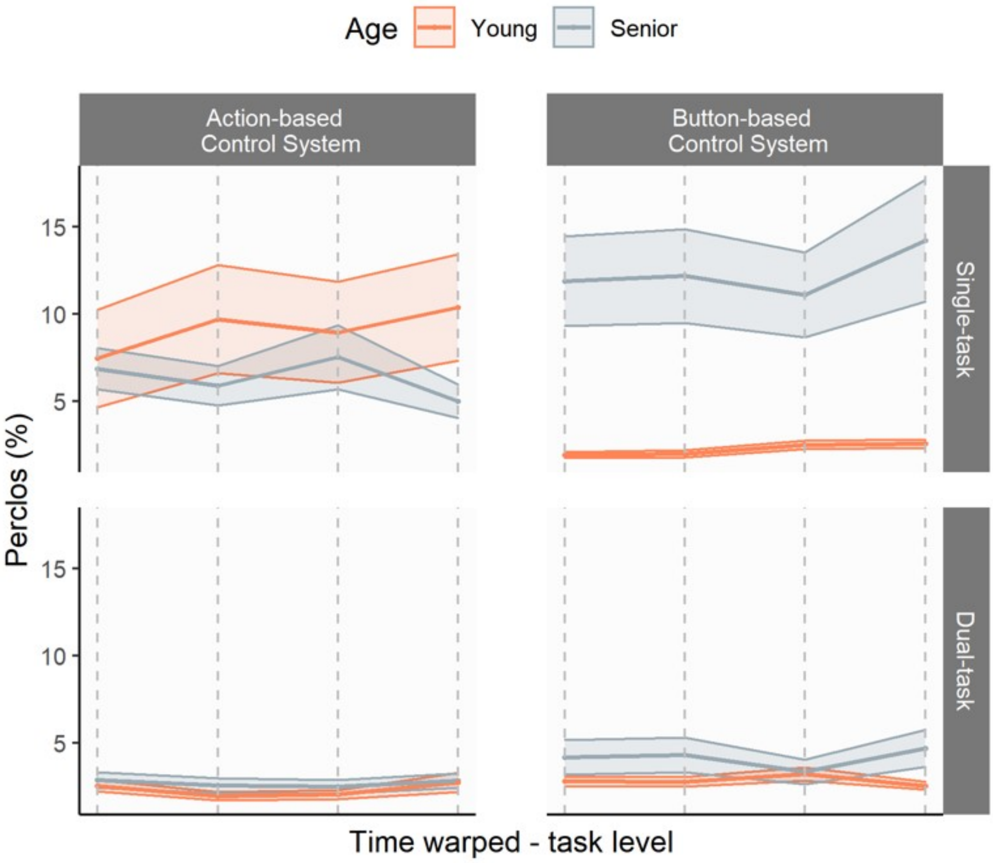

FIGURE 9.8: Perclos levels in young (orange) and senior (grey) users as a function of Task load (single-task, dual-task) and Control system (button-based, action-based). Each point represents the perclos computed over 5 trials.

load and Item yielded significant differences between single- and dual-task in the following items: Mental Demand (p < .0001), Performance (p < .01), Effort (p < .0001) and Frustration (p < .0001). Differently, post-hoc tests on the interaction between Item and Control System and Item and Age did not show any significant difference between the button- and action-based conditions for any of the items. Regarding post-hoc contrasts involving the factor Age, the tests executed on the interaction between Age and Item showed how the score at the item Physical Demand was significantly lower (p < .05) in young as compared to senior users. All other contrasts did not reach the significance threshold.

FIGURE 9.9: NASA-TLX questionnaire in young (orange) and senior users (grey) as reported after single- and dual-task executed both via action- and button-based control systems. Results are reported considering the response range of 1-20.

Individual factors. The averaged individual attitudes towards technology as self-reported by young and senior users are depicted in Figure 9.10. The Wilcoxon test showed how young participants demonstrated significantly higher knowledge about VR compared to senior users (W = 53.5, p=.017), while the frequency of use of VR devices did not differ significantly between young and senior users (W = 90, p = 0.37). Additionally, the statistical test revealed higher learnability skills for young compared to senior participants (W = 35.5, p = .002) and no significant differences in terms of problem-solving and trust in technology between the two groups (problem-solving: W =72.5; p =.10; trust in technology: W = 83.5; p = .23). Only one young and one senior participant reported to have already used a robot at least once. Finally, in-

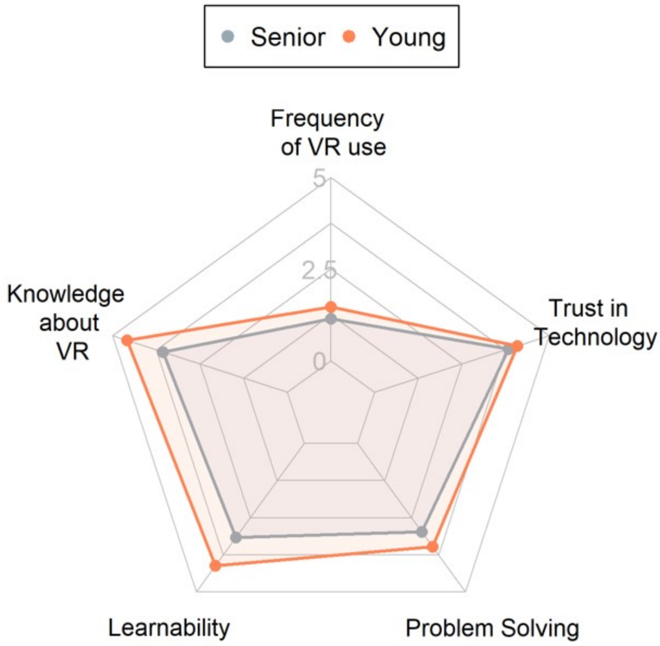

FIGURE 9.10: Averaged individual factors as self-reported by young (orange) and senior (grey) users. All items were presented on a Likert scale that ranged between 1 and 5.

dividual preferences for actions- or button-based control systems expressed before and after the experiment are shown in Figure 9.11. Before the experiment, 73.33% of the younger participants and 78.57% of senior participants expressed a preference for guiding the virtual cobot via action-based controls, while after the experiment their percentage increased to 86.67% for the younger group and to 85.71% for the senior group.

Sense of presence. The model analyzing age differences in sense of presence (Figure 9.12), instead, yielded significant main effects for Item (X^2 = 85.10, p < .0001) but not for Age (X^2 = 2.06, p = .15). Furthermore, their interaction did not reach the significance threshold (X^2 = 13.43, p = .06) and none of the post hoc contrasts revealed to be statistically significant.

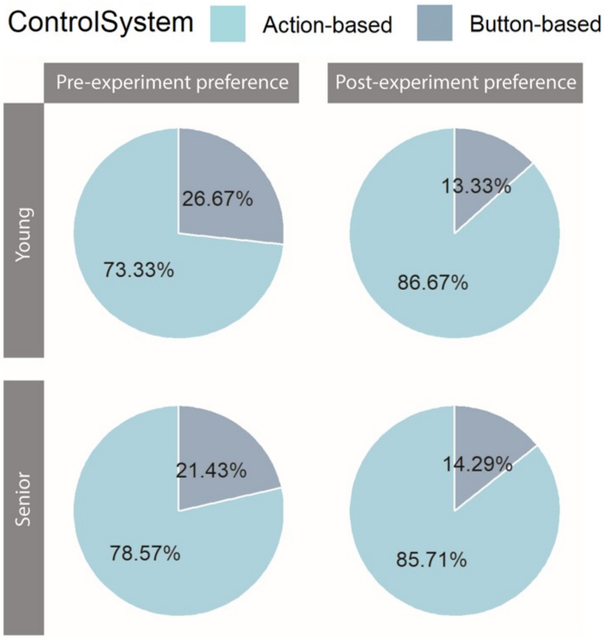

FIGURE 9.11: Self-reported preference for guiding the virtual robot via actions- (light blue) or button-based Control systems (dark blue) as expressed before and after the experiment by both young and senior participants.

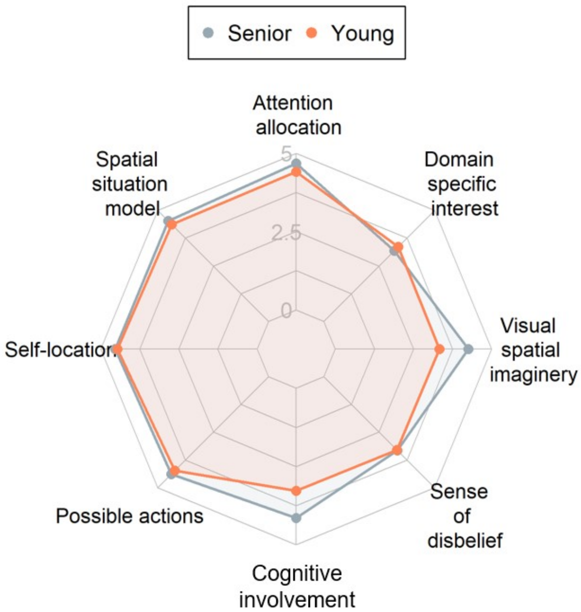

FIGURE 9.12: Averaged sense of presence as self-reported by young (orange) and senior (grey) users in each item of the MEC-SPQ scale

9.3.4 Correlation measures

Results of the correlation matrix between performance at the pick-and-place task (i.e., operation time, error rate), individual factors (i.e., trust in technology, problem solving, learnability, sense of presence, knowledge about VR, frequency of use of VR) and workload (i.e., alla NASA-TLX dimensions, pupil size and perclos) are shown in Figure 9.13.

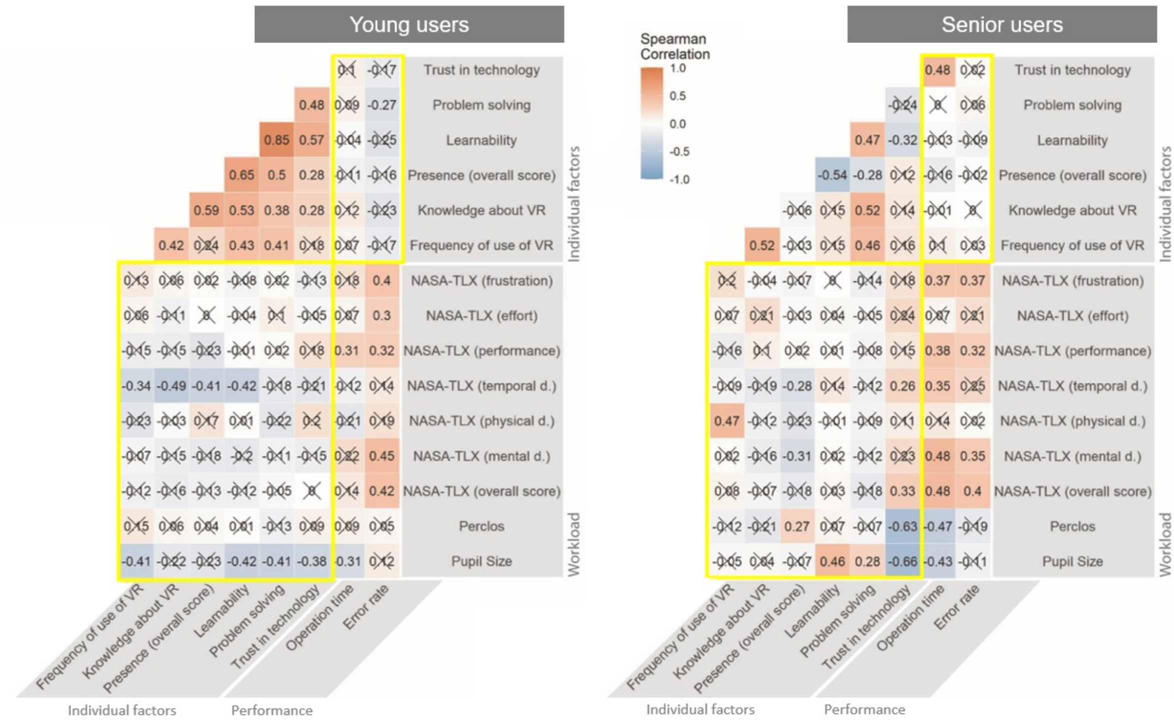

FIGURE 9.13: Correlation matrix between performance, individual factors and workload. The numbers in each cell indicate the Spearman R coefficients. Not significant correlations are marked with an X (p>.05). The color of the cells reflects the values of Spearman's R: stronger correlations are characterized by more saturated colors, blue if negative, red if positive. The yellow squares delimit the correlations between individual factors and performance, and individual factors and workload.

9.4 Discussion

In this research work, we intend to emphasize the needs and capabilities of older workers within the frame of the latest industrial virtual revolution. Specifically, we aimed at better understanding the following aspects: first, whether and how senior users can perform a number of VR-based robotic simulated teleoperations; second, how their workload capacity responds to such a repetitive task over time; third, if they tend to benefit more from action-based rather than button-based teleoperations, as previously observed in young users; forth, if senior and young workers demonstrate different individual attitudes to this new interactive technology; finally, whether some individual factors are more important then others in enhancing performance and lowering workload. With these intentions, we tested 15 young (< 30 years old) and 15 senior (>50 years old) individuals while driving the VR-based reproduction of an industrial robotic arm through a pick-and-place task. They completed the task in single- and dual-task conditions, using both action and button control systems. In the following subparagraphs, we discuss age-related differences in performance, eye-tracking, and self-report data as a function of the imposed task load and the different teleoperation control systems involved.

9.4.1 Age-related effects on behavioral performance

Overall, older individuals demonstrated higher difficulties when performing VR-based tasks than younger ones. We deduced that from the higher dual-task cost at the arithmetic task and the slower pick-and-place performance of seniors compared to younger participants, which aligns with previous VR-based teleoperation research (Grabowski et al., 2021). Interestingly, and in contrast with prior works, the task accuracy was unaltered: older participants could perform as accurately as younger ones, but at higher workload costs (as revealed by pupil size variations). Such a finding is particularly relevant, as it demonstrates how senior participants can drive an industrial robotic arm via VR as accurately as younger ones. Another interesting result is that senior operators slowed their performance and committed more errors under dual-tasking, particularly when executing precise robot maneuvering (i.e., pick phase) and when driving the robot via action controls (i.e., action-based condition). Such findings likely reflect the different levels of precision required in the pick and the place phases, besides the already mentioned difficulties related to the physical and motor complications that come with aging (Ilmarinen, 2001). Indeed, the higher the precision required by the teleoperation, the more senior participants' performance was affected by our task load manipulation.

9.4.2 Age-related effects on mental workload

In some conditions, older participants demonstrated higher workload levels and fatigue than younger users. Similar results were previously observed in physical repetitive mounting and assembly tasks (Gilles et al., 2017; Qin et al., 2014) and virtual robotic teleoperation tasks (Grabowski et al., 2021). Nonetheless, unlike previous research that based workload-related assumptions mainly on self-reports, our statement also builds on implicit workload metrics. A greater pupil size variation was observed in the senior compared to the younger group after the repeated execution of the same operation over time (task level) in the action-based condition, but also at the end of those single operations that took longer (trial level), in the button-based condition. For the latter result, at trial level, it is to notice that senior participants were particularly slow when driving the robot via buttons: compared to younger individuals, on average, they were 1.65 sec slower in the pick and 1.1 sec slower in the place actions respectively (Figure 9.3). Therefore, the revealed workload of the place phase plausibly built up on the fatigue that originated from the pick phase and continued to increase over time, reaching significantly higher levels in the last part of the operation than younger participants. For what concerns the higher pupil size of senior participants observed at task-level in the action-based condition, instead, such a result seems to be in line with the previously mentioned physical difficulties demonstrated by the senior group. Indeed, when driving a virtual robot over the same task repeatedly, using physical actions as a control system seems to be particularly strenuous for older individuals, who thus reveal higher pupil size variation toward the end of the physical teleoperation.

Furthermore, the analysis of the perclos throughout the task execution allows the discussion of vigilance-related dynamics. Literature suggests that low levels of perclos typically reflect the user's tendency to reduce eye closures to avoid missing upcoming visual information (Marquart et al., 2015). First, we found that our task load manipulation affected perclos only in the young group, particularly when driving the robot via actions. Our interpretation in this regard is that, in our action-based condition, younger participants could afford to lower their level of vigilance in the single-task compared to dual-task. In contrast, older individuals always needed to maintain high levels of vigilance to provide proper visuomotor coordination during action-based robot control.

Overall, although being a more efficient modality for teleoperating a virtual robot, the findings presented so far demonstrate how the action-based controls particularly challenge senior users' workload capacities, and more in general, that senior workers might experience higher workload compared to younger ones when teleoperating a virtual robot. On the other hand, there were also conditions in which age did not affect users' workload or vigilance. Specifically, pupil size variations did not differ between age groups in any of the trials (trials level) performed under dual-task, and also within the whole task (task level) executed via button-based controls. In addition, any of the contrasts between young and senior participants reached the significance threshold in the analysis of perclos, and similar mental demand was self-reported in the NASA-TLX after all tasks. The latter result is in contrast with previous research (Grabowski et al., 2021), and is particularly important for the interpretation of users' implicit (as registered via eye-tracking) and explicit (as reported in the self-reports) workload. Indeed, the higher levels of implicit workload revealed in action-based operations was not subjectively noted, which indicates a greater resilience of seniors compared to younger users.

9.4.3 Age, button- and action-based control systems

Another research question addressed in this study specifically asked whether the same tendency for a greater benefit of action- compared to button-based teleoperation systems observed in young users in our Study 2 (Chapter 7) also applies to older users. Our data seem to confirm this hypothesis in terms of efficiency of the teleoperation (i.e., time on task, error rate) and expressed preference in the self-reports, but not exactly in terms of implicit workload (i.e., pupil size). Indeed, it seems that, particularly when teleoperating the robot with the action-based control, older users can perform similarly to younger ones at the cost of a higher implicit workload. Yet, the advantages of the action- compared to the button-based controls are clear, in both groups. Broadly speaking, such effects could be due to the recruitment of intuitive and embodied sensorimotor activations (BRAUD et al., 2022) elicited by the direct physical control of the virtual robot in the 3D environment. Otherwise, when driving the robot via buttons, the user's spatial intents must be changed from a 3-dimensional perspective to two static axes (forward-backward, left-right), thus requiring a mental transformation. Despite the physical-related difficulties that come with age, which reveal in action-based teleoperations, such a control system is still more efficient than the button-based one. However, in this regard, we need to highlight some relevant points. Literature on this matter is contradictory: some reported that more natural control systems (e.g., action-based) could be advantageous, especially for elderly users (Bobeth et al., 2014), while others evidenced better performance and lower physical effort with traditional mouse systems (Gerling et al., 2013). In our case, young participants outperformed older ones in the action-based, but not in the button-based condition. Such result might be due to the fact that older individuals are prone to muscle pain and musculoskeletal dysfunctions (Ilmarinen, 2001), and they also strive to match similar motor strategies as young workers (Qin et al., 2014), so they naturally move slower. Younger individuals that lack physical and motor difficulties are instead particularly fast when using the action-based controls, likely uncovering age-related differences in operation times. In other words, the fact that such an advantage of action-based controls was particularly evident in young users was likely due to their ability to move faster and have better visuomotor coordination (Coats et al., 2016), also in VR.

9.4.4 Age-dependent differences in individual factors

Our participants reported a low level of experience in the field of robotics (i.e., only one young and one senior participant reported having already used a robot at least once). Mostly, they were also infrequent users of VR technology, regardless of their age. They also reported similar problem-solving skills, trust in technology and sense of presence. Therefore, we can assume that all differences observed between young and senior participants did not depend on these individual factors. Differently, it is important to notice that young users reported a significantly higher knowledge about VR and learnability skills compared to senior users, who thus lacked flexibility in learning how to use new technologies, such as VR. These factors could thus have contributed to the slowed teleoperation of seniors compared to younger participants.

9.4.5 Age-dependent relations between performance, workload and individual factors

None of the analyzed individual attitudes had a strong effect on the users' performance, not in the young, not in the senior group. In other words, our participants' operation time and error rate at the teleoperation task were not affected by their problem solving, learnability skills, knowledge or frequency of use of VR, and, contrary to what hypothesized in previous research (Maneuvrier et al., 2020; Toet et al., 2020), not even by the sense of presence. However, a moderate relation between trust in technology and operation time was observed for the senior participants. Specifically, the higher the trust in technology, the longer the operation times at the pick-and-place task. One interpretation might be that those individuals that trust technology the most are also those that put higher commitment in the task, eventually prolonging the task duration.

On the effects of individual attitudes on users' workload, instead, none of them significantly correlated with the reported levels of mental demand or effort, which are usually the most relevant aspects in the work and manufacturing sector. However, particularly in young participants, the higher the learnability skills, sense of presence and knowledge about VR, the lower they perceived the temporal demand during the teleoperation task. All relative R coefficients indicated moderate correlations between these factors. Interestingly, the same was not observed for the older group, except for the sense of presence, which weakly and negatively correlated with the perceived temporal demand. Such a result would be explained by the fact that those who felt higher immersion and telepresence are less likely to feel the temporal pressure of the task. On the other hand, particularly in senior users, higher trust in technology helped reduce workload and vigilance (i.e., pupil size, perclos). These were the only strong correlations observed in our data, as indicated by their relative R coefficients, and generate important insights on how such an individual factor can influence the experienced task demand. Furthermore, pupil size was demonstrated to be significantly affected by the frequency of use of VR and problem solving in the young group, and learnability skills in both groups - despite with a reverse effect. In this regard, it is well known how task engagement, arousal and pupil size affect one another (Wang et al., 2018). Therefore, we can assume that the VR-based task was less exciting or engaging for those who used VR more frequently, likely generating smaller pupil size variations. Similarly, it is possible that greater learnability skills made the VR-based task more attractive, thus leading to higher pupil size variations.

9.5 Conclusions

What evinced by our research work is that senior users were capable of performing VR-based robotic teleoperations as accurately as younger ones, but at a slower pace. Furthermore, the use of action-based compared to button-based control systems allowed faster and more accurate operations, and it was also largely preferred by both groups, yet implying higher level of effort for senior users. Remarkably, overall participants were inexperienced with VR and robotics and reported high levels of trust in technology and sense of presence, which is known to promote performance in virtual environments. As regards workload-related aspects, net of a similar self-reported mental demand, implicit eye parameters revealed how senior workers got more fatigued in prolonged tasks and when higher visuomotor coordination was required. Such disparity between conscious perception and implicit data (i.e., eye parameters) revealed that for maintaining the same level of accuracy (i.e., error rate) senior participants' needed more time to carry out the pick task. Such results shed light on the feasibility of VR-based industrial operations for all ages, and generate insights into the pros and cons of more natural teleoperation control systems (action-based) compared to the most traditional ones (button-based) in immersive environments.

9.5.1 Limitations

Nevertheless, we acknowledge the following limitations. First, our results were obtained from a sample of individuals that not only had rarely used VR before, but also never had previous experience with robot teleoperations. Testing experts in robotics and teleoperations, or at least industry employees, instead, might provide findings that would be generalizable to a working context. Similarly, using an expert sample when examining how individual attitudes and skills affect performance and workload could lead to more reliable findings. Nonetheless, our assumption is that, if results were promising on such an inexperienced sample, let alone on robot teleoperators that already know what an industrial robot usually does. Furthermore, our VR interface was meant to simulate robotic teleoperations but was not directly linked to the physical robot workstation. Testing the overall system would be essential as some technical factors such as the data transmission latency could impact user experience and sense of presence in VR. Finally, given its close relation with arousal, additionally assessing the perceived level of engagement could help to better understand possible relations between pupil size and task-related workload.

9.5.2 Future directions

The strengths of our study are enclosed in the diverse human factor aspects that characterize the older workers and that were here systematically assessed in a multimodal setup. We hence emphasized the needs and capabilities of older operators who are called to drive an industrial robot in VR by entangling behavioral and cognitive trends. Future research should test such new technological solutions for industry workers on both young and senior individuals, as they might reveal different trends in performance and/or workload.

Chapter 10

Study 5 - Is it my arm? Embodiment, sense of agency and ownership of an industrial robotic arm in VR

As addressed in Chapter 3, human-robot interfaces leveraging intuitive body motion and gestures, such as Body-Machine Interfaces (BoMIs), likely trigger mechanisms of embodiment. In these cases, users might forget about the mediator (interface) and act more naturally. This will eventually increase the transparency of the teleoperation, leading to higher telepresence and even altered body ownership (Kilteni et al., 2012a). All these mechanisms were proposed to potentially improve teleoperation performance (Toet et al., 2020), even though empirical findings supporting this hypothesis in the telerobotics sector are still missing.

In this research work, we thus cover these aspects and investigate whether and how mechanisms of embodiment are triggered in industrial VR-based telerobotics leveraging a BoMI. Specifically, we asked participants to drive the same VR replica of the industrial robot UR10e, but in this case, a high-precision task and a controller-free human-robot interface (i.e., BoMI) were implemented. The task simulated the action of cutting a metal plate via a robotic arm provided with a cutter. Indeed, the robot effector was purposely modified to take the shape of a pointer, and participants drove it along a specified path by simply using their own arm's movements, thanks to the use of a VR body tracker. Time and space human-robot inconsistencies were introduced to explore their single and combined contributions in generating a feeling of embodiment within our non-anthropomorphic robotic arm. Given the abstraction and the psychological nature of such mechanisms, we relied upon a multimodal setup for measuring both explicit and implicit metrics related to embodiment, presence, workload and motor control. Specifically, in addition to various self-report questionnaires, we collected performance data computed through the VR device, in combination with a Mobile Brain/Body Imaging (MoBI) approach for additionally measuring brain dynamics throughout the task execution. This study was conducted during the research visiting at the Berlin MoBI Laboratory (BeMoBIL), which offers dedicated tools and approaches for measuring and analyzing EEG during free motion. Please, consider that the results illustrated in the present Chapter are preliminary.

10.1 Hypotheses and research questions

10.1.1 Embodiment, temporal and spatial inconsistencies

At first instance, we explored the single and combined roles of human-robot temporal and spatial consistencies in generating embodiment into our non-anthropomorphic robotic arm. Literature on this matter is resumed in subparagraph 3.1. We thus designed task conditions with temporal (i.e., visuo-motor desynchronization, induced via delayed robot's movements) and spatial (i.e., spatial dislocation, created via a spatial offset of the displayed robotic arm) inconsistencies, and therefore asked if and to which degree the sense of embodiment is affected by any of them (i.e., visuo-motor desynchronization, spatial dislocation) and also by their combination. We relied on self-reports and expected the visuomotor desynchronization to create disembodiment/loss of embodiment in most of the embodiment dimensions (i.e., agency, ownership, self-location, control) (Aymerich-Franch and Ganesh, 2016; D'Angelo et al., 2018; Farizon et al., 2021; Kokkinara and Slater, 2014). Furthermore, for the spatial dislocation, literature is quite contrasting as in some cases it reported that a spatial manipulation would affect only the perceived self-location and ownership (Ratcliffe and Newport, 2017), while in other cases it didn't affect embodiment at all (Miura et al., 2021; Newport et al., 2010). Therefore, we explored effects of spatial inconsistencies over the self-reported embodiment and generally expect it to have a smaller influence then the temporal inconsistencies (Newport and Preston, 2011; Pritchard et al., 2016; Ratcliffe and Newport, 2017).

10.1.2 Embodiment and teleoperation performance

Additionally, there are researches advancing the hypothesis that higher embodiment is related to a better teleoperation performance (Toet et al., 2020). However, to the best of our knowledge, proper evidence from telerobotics that supports this hypothesis is still missing. We thus explore this aspect by leveraging self-reports on the sense of embodiment and users' performance (i.e., accuracy and velocity of the teleoperation task). We specifically suppose that higher embodiment is related to higher motor control (Aymerich-Franch and Ganesh, 2016; Kokkinara and Slater, 2014; Iwasaki et al., 2022; Tsakiris et al., 2006; Verhagen et al., 2020), which will hence lead to a more accurate and fast teleoperation performance.

10.1.3 Embodiment and workload

Literature suggests that feeling embodied into a robot might rouse embodied sensorimotor mechanisms that would make the teleoperation more efficient while also reducing workload (BRAUD et al., 2022; Verhagen et al., 2020). A few studies systematically proved such a connection between embodiment and workload (Richard et al., 2021), while others deeply reviewed the theoretical endorsement coming from the embodied cognition approach that would support a relation between embodiment and workload (Verhagen et al., 2020). However, to the best of our knowledge, this hypothesis was not corroborated in the industrial telerobotics sector, especially when involving a non-anthropomorphic robot. We here tested this hypothesis and assumed that, for higher self-perceived embodiment, the level of workload will decrease significantly.

10.1.4 Neurophysiological signatures of embodiment and motor control

EEG event-related spectral perturbations (ERSPs) were previously leveraged to measure mechanisms of embodiment and motor control (Alchalabi et al., 2019; Ding et al., 2020; Evans and Blanke, 2013; González-Franco et al., 2014). All these studies (better unfolded in subsection 5.3.1) showed how a µ-ERD over centro-parietal brain areas positively correlates with a feeling of embodiment, particularly under synchronous multimodal stimulations. However, such effects were always observed when embodying humanoid external objects (Evans and Blanke, 2013) and, to the best of our knowledge, were never addressed in an industrial robotic arm. Therefore, we here assessed whether changes in the µ frequency band are modulated by the sense of embodiment also in our industrial robotic arm. In this case, we would expect stronger µ-ERDs over motor and sensorimotor areas for higher levels of self-perceived embodiment.

10.2 Methods

10.2.1 Sample

A prior power analysis conducted on Gpower (Erdfelder et al., 1996) indicated that, for our within-subjects experimental design, a total sample of 24 participants was needed to detect a medium effect size (d = 0.5) with 80% power. 29 participants, 15 females and 14 males (M_{age}= 26.75; SD_{age} = 4.21), voluntarily took part in the experiment after signing informed consent. They all reported being right-handed, having normal or corrected-to-normal visual acuity (via contact lenses), normal color vision, and no current or past neurological or psychiatric problems. Participation was in exchange for either course credits or a 14€ per hour monetary compensation. The experimental protocol was approved by the ethics committee of the Institute of Psychology and Ergonomics TU Berlin, and the study was conducted following the principles of the Declaration of Helsinki.

In this Chapter, we only present preliminary results on the first 26 participants, 13 females and 13 males (M_{age} = 26.92; SD_{age} = 4.41). Among these, 3 participants were excluded from the EEG analysis for excessive noise observed in the data.

10.2.2 Technical setup

VR device. All participants used the same VR headset employed in our previous investigations (HTC VIVE Pro Eye). Unlike previous setups, a VIVE tracker was placed over the participant's right wrist through the dedicated wristband, which is depicted in Figure 10.1. In this way, the human arm's movements were continuously tracked and translated into the robot's movements in real-time.

FIGURE 10.1: The VR body tracker applied on the participants' right wrist through the dedicated wristband. Picture from *https://www.amazon.de/-/en/dp/B07P93TSDK*

EEG device. Additionally, continuous EEG was recorded via 64 channels (BrainAmps, Brain Products) positioned in an elastic cap (EASYCAP) arranged according to the 10 percent system (Chatrian et al., 1985). One electrode below the left eye (vEOG) recorded vertical eye movements.

Experimental room. The experiment took place in a dedicated room of the BeMoBIL. A height-adjustable table was employed, which was regulated based on the participants' height. Participants stood in front of the physical table; when in VR, they saw a similar table in front of them (as depicted in Figure 10.2), which was displayed at the same height as the physical one. In this way, they could touch the VR table and simultaneously feel the perceptual feedback from the physical table.

10.2.3 Procedure, experimental tasks and design

Experimental scenario

We aimed to present the robotic arm UR10e in VR in a way that it could elicit a sense of embodiment in the participant. However, it has to be acknowledged that this robot is usually isolated and not attached to a humanoid-like

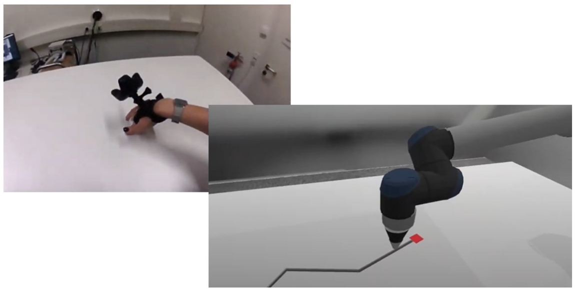

FIGURE 10.2: The physical table captured from the external cameras of the VIVE Pro Eye VR headset on the up left corner, and the virtual table concurrently captured from the virtual environment on the bottom right corner.

shape. Additionally, like many other industrial robots, it is provided with multiple joint angles allowing a higher degree of freedom in movements, while mismatching the humanoid shape of the arm and the related movement possibilities. In the end, this robotic arm completely mismatches the existing visual body representation.

In such a scenario, a feeling of embodiment could be facilitated via the synchronization of the own arm's movements and the robot's movements, through synchronous visuo-tactile perceptions, and through the spatial colocation of the human and the robotic arms. Therefore, we leveraged these aspects to increase the possibility for participants to feel embodied within our industrial robot. Specifically, the robotic arm was always shown from a 1st view perspective, as it is known to elicit higher sense of embodiment (Gorisse et al., 2017; Kilby and Whitehead, 2017). Furthermore, as shown in Figure 10.2, the robot came from the participants' right shoulder, and, for each movement of the human arm in the physical space, the robotic arm equally moved in the same direction of the virtual space. Finally, the presence of a physical table in front of the participants which matched the virtual table shown in VR allowed them to have congruent tactile feedback when touching the virtual table. To facilitate embodiment, we also suggested the participant to use their finger (as shown in Figure 10.2 and 10.3) to guide the robot over the worktable. Such an approach was adopted after the first experimental pilots, in which we observed that most of the participants started the experiment by using their hands in different ways (some using their fists, others placing all their fingers on the table), but most of them ended up in using their index finger, probably simulating the shape of the robot's pointer.

Experimental task

A high-precision teleoperation task was purposely designed for this experiment, in which participants were asked to guide the robotic arm through a path that was indicated on a metal plate placed on the worktable, right in front of the participants' view (as shown in the bottom right part of Figure 10.2). The structure of each trial is depicted in Figure 10.3. Specifically, each trial started with a *positioning phase*, in which a blue circle appeared on the left side of the worktable. Participants were thus instructed to place the robot pointer on the circle and keep their hands still for 500ms until it became green. Afterward, the path to follow appeared in front of them and the task started.

In the *task phase*, participants were asked to drive the robot over the indicated path from the start (green) to the end point (red) as indicated in Figure 10.3, while being as accurate and as fast as possible. The robot was complemented with a laser that reflected a green point on the worktable. In this way, they could always have an idea of their exact position on the metal plate. Indeed, as the robotic arm moved through the space, the robot's pointer also moved through the worktable. Once reached, the end point turned from red to green, the next blue circle of the positioning phase was enabled, and they could move on to the next trial.

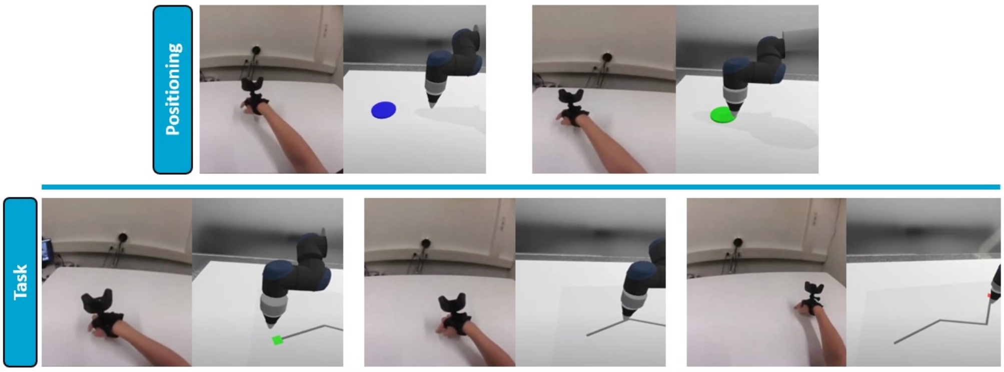

FIGURE 10.3: The steps of one trial of the teleoperation task. In the positioning phase, participants were asked to place the robot's pointer on the blue circle for 500ms until it turned green. Therefore, the task started, and they could drive the robot over the indicated path starting from the green point until the red point.

Notably, each presented path was composed of three segments of the same length, which created two angles: a narrow angle (110,63°) and a wide angle (142,11°). Figure 10.4 shows an example of a path used for the experimental task. We thus designed 7 possible paths (depicted in Figure 10.5) by combining the two angles in different orders and directions. Overall, in each task condition, each path was randomly presented 16 times, for a total of 112 trials. In the end, we analyzed 112 whole paths per condition. Furthermore, as each path was composed of the same wide and narrow angles shown in Figure 10.4, this allowed us to have 112 repetitions of the wide angle, and 112 repetitions of the narrow angle, significantly increasing the power of our analysis.

Experimental design and conditions

As depicted in Figure 10.6, the experimental design is a 2x2 within-subjects over the factors Time (synchrony, delay) and Space (co-location, offset). Each

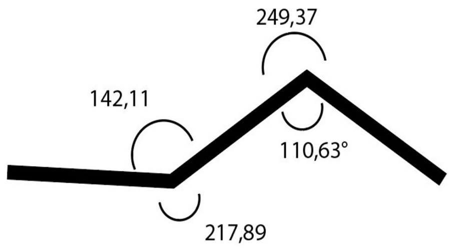

FIGURE 10.4: An example of a path used for the experimental task. All paths were composed of three segments creating two angles, a wide and a narrow angle. The numbers indicate each angle degree.

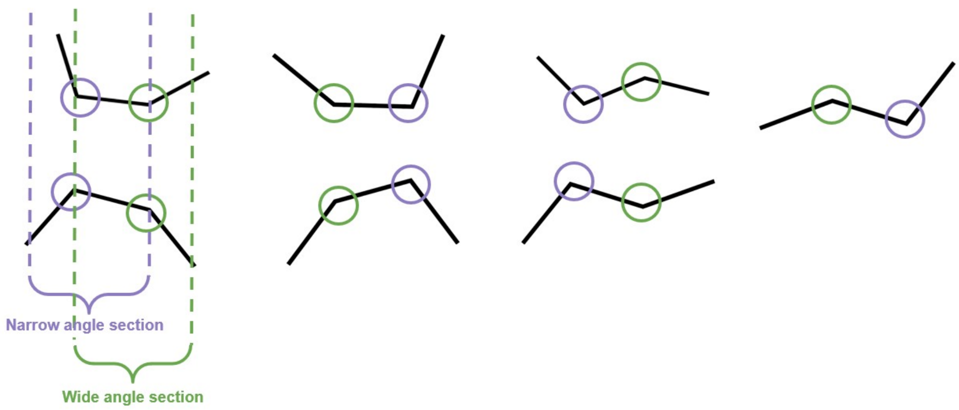

FIGURE 10.5: The seven paths designed for the present experiment. In each task condition, each of them was presented 16 times, for a total of 112 trials. Each path also included the same narrow (in violet) and wide (in green) angles, each of which was thus presented 112 times per task condition.

condition comprises 112 trials, each of which includes the positioning and the task phases depicted in Figure 10.3. The experimental conditions are explained as follows, and also depicted in Figure 10.9:

- *No manipulation*: when moving the own arm, the robotic arm moves at the same speed and in the same direction; therefore, there is visuo-motor consistency (visuo-motor synchrony). Furthermore, the position of the robotic arm (and specifically of the robot's pointer) is co-located with the own arm (specifically with the own hand); therefore, there is spatial consistency too (co-location).
- *Temporal manipulation*: when moving the own arm, the robotic arm moves in the same direction but with 0.5 sec of delay; this creates a visuo-motor inconsistency (delay). Yet, the spatial consistency (co-location) is maintained.
- *Spatial manipulation*: there is a 20cm offset between the position of the robotic arm (specifically, of the robot's pointer) and the position of the own arm (specifically, of the own hand); this creates a spatial inconsistency (offset). Yet, the temporal consistency is maintained (visuo-motor synchrony).
- *Temporal+spatial manipulation*: there are both the offset between the position of the robotic arm and the position of the own arm, and the temporal delay between the own movements and the robot's movements; this creates concurrent spatial and temporal inconsistencies (offset and delay).

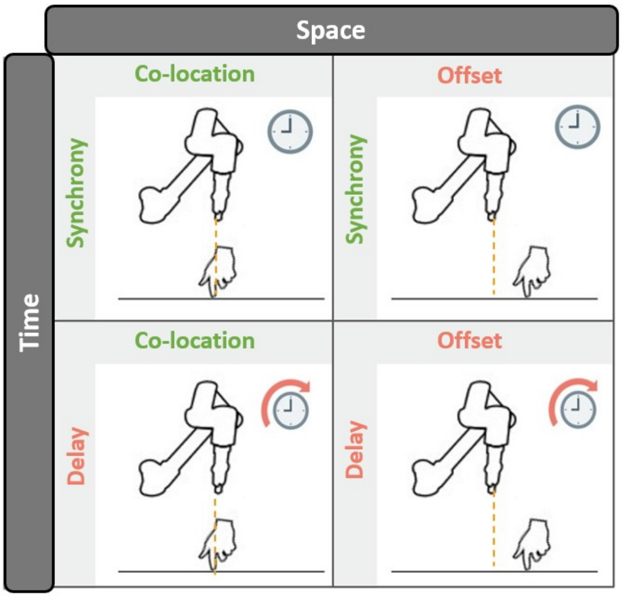

FIGURE 10.6: The experimental design.

Experimental procedure

After signing informed consent, participants were asked to fill out a demographic questionnaire. Meanwhile, both the VR environment and the physical table were adjusted based on their reported height. In this way, the whole setting was tailored to each participant, in a way that all were provided with the same view of the robot and the worktable. Thereafter, the participant was equipped with the EEG system in a dedicated room; the gelling session took about 20 minutes. Then, the experimental setup was finalized by equipping the participant with the tracker on their right wrist (shown in Figure 10.1) and the VR headset. The experimental instructions were all presented in text format in the virtual environment. Furthermore, all experimental phases (including tasks and questionnaires) were implemented in a unique VR program, in a way that participants could autonomously pass from one phase to the next one. The experimenter thus monitored the participant from the control room, without interfering with the task execution. At the beginning of the experiment, the participants were informed that they would have executed a **training phase** to familiarize themselves with the task, after which **four task blocks** would have followed. Each task block included an embodiment session, and a task session. At the end of each task block, participants always had the possibility to take a break. However, if they needed to stop the experiment in the course of a task block, they could raise their hand at any time and the experimenter immediately entered the room to help them. Finally, after the last task block, participants were provided with the last questionnaires and were provided with all the necessary products to wash their hair. The whole experiment, from the participant's arrival to the cleaning part, took about 2,5 hours on average. Please, refer to Figure 10.7 for a schematic overview of the task flow.

Task training. During the task training, participants were provided will all information needed for performing the task, and then executed three trials without manipulations (i.e., visuomotor synchronization and spatial colocation). Most importantly, they were not informed that some space and time manipulations would follow in the experimental session. Starting from the training, participants were always presented with a first-person perspective of the robotic arm. Specifically, as shown in Figure 10.2, the robot position and control system conveyed the impression that it was the participants' own arm: it came from the participants' right shoulder, and, for each movement of the human arm in the physical space, the robotic arm equally moved in the same direction of the virtual space. In case they needed to repeat the task more than three times during the training, they could repeat it as many times as they needed. Only when they confirmed to have understood all the task rules, they then autonomously passed to the actual experiment.

Embodiment session. At the beginning of each task block, during the embodiment session, the experimenter entered the room to induce embodiment (*inducing embodiment* in Figure 10.7). Particularly, we used visuo-motor and

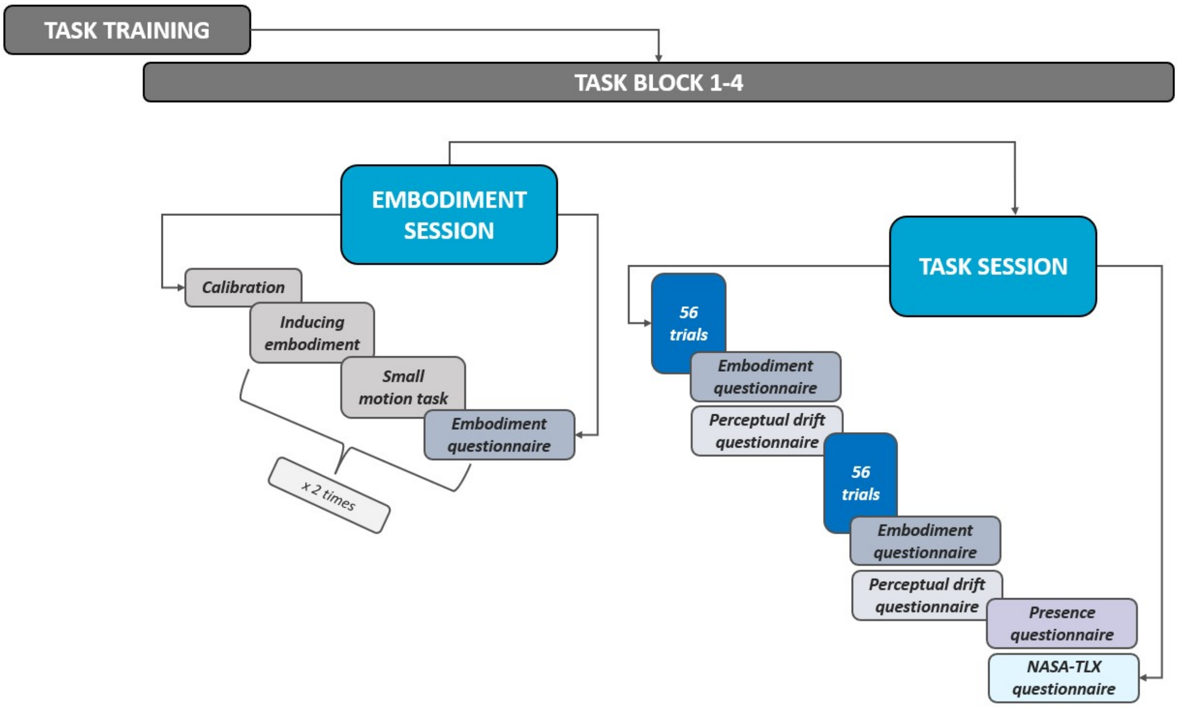

FIGURE 10.7: A schematic overview of the task flow

visuo-haptic stimulations, which are standardized methods that were previously proven to induce a feeling of embodiment in external objects (Farizon et al., 2021; Iwasaki et al., 2022; Seinfeld et al., 2022). Specifically, a blue circle appeared on the worktable, and participants were asked to place the robot pointer on that circle and keep it there, waiting for the experimenter to approach them. This phase is called *calibration* in Figure 10.7. Then, in the *inducing embodiment* phase of the same Figure 10.7, the experimenter used a body tracker with a plastic pen attached to it, which was reflected in the virtual environment as a white hand. Therefore, in the physical environment, the experimenter approached the participant's hand and brushed it with the plastic pen. Concurrently, in the virtual environment, the participant saw his robotic arm in front of him/her, and the experimenter's white hand approaching and touching the robot pointer. As shown in Figure 10.8, at the exact time when the white hand touched the robot pointer, the participant concurrently felt a touch on his/her hand, thus creating perceptual integration of visual and haptic information. Afterward, a *small motion task* (Figure 10.7) was introduced, in which five numbered circles were shown on the worktable, and the participant was asked to touch them in ascending order with the robot pointer. This task was designed to further create perceptual integration of visual and motor information, as it allowed to exert control on the robotic arm by directly moving the own arm. Finally, an *embodiment questionnaire* was presented directly printed on the worktable. Participants could answer the question by always using the robot's pointer. In this way, the feeling of embodiment created during the *inducing embodiment* and the *small motion task* was not disrupted. The whole embodiment session (from the *calibration* to the *embodiment questionnaire*) was repeated twice to establish a good level of embodiment before starting the task session.

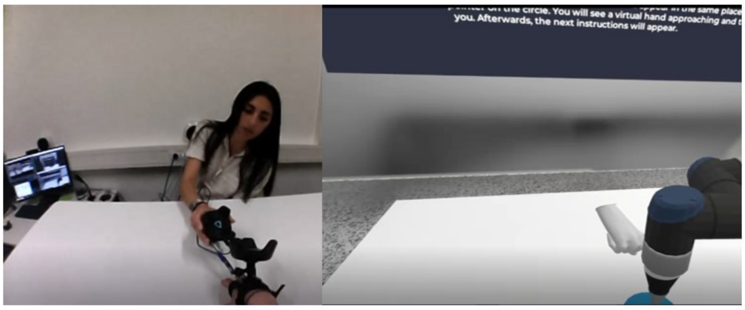

FIGURE 10.8: Embodiment session. On the left, the experimenter touching the participant's hand with a plastic pen; on the right, the concurrent participant's view in VR, who saw a white hand touching the robot pointer. This picture was captured during one data collection on an actual participant, using the software OBS.

Task session. Then, during the task session, participants autonomously executed a total amount of 112 trials in two phases, each including 56 trials, which were interspersed with the *embodiment questionnaire* and the *perceptual drift questionnaire*. Each trial was structured as shown in Figure 10.3. Once the last 56 trials were completed, the *embodiment questionnaire* and *perceptual drift questionnaire* were presented again, which were followed by the *presence questionnaire* and the *NASA-TLX questionnaire*. Based on the experimental design shown in Figure 10.6, each task session will be referred to as: *no manipulation*, *spatial manipulation*, *temporal manipulation*, *spatial + temporal manipulation* (see Figure 10.9).

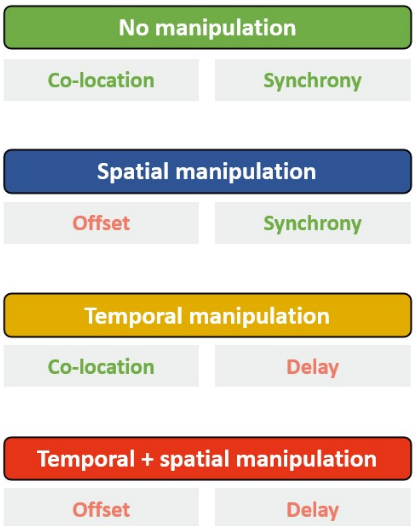

FIGURE 10.9: Task conditions.

10.2.4 Measurements

Self-reports measures

Embodiment questionnaire. Starting from questionnaires employed in previous studies (Farizon et al., 2021; Iwasaki et al., 2022), we here measured subjective feelings related to *appearance, ownership, agency, self-location* and *control*. This questionnaire was presented twice during the embodiment session, once after the first half of trials (n=56 trials) and once after the second half. We thus used the last questionnaire of the embodiment session as a baseline, and compared the answers with those given at half experimental task, and at the end of the task. In this way, we explored whether the sense of embodiment changed throughout the task execution. Specifically, participants were asked to respond to the following statements on a Likert scale from 1 to 7, from very low to very high.

- *Appearance*. I have the feeling that my arm started to resemble the robotic arm.
- *Ownership*. I have the feeling the robotic arm is part of my body.
- *Agency*. I have the feeling that the movements of the robotic arm could reproduce my own movements.
- *Agency*. I have the feeling that the movements of my own arm could reproduce the robotic arm's movements.
- *Self-location*. I have the feeling that my arm is in the place of the robotic arm.
- *Control*. I have the feeling that I can control the robotic arm's movements as it is my own arm.

Perceptual drift questionnaire. During the task session, a perceptual drift questionnaire was also administered to assess whether participants demonstrated changes in the perception of temporal delay and spatial offset with the task execution. Particularly, we asked the following questions once after first half of the experimental task, and once in the end of the task. Participants could answer the questions by sliding a cursor on a scale that went from 0cm to 50cm for the spatial drift, and from 0 to 1000ms for the temporal drift.

- *Spatial drift*. I have the feeling that the position of the robot pointer is at __cm from my hand.
- *Temporal drift*. I have the feeling the movements of the robot follow my movements after __ms.

Presence questionnaire. Furthermore, the presence questionnaire from Slater and Steed (2000) was administered at the end of each task block. Through this questionnaire, we aimed to explore possible relations between embodiment and presence. However, this aspect won't be addressed in the present thesis.

NASA-TLX questionnaire. As in all our previous contributions, we additionally administered the NASA-TLX questionnaire (Hart, 2006) to measure possible relations between perceived mental workload and i) temporal delay, ii) spatial offset of the robot, and, more in general, iii) the sense of embodiment.

Individual factors questionnaire. Before starting the experiment, participants were asked to answer the following questions: "Did you ever use VR?"; "If yes, how many times?"; "Would you consider yourself a practical user of VR?". These questions help in better defining the level of expertise of our sample in the VR domain.

Performance measures

For computing all performance measures, we considered the wide and narrow angles independently. Specifically, in the VR headset datalog, we saved the position of the robot pointer on the three axes (i.e., x, y, z) in each data frame (refresh rate of the VR headset: 120Hz). Furthermore, we saved the coordinates of the start, end and intermediate angles for each of the presented paths (Figure 10.5) in each trial. In this way, for each angle, we were able to rotate the spatial data of the robot pointer to match them over one unique angle, composed by two segments. In the end, we obtained 112 performance executed by each participant for each condition (i.e., *no manipulation*, *spatial manipulation*, *temporal manipulation*, *spatial + temporal manipulation*), and for each angle (i.e., narrow angle, wide angle, shown in Figure 10.10).

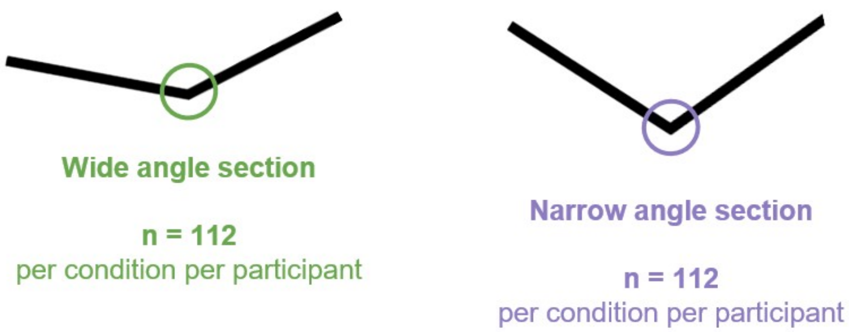

FIGURE 10.10: The number of performance repetitions obtained after rotating the spatial data.

Error distance. After rotating the spatial data, for each frame, we computed the error distance of the robot's pointer from the indicated path. For better visualizing and analyzing the data, we considered the absolute value of error distance. As that the VR headset recorded the behavioral data at 120Hz, we thus obtained a continuous index of precision of the teleoperation task.

Velocity. Furthermore, we computed the velocity at which the participant drove the robot over the indicated path. For computing this metric, we first applied dynamic time warping on the continuous behavioral data, in a way that they fitted 40 data points for each angle (20 data points for each segment composing the angle). Then, for each datapoint, we computed the velocity as the spatial distance executed during the data point divided by the time needed to execute that distance (space/time).

Electroencephalography (EEG)

For exploring the µ-desynchronization over the teleoperation task, we computed event-related spectral perturbations (ERSPs). This method allows to analyze changes in power spectrum of the EEG signal in response to an event, such as a stimulus or task, over time.

Pre-processing. Data were collected from 64 channels with a sampling rate of 500Hz and referenced to FCz. All impedances were kept below 10 kOhm. The data pre-processing was performed using the EEGLAB toolbox (Delorme and Makeig, 2004) in Matlab. Specifically, we used the *BeMoBIL Pipeline* from Klug et al. (2022), which provides dedicated functions for analyzing EEG and motion data. Data from each participant were down-sampled at 250Hz. Channels that were contaminated with artifacts were rejected and then interpolated via the dedicated function of the *BeMoBIL Pipeline* (Klug et al., 2022). We then copied the dataset (which we will call dataset A) and created a second dataset (which we will call dataset B). Dataset B was first filtered with a 1Hz high-pass filter and then underwent the AMICA processing (Palmer et al., 2012) to identify independent components (ICs) reflecting eye-movements activity. The ICs components were thus automatically classified via IClabel (Pion-Tonachini et al., 2019) in the following classes: 'brain', 'muscle', 'eye', 'heart', 'line noise', 'channel noise', and 'other'. Subsequently, we created a dataset C, from which we removed the ICs that were classified as 'eye' components with an accuracy equal to or higher than 70%. Therefore, the data was back-projected to the sensor level.

ERSPs. For computing the ERSPs, in each participant, we first removed all trials that exceeded 4SD from the averaged length of the trials' duration, as they were considered outliers. For each participant and in each condition, we computed the maximum trial duration for subsequently epoching the EEG data. We indeed epoched the data from 1sec before the event *'Successful Calibration'* (which corresponds to the moment in which the participant could start to drive the robot over the indicated path, after the 500ms of calibration period), until 1sec after the value of *maximum trial duration* previously computed in each participant and each condition. The reason why we opted for running this procedure for each condition independently is that the temporal and spatial manipulations (i.e., temporal delay and spatial offset respectively) potentially affected the velocity of the teleoperation, and therefore, also the trial length. We then identified the bad epoches via the function *pop_autorej()* of the EEGLAB package and removed them. ERSPs were finally computed for all selected channels via the function *newtimef()*, by using the median latency of the following events as anchors to timewarp the data:

- *Start*, which corresponds to the moment in which the participant placed the robot's pointer on the green point on the extremity of the path;
- *Angle 1*, which corresponds to the moment in which the robot pointer passed through the first angle of the path;
- *Angle 2*, which corresponds to the moment in which the robot pointer passed through the second angle of the path;
- *End*, which corresponds to the moment in which the robot pointer arrived at the red point at the other extremity of the path.

Finally, the ERSPs were baseline corrected by using the 500ms of Calibration time, which preceded the participant arm's first movement. Based on literature, we choose to focus on electrodes of the fronto-parietal regions and those covering the motor and sensorimotor cortex. Such locations are hypothesized to generate brain activity associated with illusory hand ownership and hand motor imagery (Kanayama et al., 2007; Munzert et al., 2009) and were investigated also in previous research assessing spectral EEG components related to embodiment and motor control (Alchalabi et al., 2019; Ding et al., 2020; González-Franco et al., 2014; Evans and Blanke, 2013). Furthermore, illusory ownership was demonstrated to modulate somatosensory evoked potentials at electrodes C3 and C4 (Peled et al., 2003; Press et al., 2008). Specifically, based on literature addressed in subsection 5.3.1, we were interested in exploring whether a stronger μ-desynchronization would reveal over contralateral motor and sensorimotor areas in conditions eliciting stronger embodiment (Ding et al., 2020; Evans and Blanke, 2013; González-Franco et al., 2014). Therefore, we specifically focused on the ERSPs over electrodes C3, CP3, P3, and additionally also F3 and FC3.

10.2.5 Statistical analysis

As a pre-analysis, we first fitted the data through the function *descdist()* of the package *fitdistrplus* (Delignette-Muller and Dutang, 2015) and chose the appropriate model setting based on data distribution. Moreover, for interpreting all post hoc contrasts, we always applied the Bonferroni method (Bonferroni, 1936). In the present thesis, for the performance data, we only present statistical analysis on the wide angle. Furthermore, for the ERSPs, we only present the preliminary plots and describe them via visual inspection and descriptive statistics.

Self-reports measures

Embodiment questionnaire. The embodiment questionnaire was administered four times in the course of each task session: twice during the *Embodiment session*, and twice during the *Task session* (see Figure 10.7). Here, we consider the last embodiment questionnaire of the *Embodiment session* as *baseline* questionnaire, the one in the middle of the *Task session* as *mid task* questionnaire, and the one administered at the end of the *Task session* as *end task* questionnaire. For analyzing the self-reported embodiment, we thus run one model for each dimension of the questionnaire (i.e., appearance, ownership, agency1, agency2, self-location, control). All questionnaire dimensions can be found in paragraph 10.2.4. Each model included the following factors: Condition (*no manipulation*, *spatial manipulation*, *temporal manipulation*, *temporal+spatial manipulation*) and Administration time (*baseline*, *mid task*, *end task*). To better visualize the results, we subtracted the baseline scores from all scores, so that, in the end, we looked at the embodiment score variations throughout the task execution (as shown in Figure 10.12). The raw data are shown instead in Figure 10.11.

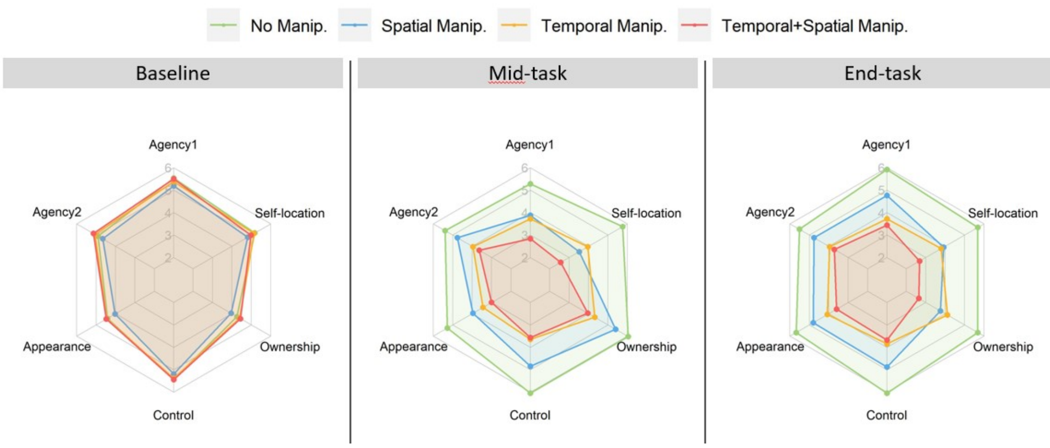

FIGURE 10.11: Averaged raw responses at each dimension of the embodiment questionnaire as reported right after the embodiment session (baseline), at half task (mid-task) and in the end of the task session (end-task). The colors indicate the task conditions.

Perceptual drift questionnaire. The perceptual drift questionnaire was administered twice in the course of each *Task session*: once after half of the trials, and once at the end of the *Task session* (see Figure 10.7). For analyzing the responses, we thus run a GLMM including the factors Space (offset, colocation), Time (delay, synchrony) and Administration time (mid-task, end-task).

NASA-TLX questionnaire. For analyzing the NASA-TLX questionnaire, we run a model including the factors Space (offset, colocation), Time (delay, synchrony) and Item (mental demand, temporal demand, physical demand, performance, effort, frustration). Furthermore, for exploring possible relations between the self-reported workload and the sense of embodiment, we computed correlation matrices between the scores reported at the NASA-TLX questionnaire and at the embodiment questionnaire by using Pearson's linear correlation test. Specifically, we first concatenated the embodiment and NASA-TLX scores of all conditions (*no manipulation*, *spatial manipulation*, *temporal manipulation*, *temporal+spatial manipulation*) in order to have full variability in the embodiment and workload scores. Thereafter, we computed the correlation matrix.

Performance measures

All performance measures (i.e., error distance, velocity) were analyzed via Generalized Linear Mixed Effect Models (GLMM) including the factors: Space (co-location, offset), Time (synchrony, delay) and Window (1, 2, 3, 4, 5, 6, 7, 8). Specifically, after applying the dynamic time warping, we divided the performance into 8 windows, to see whether the effects of temporal and spatial manipulations affected the teleoperation performance in time.

EEG measures

ERSPs. We further explored power spectral changes over time through the computation of ERSPs. Specifically, we were interested in observing motor-related mu-desynchronization dependently on Space and Time manipulations. Table 10.3 reports the descriptive statistics on the averaged power observed in the alpha frequency band (8-13Hz) in the windows *Start-Angle1*, *Angle1-Angle2*, *Ange2-End* and in the following channels: F3, FC3, C3, CP3, P3. These channels cover the contro-lateral pre-motor, motor and sensorimotor areas and were selected based on previous studies (Alchalabi et al., 2019; Ding et al., 2020; Evans and Blanke, 2013; González-Franco et al., 2014).

10.3 Results

10.3.1 Self-reports measures

Embodiment questionnaire. For all dimensions of the embodiment questionnaire, results evidenced a significant main effect of Condition (agency1: X^2 = 67.76, p<.0001; agency2: X^2 = 51.01, p<.0001; appearance: X^2 = 88.39, p<.0001; control: X^2 = 96.52, p<.0001; ownership: X^2 = 57.46, p<.0001; self-location: X^2 = 109.6, p<.0001) and a significant interaction between Condition and Administration time (agency1: X^2 = 35.7, p<.0001; agency2: X^2 = 39.9, p<.0001; appearance: X^2 = 43.02, p<.0001; control: X^2 = 58.52, p<.0001; ownership: X^2 = 56.99, p<.0001; self-location: X^2 = 52.38, p<.0001). In the following paragraphs, we report post hoc contrasts on the score as self-reported in the baseline vs. the mid-task and vs. the end-task in each experimental condition. Post hoc contrasts on the embodiment score between the different conditions as self-reported in the mid- and end-tasks (as also depicted in Figure 10.11) are instead summarized in Table 10.1.

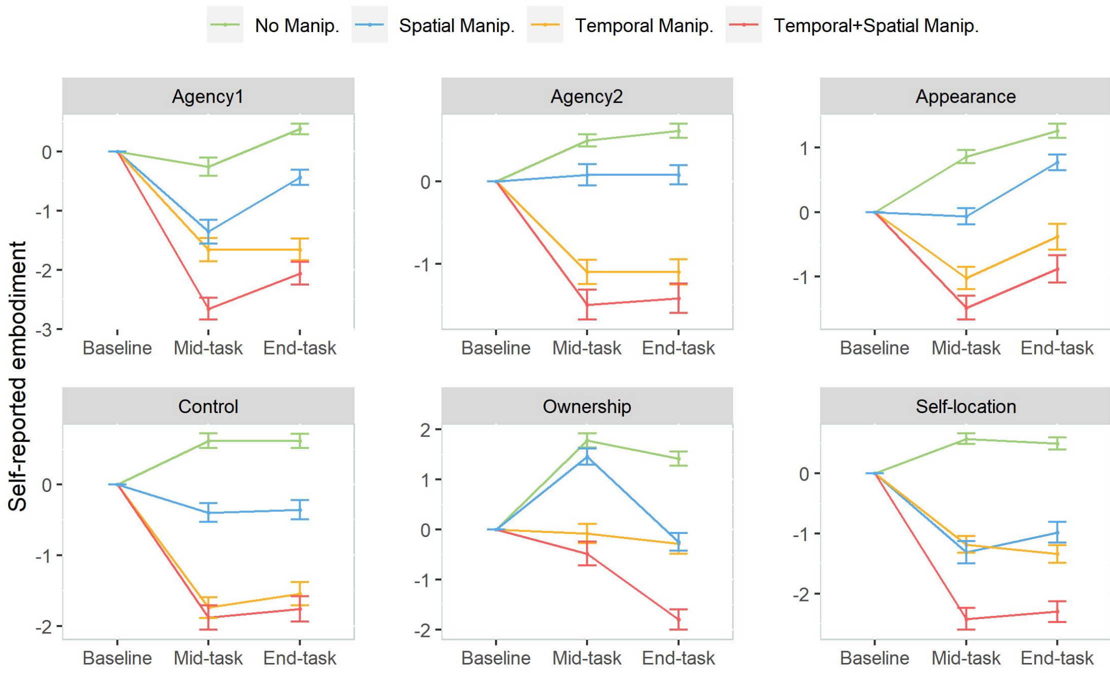

FIGURE 10.12: Averaged responses baseline-corrected at each dimension of the embodiment questionnaire. The bars represent the standard error.

TABLE 10.1: *p-values* from the post hoc contrasts between Condition and Administration time in each of the embodiment questionnaire's dimensions. NM = *no manipulation*; SM = *spatial manipulation*; TM = *temporal manipulation*; TSM = *temporal+spatial mnaipulation*; ns = not significant.

	NM vs. SM		**NM vs. TM**		**NM vs. TSM**		**TM vs. SM**		**SM vs. TSM**		**TM vs. TSM**	
	Mid-task	*End-task*	*Mid-task*	*End-task*	*Mid-task*	*End-task*	*Mid-task*	*End-task*	*Mid-task*	*End-task*	*Mid-task*	*End-task*
Agency1	p<.05	p∼.05	p<.05	p<.001	p<.001	p<.0001	ns	ns	ns	ns	ns	ns
Agency2	ns	ns	ns	p<.01	p<.001	p<.0001	ns	ns	ns	ns	ns	ns
Appearance	p<.05	ns	p<.01	p<.05	p<.001	p<.001	ns	ns	ns	ns	ns	ns
Control	ns	p∼.05	p<.0001	p<.0001	p<.001	p<.001	ns	ns	ns	ns	ns	ns
Ownership	ns	p<.0001	p<.01	p<.01	p<.01	p<.0001	ns	ns	ns	ns	ns	p<.05
Self-location	p<.0001	p<.01	p<.0001	p<.01	p<.01	p<.0001	ns	ns	ns	ns	p∼.05	ns

When looking at the post hoc contrasts for the dimension **agency1**, we notice how the scores reported during the *no manipulation* and the *spatial manipulation* conditions did not differ significantly when moving from the baseline, to the mid-task and to the end-task. Differently, in both the *temporal manipulation* and *temporal+spatial manipulation* conditions, the self-reported agency1 decreased significantly in the mid- (temporal manipulation: p<.01; temporal + spatial manipulation: p<.0001) and end-task (temporal manipulation: p<.01; temporal + spatial manipulation: p<.01) compared to the baseline.

For the dimension **agency2**, instead, the self-reported scores did not differ significantly from the baseline to the mid-task, to the end-task, in none of the task conditions.

Post hoc on the dimension **appearance**, instead, demonstrated a significant increase from the baseline to the end-task in the *no manipulation* condition (p<-05), and a significant decrease from the baseline to the mid-task in the *temporal + spatial manipulation* condition (p<.05). No significant changes in this dimension were observed throughout the task in the *spatial* and the *temporal manipulation* conditions.

The post hoc contrasts on the dimension **control** revealed a significant decrease from the baseline to the mid-task in both *temporal* (p<.01) and *temporal+spatial* conditions (p<.01), and also from the baseline to the end-task in the same *temporal* (p<.05) and *temporal+spatial* conditions (p<.01). No significant changes in feeling of control were observed throughout the task in the *spatial manipulation* and in the *no manipulation* conditions.

The sense of **ownership** was self-reported as significantly higher from the baseline to the mid-task (p<.0001) and from the baseline to the end-task (p<.05) in the *no manipulation* condition. Similarly, higher ownership was reported in the mid-task compared to the baseline in the *spatial manipulation* condition (p<.05), which however then decreased significantly again from the mid-task to the end-task (p<.01). Finally, in the *temporal+spatial manipulation* condition, the self-reported ownership decreased significantly from the baseline to the end-task (p<.01). No significant changes in sense of ownership throughout the task were observed in the *temporal manipulation* condition.

Finally, post hoc tests on the **self-location** demonstrated a significant decrease from the baseline to the mid-task in the *spatial* (p<.05) and *temporal+spatial manipulation* conditions (p<.0001), and from the baseline to the end-task in the *temporal+spatial manipulation* condition (p<.0001).

Furthermore, as can be observed in Figure 10.11, the post hoc contrasts on the **mid-task** revealed significantly lower agency1 (p<.05), appearance (p<.01), control (p<.0001), ownership (p<.01) and self-location (p<.0001) when comparing the *no manipulation* condition with the *temporal manipulation* condition. Differently, when comparing the *no manipulation* condition with the *spatial manipulation* condition, significantly lower agency1 (p<.05), appearance (p<.05) and self-location (p<.0001) were observed. When comparing the *no manipulation* with the *temporal+spatial manipulation* conditions, the scores at all dimensions significantly decreased (agency 1: p<.001; agency2: p<.05; appearance: p<.001; control: p<.0001; ownership: p<.01; self-location: p<.01).

Finally, compared to the *no manipulation* condition, post hoc contrasts on the **end-task** revealed significantly lower scores in all dimensions in the *temporal manipulation* condition (agency 1: $p<.001$; agency2: $p<.01$; appearance: $p<.05$; control: $p<.0001$; ownership: $p<.01$; self-location: $p<.01$) and also in the *temporal+spatial manipulation* condition (agency 1: $p<.0001$; agency2: $p<.05$; appearance: $p<.001$; control: $p<.001$; ownership: $p<.0001$; self-location: $p<.001$).

Perceptual drift questionnaire. The analysis of the responses to the perceptual drift questionnaire yielded significant effects for the factors Time (X^2 = 17.30, $p<.001$) and Space (X^2 = 142.76, $p<.0001$), but not for their interaction, not for the factor Administration Time. Results are shown in Figure 10.13.

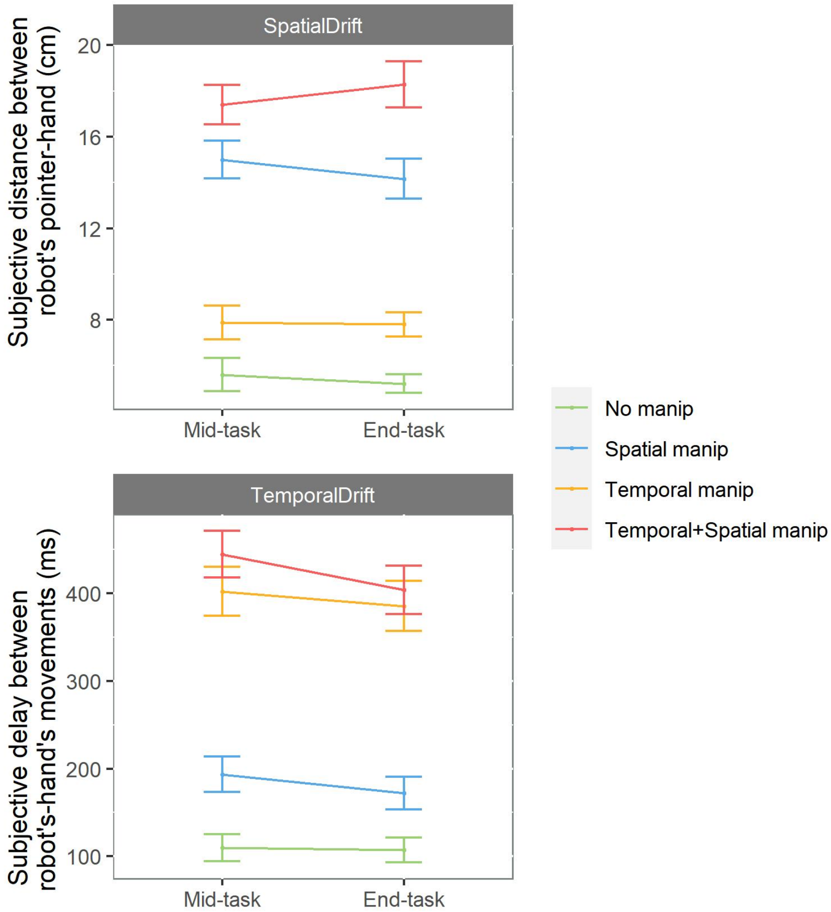

FIGURE 10.13: Averaged responses at the perceptual drift questionnaire, as reported in the mid-task, and end-task.

NASA-TLX questionnaire. Results demonstrated significant effects of Time (X^2 = 99.95, $p<.0001$), Space (X^2 = 14.04, $p<.001$) and Item (X^2 = 47.4, $p<.0001$). Furthermore, a significant interaction between Time and Item was observed (X^2 = 34.82, $p<.0001$). However, the full interaction between Time,

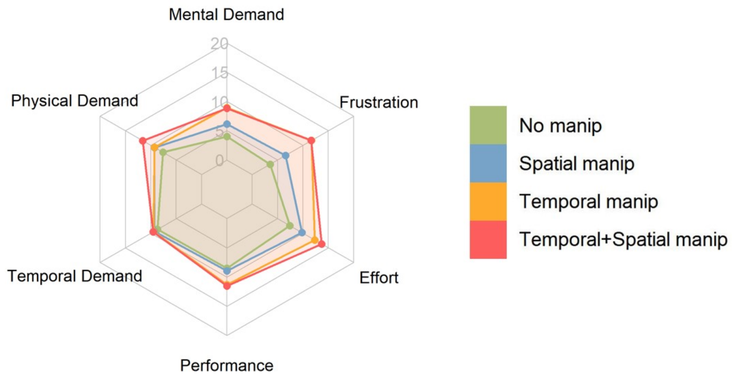

FIGURE 10.14: Averaged responses at the NASA-TLX questionnaire in each dimension and condition.

TABLE 10.2: Descriptive statistics of performance measures

	Error distance (cm)		Velocity (m/s)	
	mean	*SD*	*mean*	*SD*
No manipulation	0.86	0.15	0.3	0.09
Spatial manipulation	0.93	0.09	0.26	0.08
Temporal manipulation	1.25	0.20	0.18	0.04
Temporal+spatial manipulation	1.21	0.10	0.15	0.03

Space and Item, and also between Space and Item, did not reach the significance threshold. Post hoc on the interaction between Time and Item revealed that participants felt higher mental demand (p<.0001), effort (<.0001) and frustration (p<.0001) when a temporal delay was introduced. The averaged responses to the NASA-TLX questionnaire are depicted in Figure 10.14. The correlation matrix between the scores reported in the NASA-TLX and those reported in the embodiment questionnaire, instead, are depicted in Figure 10.15.

Individual factors questionnaire. In our sample, only 2 participants reported having used the VR headset more than 5 times, and 7 out of 25 participants reported that they would consider themselves practical users of VR devices.

10.3.2 Performance measures

Error distance. The analysis of error distance (upper plot in Figure 10.16) demonstrated significant main effects of Time (X^2 = 3367.10, p<.0001), Space (X^2 = 83.39, p<.0001) and Window (X^2 = 139.12, p<.0001). Furthermore, significant interactions were observed between Time and Space (X^2 = 24.9, p<.0001) and Time and Window (X^2 = 25.99, p<.0001). The post hoc tests revealed how the error distance was significantly higher in the 'temporal

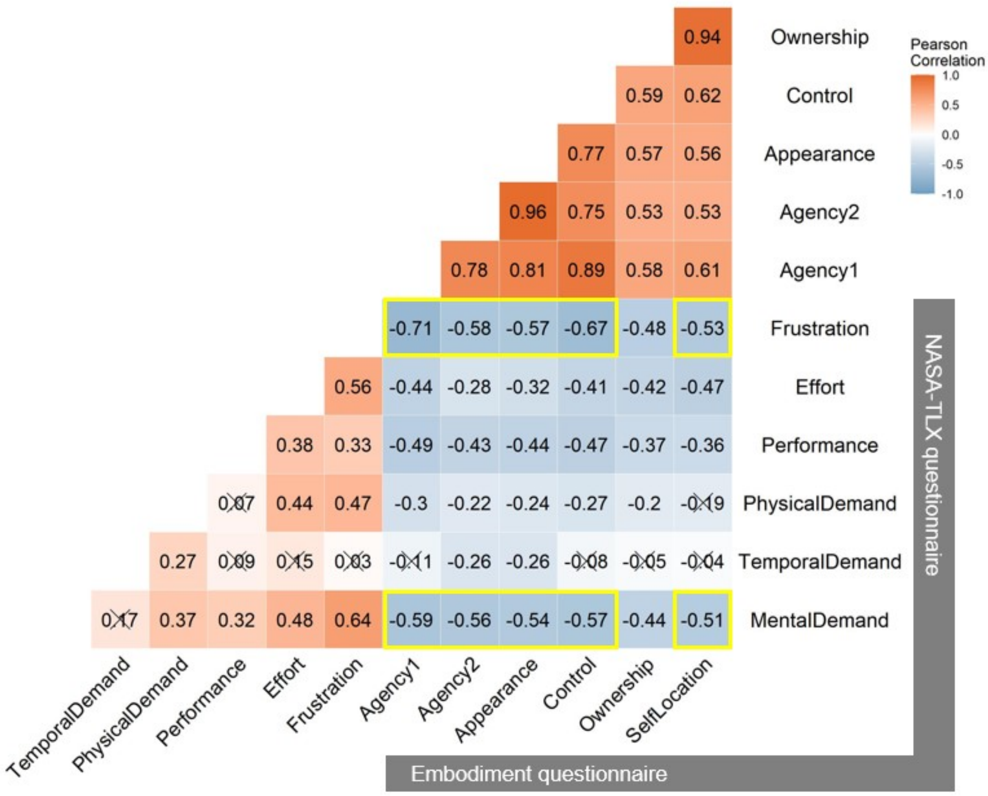

FIGURE 10.15: Correlation matrix between NASA-TLX questionnaires and embodiment questionnaires in all conditions. The numbers in each cell indicate Pearson's R coefficients. Not significant correlations are marked with an X (p>.05). The color of the cells reflects the values of Pearson's R: stronger correlations are characterized by more saturated colors, blue if negative, red if positive. Correlations that exceed an R of 0.5 are marked in yellow.

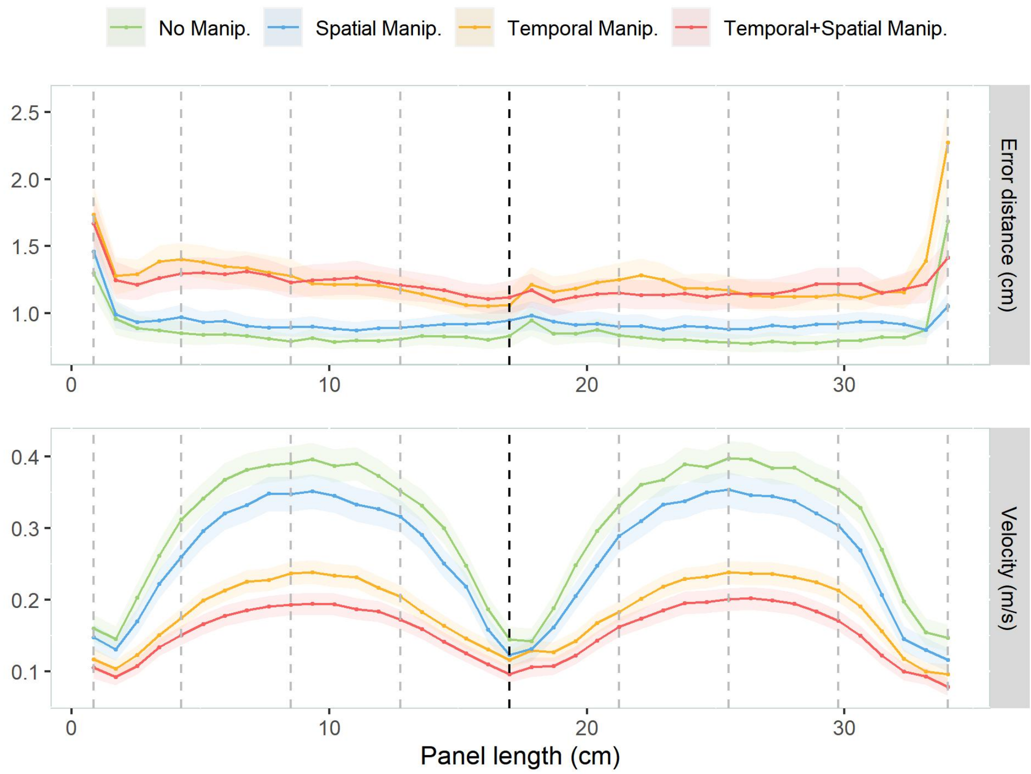

FIGURE 10.16: Error distance and velocity at the teleoperation task. The black dashed line indicates the wide angle; on its left, there is the first segment; on its right the second segment. All dashed lines indicate the windows created for analyzing the effects of our manipulations on participants' performance in time. The x-axis indicates the data points obtained via dynamic time warping, which were then re-converted in cm.

+ spatial manipulation' compared to the 'no manipulation' (p<.05) and the 'spatial manipulation' conditions (p<.05) but not compared to the 'temporal manipulation' condition (p>.05). Furthermore, the error distance in the 'temporal manipulation' was significantly higher compared to 'no manipulation' (p<.01) and 'spatial manipulation' (p<.05). Finally, the error distance in the 'spatial manipulation' did not differ significantly from the 'no manipulation' condition (p>.05). Descriptive statistics of the error distance are resumed in Table 10.2.

Velocity. The velocity at which participants executed the teleoperation task (depicted in the bottom plot in Figure 10.16) was significantly affected by Time (X^2 = 2268.4, p<.0001), Space (X^2 = 1559.83, p<.0001) and Window (X^2 = 4.5, p<.05). Furthermore, significant interactions were observed between Time and Space (X^2 = 62.42, p<.0001) and Space and Window (X^2 = 5.06, p<.05). The post hoc contrasts showed significantly slower operations in the 'temporal + spatial manipulation' compared to the 'no manipulation' (p<.0001) and the 'spatial manipulation' conditions (p<.001) but not compared to the 'temporal manipulation' condition (p>.05). Furthermore, participants in the 'temporal manipulation' were significantly slower compared to 'no manipulation' (p<.001) and the 'spatial manipulation' conditions (p<.01). Finally, the task velocity in the 'spatial manipulation' did not differ significantly from the 'no manipulation' condition (p>.05). The task velocity trend of each participant over the wide angle across conditions is also shown in Figure 10.17. Descriptive statistics of the task velocity are resumed in Table 10.2.

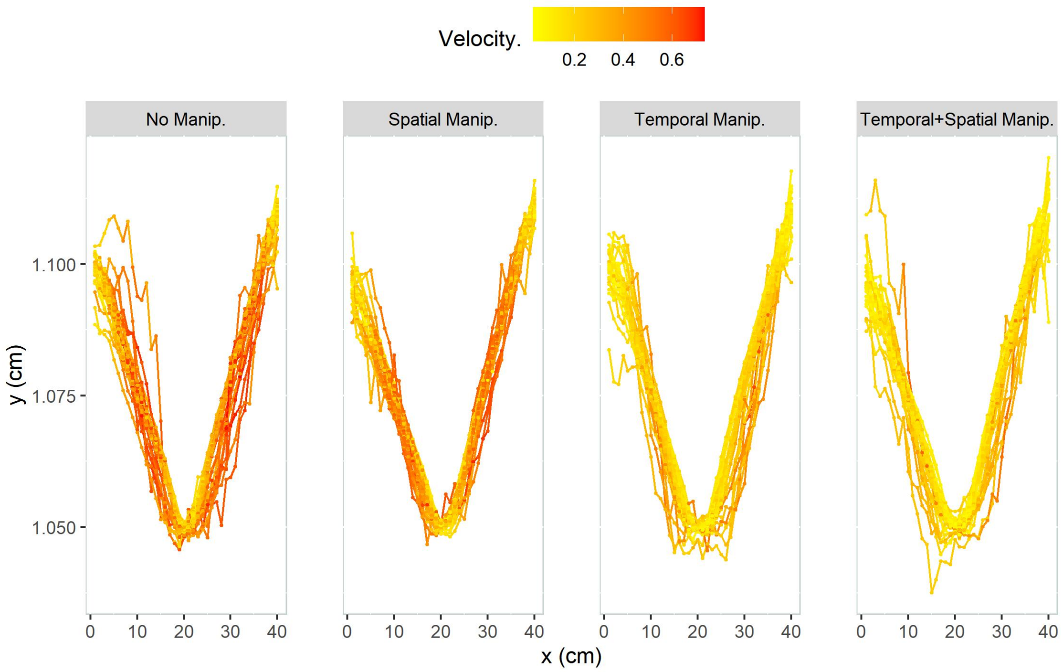

FIGURE 10.17: Velocity of the teleoperation through the wide angle. In each plot (one for each condition), each line represents the averaged path executed by each participant. The participants' paths are colored on a gradient from yellow (slower performance) to red (faster performance) based on the velocity at which they were executed.

10.3.3 EEG measures

ERSPs. The ERSPs computed for each experimental condition (i.e., *no manipulation*, *spatial manipulation*, *temporal manipulation*, *temporal+spatial manipulation*) are depicted in Figure 10.18. Furthermore, descriptive statistics relative to the averaged power in the μ frequency band (8-13 Hz) are shown in Table 10.3.

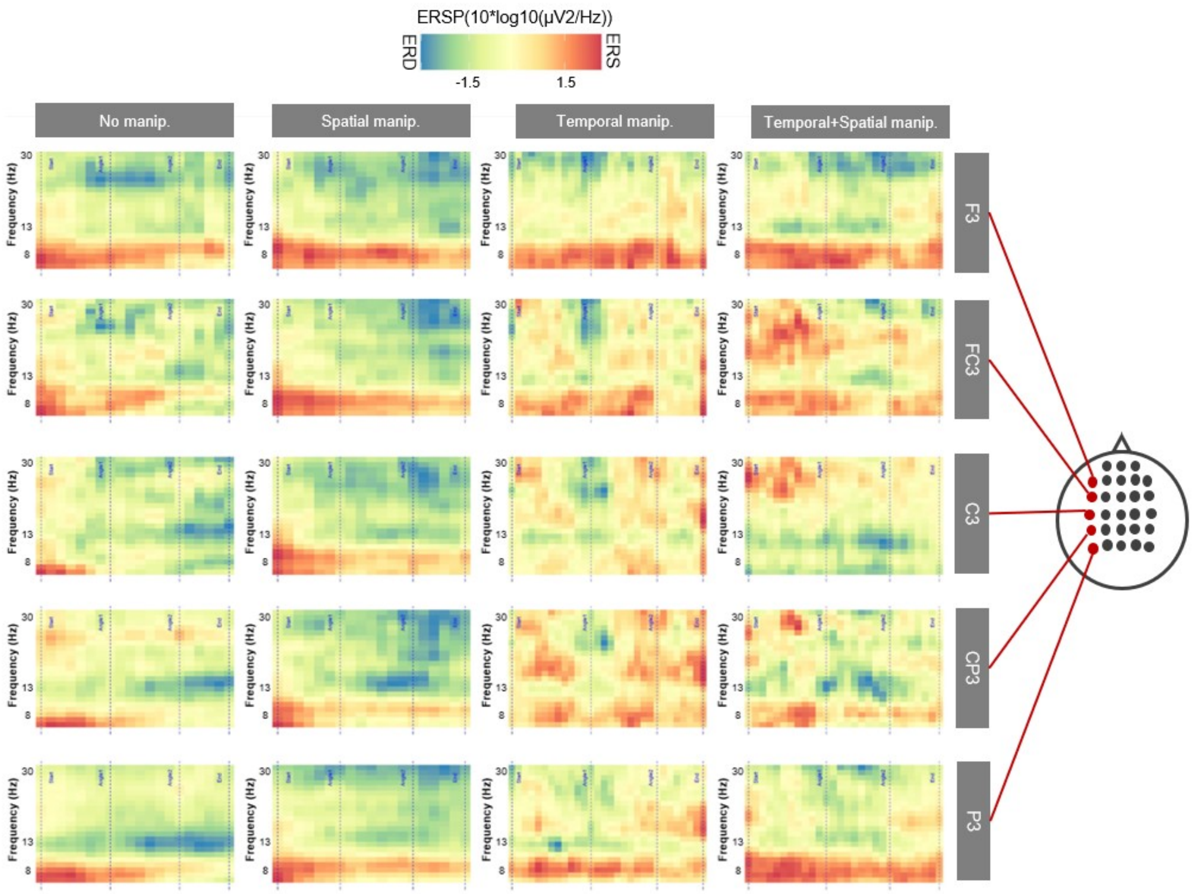

FIGURE 10.18: ERSPs baseline corrected depicted for each condition (i.e., *no manipulation*, *spatial manipulation*, *temporal manipulation*, *temporal+spatial manipulation*) and obtained from the following EEG electrodes: F3, FC3, C3, CP3, P3. The blue dashed lines indicates the following events: *start*, *angle1*, *angle2*, *end*. "Manip." is the abbreviation of "manipulation".

TABLE 10.3: Descriptive statistics of averaged μ power (8-13Hz) in each window obtained from the ERSPs

		No manip. *mean (SD)*	Spatial manip. *mean (SD)*	Temporal manip. *mean (SD)*	Temporal+Spatial manip. *mean (SD)*
Start-Angle1	F3	-0.25 (0.9)	0.2 (0.9)	-0.33 (1.1)	-1.15 (3.0)
	FC3	-0.45 (1.2)	0.13 (1.0)	-0.29 (1.2)	-0.28 (1.2)
	C3	-0.32 (1.3)	0.3 (1.1)	-0.36 (0.9)	-0.2 (1.2)
	CP3	-0.19 (1.0)	0.21 (1.0)	-0.11 (0.9)	-0.12 (1.0)
	P3	-0.15 (1.2)	0.09 (1.2)	0.006 (0.9)	-0.22 (0.8)
Angle1-Angle2	F3	-0.32 (0.9)	-0.19 (1.0)	-0.11 (1.0)	-1.26 (3.0)
	FC3	-0.48 (1.2)	-0.22 (1.0)	-0.16 (1.2)	-0.55 (1.0)
	C3	-0.37 (1.3)	-0.09 (0.9)	-0.18 (1.0)	-0.39 (1.0)
	CP3	-0.29 (1.02)	-0.03 (1.1)	-0.07 (1.0)	-0.30 (0.8)
	P3	-0.32 (1.2)	-0.13 (1.1)	-0.007 (1.0)	-0.42 (1.0)
Angle2-End	F3	-0.39 (1.0)	-0.43 (1.0)	-0.29 (1.0)	-1.26 (2.8)
	FC3	-0.73 (1.3)	-0.42 (1.1)	-0.40 (1.0)	-0.49 (1.0)
	C3	-0.58 (1.1)	-0.13 (1.1)	-0.31 (0.8)	-0.26 (1.0)
	CP3	-0.44 (0.9)	-0.07 (1.0)	-0.11 (0.8)	-0.24 (0.9)
	P3	-0.5 (1.2)	-0.23 (1.0)	-0.11 (0.8)	-0.40 (0.9)

10.4 Discussion

10.4.1 Embodiment, temporal and spatial inconsistencies

In this research, we were interested in exploring the single and combined roles of human-robot temporal and spatial consistencies in generating embodiment into our non-anthropomorphic robotic arm. In this regard, first, we want to point out how, as shown in Figure 10.11, the raw responses to the embodiment questionnaire at the baseline were consistently and quite high, particularly for the feeling of control, agency (specifically, agency1) and self-location. This suggests that, despite the appearance and shape inconsistencies between the human and the robotic arms, participants felt to can exert control on the robotic arm, to be placed within the robotic arm, and felt that the robotic arm could reproduce their movements. The results also show that the sense of ownership was self-reported as significantly higher from the baseline to the mid-task and from the baseline to the end-task in the no manipulation condition, which indicates how the more the participants repeated the same teleoperation task with our industrial robotic arm, the more sense of ownership gradually increased. However, when introducing the spatial offset, the sense of ownership decreased significantly throughout the task, as well as the sense of self-location. When introducing the temporal delay, the sense of agency and the feeling of control decreased the most, but all embodiment dimensions were significantly affected. Same, and even worse, happened when introducing both the temporal delay and the spatial offset, as all embodiment dimensions were negatively affected.

Such results suggest that both spatial and temporal manipulations can affect embodiment, but in different ways. Indeed, while a spatial offset between the own arm and the robotic arm to teleoperate significantly affects only self-location and ownership, a temporal delay between the own movements and the robotic arm's movements strongly disrupts the self-perceived embodiment in all dimensions, and similar effects are observed when coupling both spatial and temporal manipulation. The influence of a spatial offset over sense of ownership and self-location was previously observed by Ratcliffe and Newport (2017) when using humanoid hands, and was here reproduced with our non-anthropomorphic industrial robot. At the same time, our findings are not in line with those of Miura et al. (2021) and Newport et al. (2010), which observed no influence of a spatial dislocation on embodiment of humanoid hands or avatars. Oppositely, our findings on a general and stronger effect of a temporal delay between the own movements and the robot movement's over embodiment align with literature (Aymerich-Franch and Ganesh, 2016; D'Angelo et al., 2018; Farizon et al., 2021; Kokkinara and Slater, 2014). However, significant differences in the embodiment questionnaire scores between the *spatial* and the *temporal manipulation* conditions were not observed. It was also interesting to notice how, when combining temporal and spatial manipulations, the effects on the perceived embodiment were as strong as in the solely temporal and solely spatial manipulations, with the exception of the sense of ownership. Indeed, participants felt higher ownership in the temporal compared to the temporal+spatial manipulation at the end of the task, while no differences were found between the spatial and the temporal+spatial manipulations. Such a result could emphasize the impact that spatial manipulation has not only on the self-location, but also on the sense of ownership (Ratcliffe and Newport, 2017).

Additionally, by looking at Figure 10.13, we notice that no perceptual drifts were observed throughout the task, not in the space not in the time domain. This suggests that they did not get used to the temporal or the spatial manipulations, and consistently felt the spatial offset in the *spatial manipulation* and the time delay in the *temporal manipulation* conditions for the whole task.

10.4.2 Embodiment and teleoperation performance

There are researchers who advanced the hypothesis that a higher embodiment leads to a better teleoperation performance Toet et al. (2020); nonetheless, systematic evidence that supports such a hypothesis is lacking. We thus here explored this aspect, and specifically hypothesized that higher embodiment allows higher motor control (Aymerich-Franch and Ganesh, 2016; Kokkinara and Slater, 2014; Iwasaki et al., 2022; Tsakiris et al., 2006; Verhagen et al., 2020), which in turn will lead to a faster and more accurate teleoperation performance.

In our experiment, the sense of embodiment was manipulated via temporal and spatial inconsistencies (specifically, a temporal delay and a spatial offset between the participant's arm and the robotic arm). In the previous subsection 10.4.1, we discussed how such manipulations actually affected the sense of embodiment, with the temporal delay having a higher impact than the spatial offset, and the concurrent temporal and spatial manipulations having the greatest influence on the self-perceived embodiment. Therefore, to explore relations between embodiment and performance, we can consider the *no manipulation* condition as the one that elicited higher embodiment, followed by the *spatial manipulation*, *temporal manipulation* and finally the combination of *spatial+temporal manipulations*, which was observed to have the highest impact on embodiment.

By looking at the participants' performance dependently on these manipulations, we observed how the temporal manipulation affected both the velocity and the error distance: when introducing the temporal delay, our participants were slower and performed with lower accuracy (higher error distance) compared to the condition without spatial or temporal manipulations. Differently, when the spatial offset was introduced, both velocity and error distance did not differ significantly from the condition without manipulations. This suggests that the temporal synchrony between the own movements and the robot's movements is particularly important for allowing fast and precise teleoperation, while the spatial co-location of the own arm with the robotic arm does not play a prominent role when teleoperating an embodied industrial robotic arm in VR. Also, when considering that the *temporal manipulation* had a greater impact on embodiment than the *spatial manipulation* (see Figures 10.12 and 10.11, and Table 10.1), the assumption of a relation between performance and embodiment is likely corroborated. Furthermore, the combination of *temporal+spatial manipulations* (offset + delay) led to higher error distance and slower performance than the teleoperation task performed with only a spatial offset and without temporal or spatial manipulations. On the other hand, there were no significant differences in error distance or velocity when performing the task under both spatial and temporal manipulations, compared to the condition in which there was only a temporal delay. These results confirm our previous interpretation of a greater weight of the temporal delay in affecting the sense of embodiment, and now also VR-based teleoperation performance, compared to the spatial offset.

Notably, such effects were revealed independently from the part of the segment on which the participant drove the robotic pointer, as the factor Window did not yield any significant interaction with the other factors. However, by visually inspecting Figure 10.16, we can notice how particularly pronounced were the effects of *temporal* and *spatial manipulations* on the velocity of the operation, especially in the course of the segments. On the segments' extremities (i.e., *start*, *end*) and in the vicinity of the angle, instead, participants always slowed down independently from the spatial or temporal manipulations, probably to maintain motor control.

10.4.3 Embodiment and workload

Our results from the NASA-TLX questionnaire evidenced how, when executing a teleoperation task in VR, a temporal delay significantly increased the perceived mental demand, effort and frustration. Differently, no significant differences in the responses at the same items were observed when teleoperating the robotic arm with a spatial offset from the own arm. Only the overall score at the NASA-TLX resulted to be slightly higher under the spatial manipulations (M=9.74; SD=4.75) compared to conditions of co-location (M=8.56; SD=5.03), while a bigger difference in the overall NASA-TLX score was observed when introducing the temporal delay (M=10.08; SD=4.68) compared to executing the same task with visuomotor synchrony (M=7.49; SD=4.59). These are important results, as they evidence how the visuomotor desynchronization induced via temporal delay had a stronger impact on the perceived workload than a spatial dislocation between the own arm and the robotic arm. This result aligns with our findings of a higher impact of the temporal delay on the sense of embodiment and the teleoperation performance.

Also, it is possible that a direct link between the sense of embodiment and the level of perceived workload exists. For exploring such an assumption, we computed a correlation matrix between all scores reported at the NASA-TLX and those reported at the embodiment questionnaire in all conditions (see Figure 10.15.

Our results demonstrated that, for higher embodiment as reported in all the questionnaire dimensions, participants reported significantly lower frustration and mental demand. This sheds light on the importance of correctly inducing a sense of embodiment for teleoperators performing tasks in VR, as it can potentially reduce negative feelings such as frustration and mental demands. When looking at the literature, BRAUD et al. (2022) proposed that tangible and embodied interfaces can lower the level of workload perceived by operators working with Unmanned Aerial Vehicles systems. By following this assumption, (Richard et al., 2021) measured subjective embodiment within a humanoid avatar during a coloring task with force feedback, vibrotactile feedback, and no haptic feedback. The feedback that revealed to increase the sense of embodiment the most, was also the one that elicited the lowest levels of workload. Our results are thus in line with literature, with the peculiarity of being resulted from an industrial teleoperation context involving a non-anthropomorphic robotic arm.

10.4.4 Neurophysiological signatures of embodiment and motor control

With this study, we additionally investigated whether a stronger μ-ERD is generated over motor and sensorimotor areas in teleoperation conditions eliciting stronger embodiment into our industrial robot, similar to what is usually observed in embodiment experiments involving humanoid arms (Ding et al., 2020; Evans and Blanke, 2013; González-Franco et al., 2014). As addressed in subparagraph 10.4.1, the condition that elicited higher embodiment in our study is the *no manipulation* condition, in which there was spatial co-location between the own hand and the robot's pointer, as well as perfect synchrony between the own movements and the robot's movements. When introducing a spatial offset between the own hand and the robot's pointer, the sense of ownership and self-location significantly decreased. When introducing a time delay between the own movements and the robot's movements, and also when combining such a temporal delay with a spatial offset, all embodiment dimensions were strongly affected. We thus rely on such evidence to relate the different levels of embodiment as created by the *temporal* and *spatial manipulations* to the underlying neural activity in the spectral power domain (shown in Figure 10.18).

Although preliminary, our ERSPs evidenced how, in the *no manipulation* condition, a μ-ERD was clearly observed over motor and sensorimotor areas controlateral to the participants' moving arm (electrodes C3, CP3, P3), particularly through the end of the task and more evidently over more posterior regions. Such a power decrease is quite similar to what was observed when embodying external (and in some cases virtual) humanoid hands or arms (Ding et al., 2020; Evans and Blanke, 2013; González-Franco et al., 2014), likely reflecting embodied controls even in our non-anthropomorphic humanoid arm. Under *spatial manipulation*, instead, a weaker μ-ERD revealed over central (electrode C3) and parietal regions (electrode P3), whereas a stronger μ-ERD was observed over centro-parietal areas (electrode CP3), similar to the *no manipulation* condition. Interestingly, μ-ERDs were much weaker in all electrodes under examination in the *temporal manipulation* condition, particularly over the sensorimotor cortex (electrode CP3). Furthermore, some evidence for μ-ERDs emerged in the *temporal+spatial manipulation* condition over motor and sensorimotor areas (see electrodes C3 and CP3 particularly), even though earlier in time and less strong compared to the *no manipulation* condition.

Interestingly, in our study, all observed power desynchronization was revealed in the very high alpha frequency band, which could also be considered early beta (13-30 Hz). Furthermore, by visually inspecting Figure 10.18, we can also notice how an ERD also emerged in the late beta over the same motor and sensorimotor areas, particularly in the *no manipulation* and *spatial manipulation* conditions. There are examples in literature reporting how spectral power perturbations were observed in a frequency range exceeding the alpha band. For instance, Ding et al. (2020) investigated the effects of different vibrotactile feedback on the self-perceived embodiment. They found stronger ERD in the feedback conditions eliciting higher embodiment particularly in the high alpha frequencies, but also in the high beta (22-26 Hz). Similar results seem to emerge also in our study. Furthermore, Evans and Blanke (2013) induced illusory hand ownership in VR via synchronous visuo-tactile stroking on participants' hand, and compared the underlying brain dynamics with asynchronous stroking. Besides their main results on a decreased power in the μ frequency band, they also observed a significant effect of stroking synchrony on the beta (13-30 Hz) frequencies.

Overall, from first observation of our ERSPs, it seems that a power desynchronization similar to what was observed in embodiment studies also occurred in our industrial setting, and that a relation between μ-ERD and self-perceived embodiment likely exists. However, it must be acknowledged that we did not run any statistics on brain data. Therefore, our inferences are only based on visual inspections of ERSPs and relative descriptive statistics. Furthermore, such results are only preliminary, and further analysis will be conducted to confirm or clarify what are the brain dynamics underlying the embodied teleoperation of an industrial robot in VR.

10.5 Conclusions

With this study, we demonstrated how young and inexperienced participants were able to feel embodied into our industrial robotic arm in VR. In fact, despite the appearance and shape inconsistencies between the human and the robotic arms, they reported to can exert control on the robot, to be placed within the robotic arm, and felt that the robotic arm could reproduce their movements. Furthermore, we demonstrated how spatial and temporal inconsistencies between the location or the movement of the robot and of the own arm are crucial both for maintaining a high sense of embodiment within the robot, but also for performing faster and more accurately, and lowering the levels of frustration and mental demand. When comparing the two manipulations, the temporal delay had a significantly stronger impact on the participants' performance, workload and embodiment compared to the spatial offset between the own arm and the robotic arm. Finally, from a first preliminary analysis, it seems that the neural dynamics underlying the industrial robot teleoperation under embodiment are quite similar to those observed for the embodiment of external or virtual hands and arms. These results further corroborate the assumption that embodying a virtual industrial robot is possible as long as spatial, and particularly temporal consistencies between the robot and the own arm are maintained. These findings are crucial for HRI design and telerobotics, as the feeling of being embodied into the robot actually increased the teleoperation performance.

10.5.1 Limitations

We acknowledge the following limitations. Foremost, the presented analyses are preliminary, and further research is needed to either confirm the observed results or clarify their effects. Secondly, we are aware of the high abstraction of the embodiment construct, as it is strongly related to individual feelings that are often difficult to quantify. However, we took many actions with the intent of measuring embodiment with the best attention. For example, we induced embodiment via synchronous stroking of the virtual robot and the physical participant's hand at the beginning of each task block to have a baseline. We also designed the questionnaire so that the participant could respond directly by driving the robotic arm over the worktable to avoid disruption of embodiment. Remarkably, we additionally opted for a multimodal assessment including many self-reports, advanced bio-mechanical and performance analysis, as well as brain data collected via a 64channels EEG device to fully track behaviors, perceptions and implicit dynamics characterizing the teleoperators.

10.5.2 Future directions

At first instance, we will finalize the analysis of all collected metrics, particularly delving into the EEG data. We will thus run statistics on the brain data to confirm the results presented in the present thesis. We will also run correlations between the self-reported embodiment and the μ-ERD, in order to investigate whether there is an actual link between the power perturbation and the perceived level of embodiment. Furthermore, as the embodiment questionnaire was administered once in the middle of the task, and once in the end, as a next analysis, we will try to match ERSPs to the same task windows. In this way, we could explore whether the observed μ-ERD increases in time, as it happens with some dimensions of the embodiment questionnaire. Additionally, we will better analyze data on self-reported presence as well, both in relation to performance and embodiment. As telepresence is known to play a prominent role in telerobotics, its influence might be crucial in our VR scenario too.

Arguably, future research might better explore whether additional feedback (i.e., haptic feedback) can further increase the level of embodiment in our industrial robotic arm, and if this would impact the users' performance as well. Additionally, as also mentioned in our previous contributions, developing a direct connection between our virtual simulation and the physical robot UR10e is definitely a necessary step to assess the feasibility of our VR-based interface, its efficacy in generating a sense of embodiment and whether higher levels of embodiment positively impact users' performance also during real teleoperations.

Part III

DISCUSSION AND CONCLUSION

Chapter 11

Discussion and conclusions

With this thesis, we addressed relevant questions regarding VR-based industrial telerobotics. First, in Study1 (Chapter 6), we asked whether driving a virtual replica rather than the physical industrial robot UR10e benefits the user, in terms of both implicit and explicit workloads, and eventually also industrial organizations, by increasing the operation efficiency. Second, once ensured the benefits of VR-based simulations, in Study2 (Chapter 7) we asked which is the most efficient and less demanding teleoperation modality in VR, specifically assessing the effects of button-based and action-based control systems over performance and workload. In this respect, we assumed that the action-based controls would activate more intuitive and embodied mechanisms allowing better performance and lower workloads. Exceptionally, we also assessed the sensitivity of different eye-tracking metrics to workload, to better clarify whether and to which extent we can rely on VR-integrated eye-tracking for monitoring workload during teleoperations. Third, in Study3 (Chapter 8), we analyzed whether the users' performance is affected by individual factors, such as gender, individual skills and attitudes toward technology. Furthermore, we were interested in clarifying if the effects of individual factors over performance cut across different control modalities, namely the button- and the action-based one. Forth, in Study4 (Chapter 9), we specifically addressed the topic of aging, by asking whether senior users (>50 years old) can teleoperate well enough in VR, if their workload capacity differs from young users (<30 years old), and whether their individual attitudes toward technology have a higher impact on their performance and workload during VR-based teleoperations. Fifth, and finally, after observing the strong advantages of more direct and physical control modalities (i.e., action-based controls) over the more traditional button-based ones, we took the physicality of such control modalities to the next level, by removing all interaction interfaces. In Study5 (Chapter 10), we let participants guide the robot freely via their own movements in a first-person perspective and specifically asked whether they could feel embodied into our industrial robotic arm. We further asked whether a higher embodiment is related to better performance and lower workloads. Finally, given the abstraction and the psychological nature of embodiment, we additionally assessed whether the typical neurophysiological signatures of embodiment and motor control are also shown in our VR-based industrial scenario.

All our studies share some common threads, but also present critical methodological differences. Specifically, we consistently used the same **industrial robot UR10e**, and the same **VIVE-Pro-Eye VR device** through all the experiments. Such devices are commercially widespread and commonly employed in the manufacturing domain and in the entertainment/research sectors respectively. As a consequence, this makes our results highly applicable to concrete questions.

Furthermore, by accommodating advancements and proposals of the current fifth industrial revolution (Industry 5.0), we specifically addressed VR-based solutions for telerobotics by always undertaking a **human-centric perspective** (Lu et al., 2022). In all studies, in fact, we made humans the core of our investigations, by always shedding light on their behaviors, workloads, fatigue and mechanisms of motor controls when interacting and guiding the industrial robotic system. In doing so, we always leveraged a multi-method approach for better interpreting the human processes, offering a view on both implicit (e.g., eye parameters) and explicit (e.g., self-reported) dynamics that resulted in and from mental workload changes. Furthermore, we deepened the possible influence of different individual factors in different experiments, namely gender, gaming experience, learnability skills, problem solving, trust in technology, and finally age. In this way, we re-framed the worker as an individual with unique characteristics, attitudes, preferences and capabilities that, as such, finally gained a central role in the human-robot interplay.

As it is a prominent requirement of research, we always gave particular attention to **methodological rigor**. For example, we chose a simple experimental task (i.e., the pick-and-place task) for all studies except the last one, in order to guarantee appropriate experimental control while allowing a natural behavior with the least possible constraints. Such an approach follows the concept of natural or embodied cognition (Gallagher, 2006; Ladouce et al., 2017), which poses attention to the high relevance of conducting real-world-based experiments and always allowing natural movements and interactions between the user and the environment. Additionally, all statistical analyses were conducted in respect of the data, and all ethical standards were always accommodated. The reason why we changed the experimental task in Study5 is that, differently from the previous investigations, we aimed at specifically addressing mechanisms of embodiment and motor control. Therefore, the experimental task was re-designed for imposing high motor control on the user, but also for allowing us to get the most informative motion data on the velocity and accuracy of the teleoperation - thus reflecting motor control.

Another common methodological aspect regarding the experimental tasks that is worth mentioning is that, transversely to Studies 1, 2 and 4, we adopted the dual-tasking methodology for creating different levels of workloads. Therefore, in all studies, as a first methodological control, we systematically ensured that an actual difference between single- and dual-task emerged. The reader might have noticed that, while in Study 1 we created a dual-task by introducing a secondary arithmetic task in the auditory domain, in Studies 2 and 4 we instead opted for transposing the same arithmetic task in the visual domain. The reason behind this choice resides in the fact that, in Study 1, we observed how our task load manipulation affected the participants' workload (both as implicitly observed via eye-tracking parameters and explicitly reported in the self-reports), but not their performance. Therefore, to increase the task difficulty in the dual-task condition even more, we have leveraged the notions proposed in the resource theory (Pashler, 1984), according to which dual-task costs are particularly evident when multiple concurrent tasks share the same resources. Accordingly, while in Study 1 participants could rely on different sensory channels to execute the pick-and-place in the dual-task condition (i.e., proprioceptive, visual, auditory), in Studies 2 and 4 their focus was restricted to visual and proprioceptive information channels solely. As a result, our task load manipulation successfully affected both performance and workload.

Finally, while eye-tracking was leveraged in all conducted studies, it is worth mentioning the reason why we analyzed different eye parameters throughout the experiments. Specifically, in Study 1, we leveraged two different eye-tracking devices: the Pupil Labs in the physical environment, and the VR-integrated Tobii in the VR environment. As we didn't have access to the internal algorithms leveraged by Pupil Labs for computing blink parameters, whereas in the VR environment we had the possibility to build our own algorithms for defining blinks, we opted for only leveraging pupil size variations in both environments. Differently, in Study 2, which was entirely conducted in VR, we opted for programming our customized algorithms as explained in detail in the relative Method section 7.2.4, and measured pupil size, perclos and blink parameters. Further, we computed correlation matrices between these eye parameters and the self-reported workload to better define whether and to what extent each of the VR-based eye indices actually reflected workload variations. Based on results obtained in Study 2, we then only leveraged pupil size and perclos in Study 4.

All considered, our **main results** can be outlined as follows:

- Driving a robot via VR-based controls allows significantly faster operations compared to driving the same robot in the physical environment
- VR-based operations, compared to physical operations, are advantageous in terms of implicit workload, but this advantage is not self-perceived by the users
- The VR-based simulation is highly preferred over the physical robot
- Action-based controls allow faster and more accurate operations, are highly preferred, and also require lower workloads, compared to button-based controls
- When measuring workload via VR-integrated eye-tracking, pupil size and perclos are the most sensitive metrics to mental demand and effort

- Individual factors, namely gaming experience, gender, learnability skills, problem solving, and trust in technology impacted either performance or perceived workload only when driving the robot via button-based controls, but not via action-based controls
- Senior users are able to drive the virtual robot as accurately as younger ones, but at the cost of longer operation times and higher workloads, specifically through the end of prolonged task repetitions
- When driving our industrial robot in VR, senior users perceive higher physical demand compared to younger ones
- The same tendency observed in young users for higher preference and better performance when using action-based compared to button-based controls, also applies to senior users
- Senior users' implicit workload is more impacted after the repetition of action-based operations compared to button-based ones, although they don't perceive different workloads between the two control systems
- Young or senior users' teleoperation performance is not strongly affected by their individual attitudes
- Reporting high learnability skills, sense of presence and knowledge about VR helped young users in perceiving lower temporal demand, while reporting a high trust in technology helped senior users in reducing implicit workload and vigilance
- A generally high embodiment was reported when driving our industrial robot in VR from a first-person perspective, and without any temporal or spatial inconsistencies
- A time delay between the own arm and the robotic arm's movements has a greater negative impact on users' performance, workload, embodiment and motor control compared to a spatial offset between the robotic and the own arm
- When teleoperating our industrial robot in VR through a precision task, the task performance seems to be positively affected by the level of experienced embodiment
- The neurophysiological signatures of embodiment and motor control typically observed when embodying humanoid hands or arms seem to reveal in a similar way when embodying our industrial robotic arm in VR for simulated teleoperation

These results have a strong **impact both on work and industrial organizations**, and even largely on the **workforce society**. In fact, as this project builds on the industrial sector, it greatly impacts businesses and markets. It is true that all experiments only *simulated* working activities, and that our tasks were streamlined. However, all results come from a faithful reproduction of the commercially widespread robot UR10e, and it is a realistic example of virtualization of manufacturing systems. Furthermore, it is clear how the recent expansion of the *Metaverse* is fostering the utilization of VR devices. Therefore, any research conducted in VR, if uses a reasonably realistic virtual environment, could be easily transposed to practical contexts. Additionally, VR devices also allow full tracking of operators' actions, interactions, physical movements, and lately even eye movements, without motion constraints. Those characteristics make it possible to collect bunches of behavioral data during unconstrained teleoperations that, when processed and interpreted properly, can help explain or even predict human intentions online.

Also, understanding how such technical advancement can fit users with different skills (i.e., gaming, problem-solving, learnability skills), genders, and ages makes the introduction of VR into industries even more realistic. Broadly speaking, introducing VR-based teleoperations in industry can remarkably impact the working society, but this impact mainly depends on diverse individual characteristics. Above all others, **senior workers** who are susceptible to physical stress, like those on assembly lines, can surely benefit from such virtualization. For instance, operating in VR instead of on physical robots and machines keeps users from the risk of physical accidents and minimizes their physical effort. This would be exceptionally useful, especially for senior individuals that start to experience the signs of cognitive and motor decay, and therefore, constitute the weakest and most sensible part of the workforce. However, our results evidenced how these individuals, and particularly their workloads, require special attention. Indeed, when teleoperating our VR replica of the industrial robotic arm, the results on the performance benefits of action-based compared to button-based controls were crashing both in our young and senior participants. Nevertheless, only when leveraging eye-tracking technology, senior individuals revealed a higher implicit workload toward the end of the task block specifically when using action-based controls for teleoperating the robot. Most importantly, such a workload effect was not reflected in their self-reports, on the one hand highlighting their resilience to the work task, on the other hand also suggesting the importance of multimodal assessments of human factors. In this case, we learned how the introduction of VR-based teleoperations leveraging action-based controls would surely benefit work and industrial production's efficiency, but at a workload cost particularly for senior individuals. A possible solution to this issue could be the introduction of short breaks particularly for senior workers, in order to recover from the hoarding of workload generated by the prolonged execution of teleoperation task over time.

Being **informed on the user's workload** level throughout the work shift, as inferred via eye parameter, can thus help optimize working tasks and conditions (e.g., reducing working speed, suggesting breaks, etc.). This surely applies to those individuals that are most sensitive to workload, and particularly physical demands (i.e., senior workers), but also extends to the whole workers population in a generic way. Indeed, it has been proposed that the close integration of smart robotics systems' and workers' operations significantly increased the skill demands for the workers (Doolani et al., 2020). The relation between workload and performance has been widely established in different work contexts. Indeed, if a worker is given a workload that is too heavy, they may become overwhelmed and may not be able to complete all of their tasks to the best of their ability. This can lead to decreased productivity and may even lead to burnout. As a consequence, new challenges arose to find effective ways of monitoring the workload and fatigue in the human operator throughout the task. In this respect, using VR-embedded eye-trackers can be extremely helpful, and our results on the sensitivity of the VR-integrated eye-tracking device metrics to workload support this view.

Transversely to all studies, we can also argue on the effects that different VR control systems have on the users' teleoperation performance and workload, and particularly on those leveraging the most intuitive and embodied behaviors. An important motive to the present thesis was indeed the interest for understanding how to leverage **human motion as intuitive control of VR robotic systems**. There are studies that proposed how robotic interfaces leveraging intuitive body control and gestures might activate embodied sensorimotor mechanisms that would make the operations more efficient while also reducing workload (BRAUD et al., 2022; Toet et al., 2020; Verhagen et al., 2020). This is likely what also makes direct interactions between the humans and physical industrial robots so appealing, like in the case of collaborative robotics. And yet, the lack of cognitive research in the industrial telerobotics sector also translates into a lack of experimental evidence in favor of this view. As VR is the most interactive and flexible tool for excellence, using operators' direct and physical movements for guiding a robot over a trajectory becomes surely feasible in many modalities. Indeed, in Study2, Study3 and Study4, we kept the differentiation between button-based and action-based control modalities for teleoperating our VR replica of the industrial robotic arm. Through these studies, we always discussed whether an advantage of direct and physical control modalities (i.e, action-based) emerged respect to the button-based controls. And indeed, better performance and lower workloads were consistently observed in the action-based condition. Remarkably, as specifically demonstrated in Study 3, the effects of individual factors, skills and attitudes on the teleoperation performance and workload only emerged in the button-based condition, but not in the action-based one. Such effects were supposed to be due to the recruitment of intuitive and embodied sensorimotor activations elicited by the direct physical control of the virtual robot in the 3D environment, which made it more affordable for everyone. Otherwise, when driving the robot via button-based controls, the user's spatial intents needed to be changed from a 3D perspective to two static axes (forward-backward, left-right), thus requiring a mental transformation. For further clarifying whether sensorimotor mechanisms are actually recruited in action-based teleoperations, we brought the intuitiveness of the action-based control system to the next level, and eventually let the participant embody the industrial robotic arm from a first person perspective (Study5). In this case, through the additional assessment of brain dynamics along with behavioral, self-reports and eye data, we achieved a full and broad overview of what lies behind users' actions and performances. Specifically, the illusory experience that the robot becomes part of the own body has the exceptional potential of making the human-robot interface as imperceptible and almost non-existent to the operator (what is usually called *transparency*). In such cases, a better teleoperation performance can be actually achieved, as we are naturally inclined to have a better motor control on our own body rather then on an external interface. To our best knowledge, this is a particularly novel perspective in the industrial telerobotics sector, and yet, it is an incredibly important step toward the establishment of a direct relation between embodiment, attention and task performance, which could impact significantly the teleoperation industry.

Overall, while manufacturing and industrial robotics research have always been radically technical and engineering-based, we here started a fresh research line that offers the right cognitive counterpart to complement the technical one. Insights gained from behavioral and cognitive research can help improve the performance of operators and increase awareness of the workload and fatigue that users may experience. These insights can be used to develop strategies and technologies to mitigate the negative effects of workload and fatigue, ultimately leading to better performance and safety. Eventually, this will help industrial organizations to meet the needs of their employees by designing work systems, machines, robots, and teleoperation environments that are tailored to the workers, improving significantly their individual psychological well-being.

Bibliography

Abidi, M. H., Al-Ahmari, A., Ahmad, A., Ameen, W., and Alkhalefah, H. (2019). Assessment of virtual reality-based manufacturing assembly training system. *The International Journal of Advanced Manufacturing Technology*, 105(9):3743–3759.

Adami, P., Rodrigues, P. B., Woods, P. J., Becerik-Gerber, B., Soibelman, L., Copur-Gencturk, Y., and Lucas, G. (2021). Effectiveness of vr-based training on improving construction workers' knowledge, skills, and safety behavior in robotic teleoperation. *Advanced Engineering Informatics*, 50:101431.

Ahlstrom, U. and Friedman-Berg, F. J. (2006). Using eye movement activity as a correlate of cognitive workload. *International journal of industrial ergonomics*, 36(7):623–636.

Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6):716–723.

Alchalabi, B., Faubert, J., and Labbe, D. R. (2019). Eeg can be used to measure embodiment when controlling a walking self-avatar. In *2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*, pages 776–783. IEEE.

Ansari, S. Z. A., Shukla, V. K., Saxena, K., and Filomeno, B. (2022). Implementing virtual reality in entertainment industry. In *Cyber Intelligence and Information Retrieval*, pages 561–570. Springer.

Aracil, R., Buss, M., Cobos, S., Ferre, M., Hirche, S., Kuschel, M., and Peer, A. (2007). The human role in telerobotics. In *Advances in Telerobotics*, pages 11–24. Springer.

Argelaguet, F., Hoyet, L., Trico, M., and Lécuyer, A. (2016). The role of interaction in virtual embodiment: Effects of the virtual hand representation. In *2016 IEEE virtual reality (VR)*, pages 3–10. IEEE.

Aymerich-Franch, L. (2012). Can we identify with a block? identification with non-anthropomorphic avatars in virtual reality games. In *Proceedings of the international society for presence research annual conference*, volume 6.

Aymerich-Franch, L. and Ganesh, G. (2016). The role of functionality in the body model for self-attribution. *Neuroscience research*, 104:31–37.

Aymerich-Franch, L., Petit, D., Ganesh, G., and Kheddar, A. (2016). The second me: Seeing the real body during humanoid robot embodiment produces an illusion of bi-location. *Consciousness and cognition*, 46:99–109.

Aymerich-Franch, L., Petit, D., Ganesh, G., and Kheddar, A. (2017). Non-human looking robot arms induce illusion of embodiment. *International Journal of Social Robotics*, 9(4):479–490.

Bates, D., Mächler, M., Bolker, B., and Walker, S. (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological bulletin*, 91(2):276.

Bekrater-Bodmann, R. (2022). Mind over matter: Perceived phantom/prosthesis co-location contributes to prosthesis embodiment in lower limb amputees. *Consciousness and Cognition*, 98:103268.

Benedetto, S., Pedrotti, M., Minin, L., Baccino, T., Re, A., and Montanari, R. (2011). Driver workload and eye blink duration. *Transportation research part F: traffic psychology and behaviour*, 14(3):199–208.

Berndt, D. J. and Clifford, J. (1994). Using dynamic time warping to find patterns in time series. In *KDD workshop*, volume 10, pages 359–370. Seattle, WA, USA:.

Bobeth, J., Schrammel, J., Deutsch, S., Klein, M., Drobics, M., Hochleitner, C., and Tscheligi, M. (2014). Tablet, gestures, remote control? influence of age on performance and user experience with itv applications. In *Proceedings of the ACM International Conference on Interactive Experiences for TV and Online Video*, pages 139–146.

Boenzi, F., Mossa, G., Mummolo, G., and Romano, V. (2015). Workforce aging in production systems: Modeling and performance evaluation. *Procedia Engineering*, 100:1108–1115.

Bonferroni, C. (1936). Teoria statistica delle classi e calcolo delle probabilita. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commericiali di Firenze*, 8:3–62.

Borghini, G., Astolfi, L., Vecchiato, G., Mattia, D., and Babiloni, F. (2014). Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness. *Neuroscience & Biobehavioral Reviews*, 44:58–75.

Botvinick, M. and Cohen, J. (1998). Rubber hands 'feel'touch that eyes see. *Nature*, 391(6669):756–756.

BRAUD, V., COUTURE, N., BOVET, L., and FERRARI, V. (2022). Embodied cognition and tangible user interfaces: Alternatives for a reduction of unmanned aerial vehicles operators workload.

Brizzi, F., Peppoloni, L., Graziano, A., Di Stefano, E., Avizzano, C. A., and Ruffaldi, E. (2017). Effects of augmented reality on the performance of teleoperated industrial assembly tasks in a robotic embodiment. *IEEE Transactions on Human-Machine Systems*, 48(2):197–206.

Brough, P., Johnson, G., Drummond, S., Pennisi, S., and Timms, C. (2011). Comparisons of cognitive ability and job attitudes of older and younger workers. *Equality, Diversity and Inclusion: An International Journal*.

Bruggen, A. (2015). An empirical investigation of the relationship between workload and performance. *Management Decision*.

Bugalia, N., Sen, A., Kalra, P., and Kumar, S. (2015). Immersive environment for robotic tele-operation. In *Proceedings of the 2015 Conference on Advances In Robotics*, pages 1–6.

Bunce, D. and Sisa, L. (2002). Age differences in perceived workload across a short vigil. *Ergonomics*, 45(13):949–960.

Calzavara, M., Battini, D., Bogataj, D., Sgarbossa, F., and Zennaro, I. (2020). Ageing workforce management in manufacturing systems: state of the art and future research agenda. *International Journal of Production Research*, 58(3):729–747.

Carrieri, M., Petracca, A., Lancia, S., Basso Moro, S., Brigadoi, S., Spezialetti, M., Ferrari, M., Placidi, G., and Quaresima, V. (2016). Prefrontal cortex activation upon a demanding virtual hand-controlled task: a new frontier for neuroergonomics. *Frontiers in human neuroscience*, 10:53.

Casadio, M., Ranganathan, R., and Mussa-Ivaldi, F. A. (2012). The body-machine interface: a new perspective on an old theme. *Journal of Motor behavior*, 44(6):419–433.

Chacón, A., Ponsa, P., and Angulo, C. (2021). Usability study through a human-robot collaborative workspace experience. *Designs*, 5(2):35.

Chan, A. H., Hoffmann, E. R., and Ho, J. C. (2019). Movement time and guidance accuracy in teleoperation of robotic vehicles. *Ergonomics*, 62(5):706–720.

Chatrian, G. E., Lettich, E., and Nelson, P. L. (1985). Ten percent electrode system for topographic studies of spontaneous and evoked eeg activities. *American Journal of EEG technology*, 25(2):83–92.

Chen, J. and Or, C. (2017). Assessing the use of immersive virtual reality, mouse and touchscreen in pointing and dragging-and-dropping tasks among young, middle-aged and older adults. *Applied ergonomics*, 65:437–448.

Chuan, T. K., Shin, C. N., Foo, D., Chao, W. Y., and Wenhao, Y. (2007). Modeling robot movement in a virtual environment. *IFAC Proceedings Volumes*, 40(16):177–182.

Claudon, L., Desbrosses, K., Gilles, M., Pichené-Houard, A., Remy, O., and Wild, P. (2020). Temporal leeway: can it help to reduce biomechanical load for older workers performing repetitive light assembly tasks? *Applied Ergonomics*, 86:103081.

Coats, R. O., Fath, A. J., Astill, S. L., and Wann, J. P. (2016). Eye and hand movement strategies in older adults during a complex reaching task. *Experimental brain research*, 234(2):533–547.

Coello, Y. and Fischer, M. H. (2015). *Perceptual and Emotional Embodiment: Foundations of Embodied Cognition Volume 1*, volume 1. Routledge.

Damiani, L., Demartini, M., Guizzi, G., Revetria, R., and Tonelli, F. (2018). Augmented and virtual reality applications in industrial systems: A qualitative review towards the industry 4.0 era. *IFAC-PapersOnLine*, 51(11):624–630.

Darley, W. K. and Smith, R. E. (1995). Gender differences in information processing strategies: An empirical test of the selectivity model in advertising response. *Journal of advertising*, 24(1):41–56.

De Vignemont, F. (2011). Embodiment, ownership and disownership. *Consciousness and cognition*, 20(1):82–93.

De Waard, D. and Brookhuis, K. (1996). The measurement of drivers' mental workload.

Dehais, F., Karwowski, W., and Ayaz, H. (2020a). Brain at work and in everyday life as the next frontier: grand field challenges for neuroergonomics.

Dehais, F., Lafont, A., Roy, R., and Fairclough, S. (2020b). A neuroergonomics approach to mental workload, engagement and human performance. *Frontiers in neuroscience*, 14:268.

Delignette-Muller, M. L. and Dutang, C. (2015). fitdistrplus: An r package for fitting distributions. *Journal of statistical software*, 64:1–34.

Delorme, A. and Makeig, S. (2004). Eeglab: an open source toolbox for analysis of single-trial eeg dynamics including independent component analysis. *Journal of neuroscience methods*, 134(1):9–21.

Di Pasquale, V., Miranda, S., and Neumann, W. P. (2020). Ageing and human-system errors in manufacturing: a scoping review. *International Journal of Production Research*, 58(15):4716–4740.

Ding, L., He, J., Yao, L., Zhuang, J., Chen, S., Wang, H., Jiang, N., and Jia, J. (2020). Mirror visual feedback combining vibrotactile stimulation promotes embodiment perception: an electroencephalogram (eeg) pilot study. *Frontiers in Bioengineering and Biotechnology*, 8:553270.

Djebbara, Z., Fich, L. B., Petrini, L., and Gramann, K. (2019). Sensorimotor brain dynamics reflect architectural affordances. *Proceedings of the National Academy of Sciences*, 116(29):14769–14778.

Donato, A. J., Tench, K., Glueck, D. H., Seals, D. R., Eskurza, I., and Tanaka, H. (2003). Declines in physiological functional capacity with age: a longitudinal study in peak swimming performance. *Journal of applied physiology*, 94(2):764–769.

Doolani, S., Wessels, C., Kanal, V., Sevastopoulos, C., Jaiswal, A., Nambiappan, H., and Makedon, F. (2020). A review of extended reality (xr) technologies for manufacturing training. *Technologies*, 8(4):77.

Du, G., Zhang, L., Su, K., Wang, X., Teng, S., and Liu, P. X. (2022). A multimodal fusion fatigue driving detection method based on heart rate and perclos. *IEEE Transactions on Intelligent Transportation Systems*.

Du, J., Do, H. M., and Sheng, W. (2021). Human–robot collaborative control in a virtual-reality-based telepresence system. *International Journal of Social Robotics*, 13(6):1295–1306.

Dybvik, H., Løland, M., Gerstenberg, A., Slåttsveen, K. B., and Steinert, M. (2021). A low-cost predictive display for teleoperation: Investigating effects on human performance and workload. *International Journal of Human-Computer Studies*, 145:102536.

Dye, M. W., Green, C. S., and Bavelier, D. (2009). Increasing speed of processing with action video games. *Current directions in psychological science*, 18(6):321–326.

D'Angelo, M., di Pellegrino, G., Seriani, S., Gallina, P., and Frassinetti, F. (2018). The sense of agency shapes body schema and peripersonal space. *Scientific reports*, 8(1):1–11.

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In *Proceedings of the Human Factors Society annual meeting*, volume 32, pages 97–101. Sage Publications Sage CA: Los Angeles, CA.

Endsley, M. R., Bolté, B., and Jones, D. G. (2003). *Designing for situation awareness: An approach to user-centered design*. CRC press.

Erdfelder, E., Faul, F., and Buchner, A. (1996). Gpower: A general power analysis program. *Behavior research methods, instruments, & computers*, 28(1):1–11.

Evans, N. and Blanke, O. (2013). Shared electrophysiology mechanisms of body ownership and motor imagery. *Neuroimage*, 64:216–228.

Faccio, M., Granata, I., Menini, A., Milanese, M., Rossato, C., Bottin, M., Minto, R., Pluchino, P., Gamberini, L., Boschetti, G., et al. (2022). Human factors in cobot era: a review of modern production systems features. *Journal of Intelligent Manufacturing*, pages 1–22.

Farizon, D., Dominey, P., and Ventre-Dominey, J. (2021). Insights on embodiment induced by visuo-tactile stimulation during robotic telepresence. *Scientific reports*, 11(1):1–12.

Faure, V., Lobjois, R., and Benguigui, N. (2016). The effects of driving environment complexity and dual tasking on drivers' mental workload and eye blink behavior. *Transportation research part F: traffic psychology and behaviour*, 40:78–90.

Firu, A. C., Tapîrdea, A. I., Feier, A. I., and Drăghici, G. (2021). Virtual reality in the automotive field in industry 4.0. *Materials Today: Proceedings*, 45:4177–4182.

Fogarty, C. and Stern, J. A. (1989). Eye movements and blinks: their relationship to higher cognitive processes. *International journal of psychophysiology*, 8(1):35–42.

Franzluebbers, A. and Johnson, K. (2019). Remote robotic arm teleoperation through virtual reality. In *Symposium on Spatial User Interaction*, pages 1–2.

Fratczak, P., Goh, Y. M., Kinnell, P., Soltoggio, A., and Justham, L. (2019). Understanding human behaviour in industrial human-robot interaction by means of virtual reality. In *Proceedings of the Halfway to the Future Symposium 2019*, pages 1–7.

Friard, O. and Gamba, M. (2016). Boris: a free, versatile open-source event-logging software for video/audio coding and live observations. *Methods in ecology and evolution*, 7(11):1325–1330.

Fuhl, W., Tonsen, M., Bulling, A., and Kasneci, E. (2016). Pupil detection for head-mounted eye tracking in the wild: an evaluation of the state of the art. *Machine Vision and Applications*, 27(8):1275–1288.

Gallagher, S. (2006). *How the body shapes the mind*. Clarendon Press.

Gao, Q., Wang, Y., Song, F., Li, Z., and Dong, X. (2013). Mental workload measurement for emergency operating procedures in digital nuclear power plants. *Ergonomics*, 56(7):1070–1085.

Gehrke, L., Akman, S., Lopes, P., Chen, A., Singh, A. K., Chen, H.-T., Lin, C.-T., and Gramann, K. (2019). Detecting visuo-haptic mismatches in virtual reality using the prediction error negativity of event-related brain potentials. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–11.

Gehrke, L., Iversen, J. R., Makeig, S., and Gramann, K. (2018). The invisible maze task (imt): interactive exploration of sparse virtual environments to investigate action-driven formation of spatial representations. In *German Conference on Spatial Cognition*, pages 293–310. Springer.

Gerling, K. M., Dergousoff, K. K., Mandryk, R. L., et al. (2013). Is movement better? comparing sedentary and motion-based game controls for older adults. In *Proceedings-Graphics Interface*, pages 133–140. Canadian Information Processing Society.

Gilles, M. A., Guélin, J.-C., Desbrosses, K., and Wild, P. (2017). Motor adaptation capacity as a function of age in carrying out a repetitive assembly task at imposed work paces. *Applied ergonomics*, 64:47–55.

Gliesche, P., Krick, T., Pfingsthorn, M., Drolshagen, S., Kowalski, C., and Hein, A. (2020). Kinesthetic device vs. keyboard/mouse: A comparison in home care telemanipulation. *Frontiers in Robotics and AI*, 7:561015.

Gomer, J. A. and Pagano, C. C. (2011). Spatial perception and robot operation: Should spatial abilities be considered when selecting robot operators? In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, volume 55, pages 1260–1264. SAGE Publications Sage CA: Los Angeles, CA.

González-Franco, M., Peck, T. C., Rodríguez-Fornells, A., and Slater, M. (2014). A threat to a virtual hand elicits motor cortex activation. *Experimental brain research*, 232(3):875–887.

Goodrich, M. A., Schultz, A. C., et al. (2008). Human–robot interaction: a survey. *Foundations and Trends® in Human–Computer Interaction*, 1(3):203–275.

Gorisse, G., Christmann, O., Amato, E. A., and Richir, S. (2017). First-and third-person perspectives in immersive virtual environments: presence and performance analysis of embodied users. *Frontiers in Robotics and AI*, 4:33.

Grabowski, A., Jankowski, J., and Wodzyński, M. (2021). Teleoperated mobile robot with two arms: the influence of a human-machine interface, vr training and operator age. *International Journal of Human-Computer Studies*, 156:102707.

Gramann, K., Fairclough, S. H., Zander, T. O., and Ayaz, H. (2017). Trends in neuroergonomics. *Frontiers in human neuroscience*, page 165.

Gramann, K., Ferris, D. P., Gwin, J., and Makeig, S. (2014). Imaging natural cognition in action. *International Journal of Psychophysiology*, 91(1):22–29.

Gramann, K., Gwin, J. T., Bigdely-Shamlo, N., Ferris, D. P., and Makeig, S. (2010). Visual evoked responses during standing and walking. *Frontiers in human neuroscience*, 4:202.

Gramann, K., Gwin, J. T., Ferris, D. P., Oie, K., Jung, T.-P., Lin, C.-T., Liao, L.-D., and Makeig, S. (2011). Cognition in action: imaging brain/body dynamics in mobile humans.

Grandi, F., Zanni, L., Peruzzini, M., Pellicciari, M., and Campanella, C. E. (2020). A transdisciplinary digital approach for tractor's human-centred design. *International Journal of Computer Integrated Manufacturing*, 33(4):377–397.

Greef, T. d., Lafeber, H., Oostendorp, H. v., and Lindenberg, J. (2009). Eye movement as indicators of mental workload to trigger adaptive automation. In *International Conference on Foundations of Augmented Cognition*, pages 219–228. Springer.

Green, C. S. and Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, 423(6939):534–537.

Gregersen, A. and Grodal, T. (2008). Embodiment and interface. In *The video game theory reader 2*, pages 87–106. Routledge.

Guo, Y., Freer, D., Deligianni, F., and Yang, G.-Z. (2021). Eye-tracking for performance evaluation and workload estimation in space telerobotic training. *IEEE Transactions on Human-Machine Systems*, 52(1):1–11.

Habibnezhad, M., Puckett, J., Jebelli, H., Karji, A., Fardhosseini, M. S., and Asadi, S. (2020). Neurophysiological testing for assessing construction workers' task performance at virtual height. *Automation in Construction*, 113:103143.

Hancock, P. A., Billings, D. R., Schaefer, K. E., Chen, J. Y., De Visser, E. J., and Parasuraman, R. (2011). A meta-analysis of factors affecting trust in human-robot interaction. *Human factors*, 53(5):517–527.

Hansen, L. I. N., Vinther, N., Stranovsky, L., Philipsen, M. P., Wu, H., and Moeslund, T. B. (2018). Collaborative meat processing in virtual reality: Evaluating perceived safety and predictability of robot approach. In *International Conference on Human Robot Interaction (HRI 2018)*. VAM-HRI.

Hart, S. G. (2006). Nasa-task load index (nasa-tlx); 20 years later. In *Proceedings of the human factors and ergonomics society annual meeting*, volume 50, pages 904–908. Sage publications Sage CA: Los Angeles, CA.

Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological psychology*, 45(1-3):73–93.

Hoffmann, S., Borges, U., Bröker, L., Laborde, S., Liepelt, R., Lobinger, B. H., Löffler, J., Musculus, L., and Raab, M. (2018). The psychophysiology of action: A multidisciplinary endeavor for integrating action and cognition. *Frontiers in Psychology*, 9:1423.

Holland, M. K. and Tarlow, G. (1972). Blinking and mental load. *Psychological Reports*, 31(1):119–127.

Hormaza, L. A., Mohammed, W. M., Ferrer, B. R., Bejarano, R., and Lastra, J. L. M. (2019). On-line training and monitoring of robot tasks through virtual reality. In *2019 IEEE 17th International Conference on Industrial Informatics (INDIN)*, volume 1, pages 841–846. IEEE.

Hsieh, S.-Y. and Lu, J.-M. (2018). Feasibility evaluation for immersive virtual reality simulation of human-machine collaboration: A case study of handover tasks. In *Congress of the International Ergonomics Association*, pages 364–369. Springer.

Hu, C., Meng, M. Q., Liu, P. X., and Wang, X. (2003). Visual gesture recognition for human-machine interface of robot teleoperation. In *Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)(Cat. No. 03CH37453)*, volume 2, pages 1560–1565. IEEE.

Hughes, A. M., Hancock, G. M., Marlow, S. L., Stowers, K., and Salas, E. (2019). Cardiac measures of cognitive workload: a meta-analysis. *Human factors*, 61(3):393–414.

Ijaz, K., Ahmadpour, N., Naismith, S. L., Calvo, R. A., et al. (2019). An immersive virtual reality platform for assessing spatial navigation memory in predementia screening: feasibility and usability study. *JMIR mental health*, 6(9):e13887.

Ilmarinen, J. E. (2001). Aging workers. *Occupational and environmental medicine*, 58(8):546–546.

Iqbal, S. T., Zheng, X. S., and Bailey, B. P. (2004). Task-evoked pupillary response to mental workload in human-computer interaction. In *CHI'04 extended abstracts on Human factors in computing systems*, pages 1477–1480.

ISO (2019). Ergonomics of human-system interaction—part 210: Human-centred design for interactive systems.

Iwasaki, Y., Navarro, B., Iwata, H., and Ganesh, G. (2022). Embodiment modifies attention allotment for the benefit of dual task performance. *Communications biology*, 5(1):1–7.

James, C. A., Bednarz, T. P., Haustein, K., Alem, L., Caris, C., and Castleden, A. (2011). Tele-operation of a mobile mining robot using a panoramic display: an exploration of operators sense of presence. In *2011 IEEE International Conference on Automation Science and Engineering*, pages 279–284. IEEE.

Jungnickel, E., Gehrke, L., Klug, M., and Gramann, K. (2019). Mobi—mobile brain/body imaging. *Neuroergonomics*, pages 59–63.

Kahneman, D. (1973). *Attention and effort*, volume 1063. Citeseer.

Kammers, M. P., de Vignemont, F., Verhagen, L., and Dijkerman, H. C. (2009). The rubber hand illusion in action. *Neuropsychologia*, 47(1):204–211.

Kanayama, N., Sato, A., and Ohira, H. (2007). Crossmodal effect with rubber hand illusion and gamma-band activity. *Psychophysiology*, 44(3):392–402.

Kaufeld, M. and Nickel, P. (2019). Level of robot autonomy and information aids in human-robot interaction affect human mental workload–an investigation in virtual reality. In *International Conference on Human-Computer Interaction*, pages 278–291. Springer.

Kenny, G. P., Yardley, J. E., Martineau, L., and Jay, O. (2008). Physical work capacity in older adults: implications for the aging worker. *American journal of industrial medicine*, 51(8):610–625.

Keogh, E. J. and Pazzani, M. J. (2001). Derivative dynamic time warping. In *Proceedings of the 2001 SIAM international conference on data mining*, pages 1–11. SIAM.

Ketcham, C. J., Seidler, R. D., Van Gemmert, A. W., and Stelmach, G. E. (2002). Age-related kinematic differences as influenced by task difficulty, target size, and movement amplitude. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, 57(1):P54–P64.

Kilby, C. and Whitehead, A. (2017). A study of viewpoint and feedback in wearable systems for controlling a robot arm. In *International Conference on Applied Human Factors and Ergonomics*, pages 136–148. Springer.

Kilteni, K., Groten, R., and Slater, M. (2012a). The sense of embodiment in virtual reality. *Presence: Teleoperators and Virtual Environments*, 21(4):373–387.

Kilteni, K., Normand, J.-M., Sanchez-Vives, M. V., and Slater, M. (2012b). Extending body space in immersive virtual reality: a very long arm illusion. *PloS one*, 7(7):e40867.

Kim, C.-S., Jung, M., Kim, S.-Y., Kim, K., et al. (2020). Controlling the sense of embodiment for virtual avatar applications: methods and empirical study. *JMIR Serious Games*, 8(3):e21879.

Kim, S. Y., Park, H., Kim, H., Kim, J., and Seo, K. (2022). Technostress causes cognitive overload in high-stress people: Eye tracking analysis in a virtual kiosk test. *Information Processing & Management*, 59(6):103093.

Klug, M., Jeung, S., Wunderlich, A., Gehrke, L., Protzak, J., Djebbara, Z., Argubi-Wollesen, A., Wollesen, B., and Gramann, K. (2022). The bemobil pipeline for automated analyses of multimodal mobile brain and body imaging data. *bioRxiv*.

Kokkinara, E. and Slater, M. (2014). Measuring the effects through time of the influence of visuomotor and visuotactile synchronous stimulation on a virtual body ownership illusion. *Perception*, 43(1):43–58.

Körber, M., Cingel, A., Zimmermann, M., and Bengler, K. (2015). Vigilance decrement and passive fatigue caused by monotony in automated driving. *Procedia Manufacturing*, 3:2403–2409.

Kothe, C. et al. (2014). Lab streaming layer (lsl).

Kramer, A. F. (2020). Physiological metrics of mental workload: A review of recent progress. *Multiple-task performance*, pages 279–328.

Kret, M. E. and Sjak-Shie, E. E. (2019). Preprocessing pupil size data: Guidelines and code. *Behavior research methods*, 51(3):1336–1342.

Krüger, J., Lien, T. K., and Verl, A. (2009). Cooperation of human and machines in assembly lines. *CIRP annals*, 58(2):628–646.

Ladouce, S., Donaldson, D. I., Dudchenko, P. A., and Ietswaart, M. (2017). Understanding minds in real-world environments: toward a mobile cognition approach. *Frontiers in human neuroscience*, 10:694.

Lagomarsino, M., Lorenzini, M., De Momi, E., and Ajoudani, A. (2022). An online framework for cognitive load assessment in industrial tasks. *Robotics and Computer-Integrated Manufacturing*, 78:102380.

Lamers, M. H., Verbeek, F. J., and van der Putten, P. W. (2013). Tinkering in scientific education. In *International Conference on Advances in Computer Entertainment Technology*, pages 568–571. Springer.

Lankton, N. K., McKnight, D. H., and Tripp, J. (2015). Technology, humanness, and trust: Rethinking trust in technology. *Journal of the Association for Information Systems*, 16(10):1.

Lee, J. D. and See, K. A. (2004). Trust in automation: Designing for appropriate reliance. *Human factors*, 46(1):50–80.

Lee, K. M. (2004). Presence, explicated. *Communication theory*, 14(1):27–50.

Leone, C., Feys, P., Moumdjian, L., D'Amico, E., Zappia, M., and Patti, F. (2017). Cognitive-motor dual-task interference: a systematic review of neural correlates. *Neuroscience & Biobehavioral Reviews*, 75:348–360.

Linn, C., Bender, S., Prosser, J., Schmitt, K., and Werth, D. (2017). Virtual remote inspection—a new concept for virtual reality enhanced real-time maintenance. In *2017 23rd International Conference on Virtual System & Multimedia (VSMM)*, pages 1–6. IEEE.

Lipton, J. I., Fay, A. J., and Rus, D. (2017). Baxter's homunculus: Virtual reality spaces for teleoperation in manufacturing. *IEEE Robotics and Automation Letters*, 3(1):179–186.

Liu, H. and Wang, L. (2020). Remote human–robot collaboration: A cyber–physical system application for hazard manufacturing environment. *Journal of manufacturing systems*, 54:24–34.

Longo, L., Wickens, C. D., Hancock, G., and Hancock, P. (2022). Human mental workload: A survey and a novel inclusive definition. *Frontiers in Psychology*, 13:883321.

Lu, Y., Zheng, H., Chand, S., Xia, W., Liu, Z., Xu, X., Wang, L., Qin, Z., and Bao, J. (2022). Outlook on human-centric manufacturing towards industry 5.0. *Journal of Manufacturing Systems*, 62:612–627.

Ma, R. and Kaber, D. B. (2006). Presence, workload and performance effects of synthetic environment design factors. *International Journal of Human-Computer Studies*, 64(6):541–552.

Macchini, M., Lortkipanidze, M., Schiano, F., and Floreano, D. (2021a). The impact of virtual reality and viewpoints in body motion based drone teleoperation. In *2021 IEEE Virtual Reality and 3D User Interfaces (VR)*, pages 511–518. IEEE.

Macchini, M., Schiano, F., and Floreano, D. (2021b). Data-driven personalization of body-machine interfaces to control diverse robot types.

Makeig, S., Gramann, K., Jung, T.-P., Sejnowski, T. J., and Poizner, H. (2009). Linking brain, mind and behavior. *International Journal of Psychophysiology*, 73(2):95–100.

Maneuvrier, A., Decker, L. M., Ceyte, H., Fleury, P., and Renaud, P. (2020). Presence promotes performance on a virtual spatial cognition task: impact of human factors on virtual reality assessment. *Frontiers in Virtual Reality*, 1:571713.

Marinescu, A. C., Sharples, S., Ritchie, A. C., Sanchez Lopez, T., McDowell, M., and Morvan, H. P. (2018). Physiological parameter response to variation of mental workload. *Human factors*, 60(1):31–56.

Marquart, G., Cabrall, C., and de Winter, J. (2015). Review of eye-related measures of drivers' mental workload. *Procedia Manufacturing*, 3:2854–2861.

Martín-Barrio, A., Roldán, J. J., Terrile, S., del Cerro, J., and Barrientos, A. (2020). Application of immersive technologies and natural language to hyper-redundant robot teleoperation. *Virtual Reality*, 24(3):541–555.

Mathôt, S. (2018). Pupillometry: Psychology, physiology, and function. *Journal of Cognition*, 1(1).

Matsas, E. and Vosniakos, G.-C. (2017). Design of a virtual reality training system for human–robot collaboration in manufacturing tasks. *International Journal on Interactive Design and Manufacturing (IJIDeM)*, 11(2):139–153.

Matsas, E., Vosniakos, G.-C., and Batras, D. (2018). Prototyping proactive and adaptive techniques for human-robot collaboration in manufacturing using virtual reality. *Robotics and Computer-Integrated Manufacturing*, 50:168–180.

Matthews, G., Joyner, L., Gilliland, K., Campbell, S., Falconer, S., and Huggins, J. (1999). Validation of a comprehensive stress state questionnaire: Towards a state big three. *Personality psychology in Europe*, 7:335–350.

Matthews, G., Reinerman-Jones, L. E., Barber, D. J., and Abich IV, J. (2015). The psychometrics of mental workload: Multiple measures are sensitive but divergent. *Human factors*, 57(1):125–143.

Mavridis, N., Pierris, G., Gallina, P., Moustakas, N., and Astaras, A. (2015). Subjective difficulty and indicators of performance of joystick-based robot arm teleoperation with auditory feedback. In *2015 International Conference on Advanced Robotics (ICAR)*, pages 91–98. IEEE.

McIntire, L. K., McKinley, R. A., Goodyear, C., and McIntire, J. P. (2014). Detection of vigilance performance using eye blinks. *Applied ergonomics*, 45(2):354–362.

Melluso, N., Fareri, S., Fantoni, G., Bonaccorsi, A., Chiarello, F., Coli, E., Giordano, V., Manfredi, P., and Manafi, S. (2020). Lights and shadows of covid-19, technology and industry 4.0. *arXiv preprint arXiv:2004.13457*.

Memar, A. H. and Esfahani, E. T. (2018). Eeg correlates of motor control difficulty in physical human-robot interaction: A frequency domain analysis. In *2018 IEEE Haptics Symposium (HAPTICS)*, pages 229–234. IEEE.

Menchaca Brandan, M. A. (2007). *Influence of spatial orientation and spatial visualization abilities on space teleoperation performance*. PhD thesis, Massachusetts Institute of Technology.

Mingardi, M., Pluchino, P., Bacchin, D., Rossato, C., and Gamberini, L. (2020). Assessment of implicit and explicit measures of mental workload in working situations: implications for industry 4.0. *Applied Sciences*, 10(18):6416.

Minsky, M. (1980). Telepresence.

Miura, R., Kasahara, S., Kitazaki, M., Verhulst, A., Inami, M., and Sugimoto, M. (2021). Multisoma: Distributed embodiment with synchronized behavior and perception. In *Augmented Humans Conference 2021*, pages 1–9.

Munzert, J., Lorey, B., and Zentgraf, K. (2009). Cognitive motor processes: the role of motor imagery in the study of motor representations. *Brain research reviews*, 60(2):306–326.

Mutlu, B., Osman, S., Forlizzi, J., Hodgins, J., and Kiesler, S. (2006). Task structure and user attributes as elements of human-robot interaction design. In *ROMAN 2006-The 15th IEEE International Symposium on Robot and Human Interactive Communication*, pages 74–79. IEEE.

Mystakidis, S. (2022). Metaverse. *Encyclopedia*, 2(1):486–497.

Nachreiner, F., Nickel, P., and Meyer, I. (2006). Human factors in process control systems: The design of human–machine interfaces. *Safety Science*, 44(1):5–26.

Navon, D. and Miller, J. (1987). Role of outcome conflict in dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance*, 13(3):435.

Nenna, F., Do, C. T., Protzak, J., and Gramann, K. (2021). Alteration of brain dynamics during dual-task overground walking. *European Journal of Neuroscience*, 54(12):8158–8174.

Nenna, F. and Gamberini, L. (2022). The influence of gaming experience, gender and other individual factors on robot teleoperations in vr. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction*, pages 945–949.

Nenna, F., Orso, V., Zanardi, D., and Gamberini, L. (2022a). The virtualization of human–robot interactions: a user-centric workload assessment. *Virtual Reality*, pages 1–19.

Nenna, F., Zanardi, D., and Gamberini, L. (2022b). Human-centric telerobotics: investigating users' performance and workload via vr-based eye-tracking measures. *arXiv preprint arXiv:2212.07345*.

Newport, R., Pearce, R., and Preston, C. (2010). Fake hands in action: embodiment and control of supernumerary limbs. *Experimental brain research*, 204(3):385–395.

Newport, R. and Preston, C. (2011). Disownership and disembodiment of the real limb without visuoproprioceptive mismatch. *Cognitive Neuroscience*, 2(3-4):179–185.

Nguyen Ngoc, H., Lasa, G., and Iriarte, I. (2022). Human-centred design in industry 4.0: case study review and opportunities for future research. *Journal of Intelligent Manufacturing*, 33(1):35–76.

Niemeyer, G., Preusche, C., Stramigioli, S., and Lee, D. (2016). Telerobotics. In *Springer handbook of robotics*, pages 1085–1108. Springer.

Nomura, T. and Takagi, S. (2011). Exploring effects of educational backgrounds and gender in human-robot interaction. In *2011 International conference on user science and engineering (i-user)*, pages 24–29. IEEE.

Novak, D., Beyeler, B., Omlin, X., and Riener, R. (2015). Workload estimation in physical human–robot interaction using physiological measurements. *Interacting with computers*, 27(6):616–629.

O'DONNELL, R. D. (1986). Workload assessment methodology. *Cognitive processes and performance*.

Opiyo, S., Zhou, J., Mwangi, E., Kai, W., and Sunusi, I. (2021). A review on teleoperation of mobile ground robots: Architecture and situation awareness. *International Journal of Control, Automation and Systems*, 19(3):1384–1407.

Palmer, J. A., Kreutz-Delgado, K., and Makeig, S. (2012). Amica: An adaptive mixture of independent component analyzers with shared components. *Swartz Center for Computatonal Neursoscience, University of California San Diego, Tech. Rep.*

Pang, G., Yang, G., and Pang, Z. (2021). Review of robot skin: A potential enabler for safe collaboration, immersive teleoperation, and affective interaction of future collaborative robots. *IEEE Transactions on Medical Robotics and Bionics*.

Paperno, N., Rupp, M. A., Parkhurst, E. L., Maboudou-Tchao, E. M., Smither, J. A.-A., Bricout, J., and Behal, A. (2019). Age and gender differences in performance for operating a robotic manipulator. *IEEE Transactions on Human-Machine Systems*, 49(2):137–149.

Parasuraman, R. and Rizzo, M. (2006). *Neuroergonomics: The brain at work*, volume 3. Oxford University Press.

Parasuraman, R. and Wilson, G. F. (2008). Putting the brain to work: Neuroergonomics past, present, and future. *Human factors*, 50(3):468–474.

Pashler, H. (1984). Processing stages in overlapping tasks: evidence for a central bottleneck. *Journal of Experimental Psychology: Human perception and performance*, 10(3):358.

Pavani, F., Spence, C., and Driver, J. (2000). Visual capture of touch: Out-of-the-body experiences with rubber gloves. *Psychological science*, 11(5):353–359.

Pavon-Pulido, N., Lopez-Riquelme, J. A., Pinuaga-Cascales, J. J., Ferruz-Melero, J., and Dos Santos, R. M. (2015). Cybi: A smart companion robot for elderly people: Improving teleoperation and telepresence skills by combining cloud computing technologies and fuzzy logic. In *2015 IEEE International Conference on Autonomous Robot Systems and Competitions*, pages 198–203. IEEE.

Peeters, M. C. and van Emmerik, H. (2008). An introduction to the work and well-being of older workers: From managing threats to creating opportunities. *Journal of managerial psychology*.

Peled, A., Pressman, A., Geva, A. B., and Modai, I. (2003). Somatosensory evoked potentials during a rubber-hand illusion in schizophrenia. *Schizophrenia research*, 64(2-3):157–163.

Pérez, L., Diez, E., Usamentiaga, R., and García, D. F. (2019). Industrial robot control and operator training using virtual reality interfaces. *Computers in Industry*, 109:114–120.

Petkova, V. I., Khoshnevis, M., and Ehrsson, H. H. (2011). The perspective matters! multisensory integration in ego-centric reference frames determines full-body ownership. *Frontiers in psychology*, 2:35.

Pion-Tonachini, L., Kreutz-Delgado, K., and Makeig, S. (2019). Iclabel: An automated electroencephalographic independent component classifier, dataset, and website. *NeuroImage*, 198:181–197.

Pomplun, M. and Sunkara, S. (2019). Pupil dilation as an indicator of cognitive workload in human-computer interaction. In *Human-Centered Computing*, pages 542–546. CRC Press.

Pratticò, F. G. and Lamberti, F. (2021). Towards the adoption of virtual reality training systems for the self-tuition of industrial robot operators: A case study at kuka. *Computers in Industry*, 129:103446.

Press, C., Heyes, C., Haggard, P., and Eimer, M. (2008). Visuotactile learning and body representation: an erp study with rubber hands and rubber objects. *Journal of cognitive neuroscience*, 20(2):312–323.

Preston, C. and Newport, R. (2011). Differential effects of perceived hand location on the disruption of embodiment by apparent physical encroachment of the limb. *Cognitive Neuroscience*, 2(3-4):163–170.

Pritchard, S. C., Zopf, R., Polito, V., Kaplan, D. M., and Williams, M. A. (2016). Non-hierarchical influence of visual form, touch, and position cues on embodiment, agency, and presence in virtual reality. *Frontiers in psychology*, 7:1649.

Qin, J., Lin, J.-H., Faber, G. S., Buchholz, B., and Xu, X. (2014). Upper extremity kinematic and kinetic adaptations during a fatiguing repetitive task. *Journal of Electromyography and Kinesiology*, 24(3):404–411.

Ra, S., Shrestha, U., Khatiwada, S., Yoon, S. W., and Kwon, K. (2019). The rise of technology and impact on skills. *International Journal of Training Research*, 17(sup1):26–40.

Ratcliffe, N. and Newport, R. (2017). The effect of visual, spatial and temporal manipulations on embodiment and action. *Frontiers in human neuroscience*, 11:227.

Richard, G., Pietrzak, T., Argelaguet, F., Lécuyer, A., and Casiez, G. (2021). Studying the role of haptic feedback on virtual embodiment in a drawing task. *Frontiers in Virtual Reality*, 1:573167.

Rosen, E., Whitney, D., Phillips, E., Ullman, D., and Tellex, S. (2018). Testing robot teleoperation using a virtual reality interface with ros reality. In *Proceedings of the 1st International Workshop on Virtual, Augmented, and Mixed Reality for HRI (VAM-HRI)*, pages 1–4.

Rossato, C., Pluchino, P., Cellini, N., Jacucci, G., Spagnolli, A., and Gamberini, L. (2021). Facing with collaborative robots: the subjective experience in senior and younger workers. *Cyberpsychology, Behavior, and Social Networking*, 24(5):349–356.

Rouanet, P., Béchu, J., and Oudeyer, P.-Y. (2009). A comparison of three interfaces using handheld devices to intuitively drive and show objects to a social robot: the impact of underlying metaphors. In *RO-MAN 2009-The 18th IEEE International Symposium on Robot and Human Interactive Communication*, pages 1066–1072. IEEE.

Sanchez-Vives, M. V., Spanlang, B., Frisoli, A., Bergamasco, M., and Slater, M. (2010). Virtual hand illusion induced by visuomotor correlations. *PloS one*, 5(4):e10381.

Savur, C., Kumar, S., and Sahin, F. (2019). A framework for monitoring human physiological response during human robot collaborative task. In *2019 IEEE international conference on systems, man and cybernetics (SMC)*, pages 385–390. IEEE.

Schettler, A., Raja, V., and Anderson, M. L. (2019). The embodiment of objects: Review, analysis, and future directions. *Frontiers in Neuroscience*, 13:1332.

Seinfeld, S., Schmidt, I., and Müller, J. (2022). Evoking realistic affective touch experiences in virtual reality. *arXiv preprint arXiv:2202.13389*.

Shapiro, L. (2010). *Embodied cognition*. Routledge.

Shirwalkar, S., Singh, A., Sharma, K., and Singh, N. (2013). Telemanipulation of an industrial robotic arm using gesture recognition with kinect. In *2013 International Conference on Control, Automation, Robotics and Embedded Systems (CARE)*, pages 1–6. IEEE.

Showkat, D. and Grimm, C. (2018). Identifying gender differences in information processing style, self-efficacy, and tinkering for robot teleoperation. In *2018 15th international conference on ubiquitous robots (UR)*, pages 443–448. IEEE.

Singh, A. K., Chen, H.-T., Cheng, Y.-F., King, J.-T., Ko, L.-W., Gramann, K., and Lin, C.-T. (2018). Visual appearance modulates prediction error in virtual reality. *IEEE Access*, 6:24617–24624.

Slater, M., Spanlang, B., Sanchez-Vives, M. V., and Blanke, O. (2010). First person experience of body transfer in virtual reality. *PloS one*, 5(5):e10564.

Slater, M. and Steed, A. (2000). A virtual presence counter. *Presence*, 9(5):413–434.

Souchet, A. D., Philippe, S., Lourdeaux, D., and Leroy, L. (2022). Measuring visual fatigue and cognitive load via eye tracking while learning with virtual reality head-mounted displays: A review. *International Journal of Human–Computer Interaction*, 38(9):801–824.

Statista (2019). Median age of the global labor force by region and gender 2019.

Steinfeld, A., Fong, T., Kaber, D., Lewis, M., Scholtz, J., Schultz, A., and Goodrich, M. (2006). Common metrics for human-robot interaction. In *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, pages 33–40.

Sweeney, J. A., Rosano, C., Berman, R. A., and Luna, B. (2001). Inhibitory control of attention declines more than working memory during normal aging. *Neurobiology of aging*, 22(1):39–47.

Sweller, J., Ayres, P., and Kalyuga, S. (2011). Measuring cognitive load. In *Cognitive load theory*, pages 71–85. Springer.

Syed-Abdul, S., Malwade, S., Nursetyo, A. A., Sood, M., Bhatia, M., Barsasella, D., Liu, M. F., Chang, C.-C., Srinivasan, K., Li, Y.-C. J., et al. (2019). Virtual reality among the elderly: a usefulness and acceptance study from taiwan. *BMC geriatrics*, 19(1):1–10.

Szczurek, K. A., Prades, R. M., Matheson, E., Rodriguez-Nogueira, J., and Di Castro, M. (2022). Mixed reality human–robot interface with adaptive communications congestion control for the teleoperation of mobile redundant manipulators in hazardous environments. *IEEE Access*, 10:87182–87216.

Team, R. (2021). Rstudio: integrated development for r. rstudio, pbc, boston, ma. 2020.

Teimourikia, M., Saidinejad, H., Comai, S., and Salice, F. (2014). Personalized hand pose and gesture recognition system for the elderly. In *International Conference on Universal Access in Human-Computer Interaction*, pages 191–202. Springer.

Toet, A., Kuling, I. A., Krom, B. N., and Van Erp, J. B. (2020). Toward enhanced teleoperation through embodiment. *Frontiers in Robotics and AI*, 7:14.

Tsakiris, M., Prabhu, G., and Haggard, P. (2006). Having a body versus moving your body: How agency structures body-ownership. *Consciousness and cognition*, 15(2):423–432.

Van Acker, B. B., Bombeke, K., Durnez, W., Parmentier, D. D., Mateus, J. C., Biondi, A., Saldien, J., and Vlerick, P. (2020). Mobile pupillometry in manual assembly: A pilot study exploring the wearability and external validity of a renowned mental workload lab measure. *International Journal of Industrial Ergonomics*, 75:102891.

Van Orden, K. F., Limbert, W., Makeig, S., and Jung, T.-P. (2001). Eye activity correlates of workload during a visuospatial memory task. *Human factors*, 43(1):111–121.

Ventre-Dominey, J., Gibert, G., Bosse-Platiere, M., Farne, A., Dominey, P. F., and Pavani, F. (2019). Embodiment into a robot increases its acceptability. *Scientific reports*, 9(1):1–10.

Verhaeghen, P. and Salthouse, T. A. (1997). Meta-analyses of age–cognition relations in adulthood: Estimates of linear and nonlinear age effects and structural models. *Psychological bulletin*, 122(3):231.

Verhagen, P., Kuling, I., Gijsbertse, K., Stuldreher, I. V., Overvliet, K., Falcone, S., Van Erp, J., and Brouwer, A.-M. (2020). The cross-modal congruency effect as an objective measure of embodiment. In *Companion publication of the 2020 international conference on multimodal interaction*, pages 107–111.

Verrel, J., Lövdén, M., and Lindenberger, U. (2012). Normal aging reduces motor synergies in manual pointing. *Neurobiology of Aging*, 33(1):200–e1.

Villani, V., Righi, M., Sabattini, L., and Secchi, C. (2020). Wearable devices for the assessment of cognitive effort for human–robot interaction. *IEEE Sensors Journal*, 20(21):13047–13056.

Vorderer, P., Wirth, W., Gouveia, F. R., Biocca, F., Saari, T., Jäncke, L., Böcking, S., Schramm, H., Gysbers, A., Hartmann, T., et al. (2004). Mec spatial presence questionnaire. *Retrieved Sept*, 18:2015.

Vozar, S. E. (2013). *A Framework for Improving the Speed and Performance of Teleoperated Mobile Manipulators*. PhD thesis.

Wang, C.-A., Baird, T., Huang, J., Coutinho, J. D., Brien, D. C., and Munoz, D. P. (2018). Arousal effects on pupil size, heart rate, and skin conductance in an emotional face task. *Frontiers in neurology*, 9:1029.

Wang, H., Zhang, B., Zhang, T., and Jakacky, A. (2019). Tele-operating a collaborative robot for space repairs with virtual reality. In *2019 IEEE 9th Annual International Conference on CYBER Technology in Automation, Control, and Intelligent Systems (CYBER)*, pages 175–180. IEEE.

Wascher, E., Heppner, H., and Hoffmann, S. (2014). Towards the measurement of event-related eeg activity in real-life working environments. *International Journal of Psychophysiology*, 91(1):3–9.

Wascher, E., Reiser, J. E., Rinkenauer, G., Larra, M., Dreger, F., Schneider, D., Karthaus, M., Getzmann, S., Gutberlet, M., and Arnau, S. (2020). Neuroergonomics on the go. a preview of the potential of mobile eeg for work-place evaluation and design.

Weistroffer, V., Paljic, A., Fuchs, P., Hugues, O., Chodacki, J.-P., Ligot, P., and Morais, A. (2014). Assessing the acceptability of human-robot co-presence on assembly lines: A comparison between actual situations and their virtual reality counterparts. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, pages 377–384. IEEE.

Welfare, K. S., Hallowell, M. R., Shah, J. A., and Riek, L. D. (2019). Consider the human work experience when integrating robotics in the workplace. In *2019 14th ACM/IEEE international conference on human-robot interaction (HRI)*, pages 75–84. IEEE.

Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical issues in ergonomics science*, 3(2):159–177.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic bulletin & review*, 9(4):625–636.

Wonsick, M. and Padir, T. (2020). A systematic review of virtual reality interfaces for controlling and interacting with robots. *Applied Sciences*, 10(24):9051.

Wu, C., Cha, J., Sulek, J., Zhou, T., Sundaram, C. P., Wachs, J., and Yu, D. (2020). Eye-tracking metrics predict perceived workload in robotic surgical skills training. *Human factors*, 62(8):1365–1386.

Xiao, J., Wang, P., Lu, H., and Zhang, H. (2020). A three-dimensional mapping and virtual reality-based human–robot interaction for collaborative space exploration. *International Journal of Advanced Robotic Systems*, 17(3):1729881420925293.

You, E. and Hauser, K. (2012). Assisted teleoperation strategies for aggressively controlling a robot arm with 2d input. In *Robotics: science and systems*, volume 7, page 354. MIT Press USA.

Young, M. S., Brookhuis, K. A., Wickens, C. D., and Hancock, P. A. (2015). State of science: mental workload in ergonomics. *Ergonomics*, 58(1):1–17.

Zheng, B., Jiang, X., and Atkins, M. S. (2015). Detection of changes in surgical difficulty: evidence from pupil responses. *Surgical innovation*, 22(6):629–635.

Zheng, B., Jiang, X., Tien, G., Meneghetti, A., Panton, O. N. M., and Atkins, M. S. (2012). Workload assessment of surgeons: correlation between nasa tlx and blinks. *Surgical endoscopy*, 26(10):2746–2750.

Zhu, Q. and Du, J. (2020). Neural functional analysis in virtual reality simulation: example of a human-robot collaboration tasks. In *2020 Winter Simulation Conference (WSC)*, pages 2424–2434. IEEE.