HUMAN-COMPUTER INTERACTION

Max Pascher

An Interaction Design for AI-enhanced Assistive Human-Robot Collaboration

UDE

UDE

ISBN 978-3-00-079648-7

An Interaction Design for AI-enhanced Assistive Human-Robot Collaboration

Von der Fakultät für Informatik
der Universität Duisburg-Essen

zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften (Dr. rer. nat.)

genehmigte kumulative Dissertation

von

Maximilian Michael Pascher

aus

Bocholt

1. Gutachter: Prof. Dr. Stefan Schneegaß
2. Gutachter: Prof. Dr. Jens Gerken
3. Gutachter: Prof. Dr. Michael Prilla

Tag der mündlichen Prüfung: 04. Juli 2024

DOI: 10.17185/duepublico/82229
URN: urn:nbn:de:hbz:465-20240820-073433-2

Abstract

The global population of individuals with motor impairments faces substantial challenges, including reduced mobility, social exclusion, and increased caregiver dependency. While advances in assistive technologies can augment human capabilities, independence, and overall well-being by alleviating caregiver fatigue and care receiver weariness, target user involvement regarding their needs and lived experiences in the ideation, development, and evaluation process is often neglected. Further, current interaction design concepts often prove unsatisfactory, posing challenges to user autonomy and system usability, hence resulting in additional stress for end users. Here, the advantages of Artificial Intelligence (AI) can enhance accessibility of assistive technology. As such, a notable research gap exists in the development and evaluation of interaction design concepts for AI-enhanced assistive robotics.

This thesis addresses the gap by streamlining the development and evaluation of shared control approaches while enhancing user integration through three key contributions. Firstly, it identifies user needs for assistive technologies and explores concepts related to *robot motion intent* communication. Secondly, it introduces the innovative shared control approach *Adaptive DoF Mapping Control (ADMC)*, which generates mappings of a robot's Degrees-of-Freedom (DoFs) based on situational Human-Robot Interaction (HRI) tasks and suggests them to users. Thirdly, it presents and evaluates the Extended Reality (XR) framework *AdaptiX* for *in-silico* development and evaluation of multi-modal interaction designs and feedback methods for shared control applications.

In contrast to existing goal-oriented shared control approaches, my work highlights the development of a novel concept that does not rely on computing trajectories for known movement goals. Instead of pre-determined goals, *ADMC* utilises its inherent rule engine – for example, a Convolutional Neural Network (CNN), the robot arm's posture, and a colour-and-depth camera feed of the robot's gripper surroundings. This approach facilitates a more flexible and situationally aware shared control system.

The evaluations within this thesis demonstrate that the *ADMC* approach signi-

ficantly reduces task completion time, average number of necessary switches between DoF mappings, and perceived workload of users, compared to a non-adaptive input method utilising cardinal DoFs. Further, the effectiveness of *AdaptiX* for evaluations *in-silico* as well as real-world scenarios has been shown in one remote and two laboratory user studies.

The thesis emphasises the transformative impact of assistive technologies for individuals with motor impairments, stressing the importance of user-centred design and legible AI-enhanced shared control applications, as well as the benefits of *in-silico* testing. Further, it also outlines future research opportunities with a focus on refining communication methods, extending the application of approaches like *ADMC*, and enhancing tools like *AdaptiX* to accommodate diverse tasks and scenarios. Addressing these challenges can further advance AI-enhanced assistive robotics, promoting the full inclusion of individuals with physical impairments in social and professional spheres.

Zusammenfassung

Menschen mit motorischen Beeinträchtigungen stehen vor großen Herausforderungen, einschlie lich eingeschränkter Mobilität, sozialer Ausgrenzung und zunehmender Abhängigkeit von Betreuungspersonen. Assistive Technologien können die Mobilität und Unabhängigkeit der Betroffenen fördern und ihr allgemeines Wohlbefinden verbessern. Allerdings ist zu beobachten, dass die Einbeziehung der Zielgruppe in Bezug auf ihre Bedürfnisse während der Entwicklungs- und Evaluationsphase häufig vernachlässigt wird. Darüber hinaus erweisen sich aktuelle Konzepte des Interaktionsdesigns oft als unbefriedigend. Sie stellen Herausforderungen für die Autonomie der Zielgruppe und die Bedienbarkeit dar, was zu zusätzlichem Stress für die Betroffenen führt. Künstliche Intelligenz (KI) kann die Zugänglichkeit assistiver Technologien optimieren und verbessern, insbesondere durch den Einsatz von assistiver Mensch-Roboter-Kollaboration. Es besteht insoweit eine Forschungslücke in der Entwicklung und Evaluierung von Interaktionsdesignkonzepten für KI-unterstützte assistive Robotik.

Diese Arbeit adressiert diese Forschungslücke, indem sie die Entwicklung und Evaluierung von Ansätzen für KI-unterstützte assistive Robotik optimiert und die Integration der Zielgruppe durch drei wesentliche Punkte verbessert. Erstens werden die Bedürfnisse der Nutzenden von assistiven Technologien identifiziert und Konzepte für die Kommunikation geplanter Roboterbewegungen erforscht. Zweitens führt sie das innovative *ADMC*-Konzept für eine unterstützte Bedienung des Roboters ein. Dieses Konzept basiert auf der situationsspezifischen Erfüllung der an das Mensch-Roboter-Team gestellten Aufgaben durch die Generierung von Kombinationen von Roboter-Freiheitsgraden, die dem Bedienenden vorgeschlagen werden. Drittens wird das XR-Framework *AdaptiX* zur *in-silico*-Entwicklung und Evaluation von multimodalen Interaktionsdesigns und Feedbackmethoden für Anwendungen der KI-unterstützten Bedienung vorgestellt und evaluiert.

Im Gegensatz zu bestehenden Ansätzen wird in dieser Arbeit ein neuartiges Konzept entwickelt, das nicht auf der Berechnung von Bewegungen zu bereits bekannten Zielen basiert. Anstelle von vordefinierten Endpunkten verwen-

det *ADMC* seinen inhärenten Regelmechanismus – z.B. ein Convolutional Neural Network (CNN), die Ausrichtung des Roboterarms sowie ein Farb- und Tiefenbild der Umgebung. Dieser Ansatz ermöglicht ein flexibleres und situationsangepasstes System.

Die im Rahmen dieser Arbeit durchgeführten Studien zeigen, dass der *ADMC*-Ansatz die zur Aufgabenerfüllung benötigte Zeit, die durchschnittliche Anzahl der notwendigen Wechsel zwischen Freiheitsgradabbildungen und die wahrgenommene Arbeitsbelastung der Nutzenden signifikant reduziert. Dies wurde mit einer nicht-adaptiven Eingabemethode verglichen, welche rein kartesische Freiheitsgrade verwendet. Darüber hinaus wurde die Effizienz von *AdaptiX* für die Evaluierung sowohl *in-silico* als auch für Anwendungen in der realen Welt nachgewiesen.

Insgesamt wird der positive Einfluss von assistiven Technologien auf Menschen mit motorischen Beeinträchtigungen hervorgehoben. Auch die Bedeutung von nutzerzentriertem Design, verständlichem Verhalten KI-unterstützter Anwendungen und die Vorteile von *in-silico*-Tests werden betont. Darüber hinaus werden Forschungsmöglichkeiten skizziert, wobei der Schwerpunkt auf der Weiterentwicklung von Kommunikationsmethoden, Szenarien für *ADMC* und Werkzeugen wie *AdaptiX* liegt. Lösungen in diesem Bereich können die Entwicklung von KI-gestützter assistiver Robotik vorantreiben und die vollständige Integration von Menschen mit körperlichen Beeinträchtigungen in soziale und berufliche Bereiche fördern.

Acknowledgements

The support, help, and collaboration of many remarkable people over the years made this work possible. To all of them, I extend my sincere gratitude and hope to follow their example and pay it forward to the next generation of HRI students.

First and foremost, I owe a profound debt of gratitude to my academic advisors, Prof. Dr. Stefan Schneegaß, and Prof. Dr. Jens Gerken, for their unwavering support, continuous guidance, and invaluable encouragement throughout my doctoral studies. Their expertise, mentorship, and constructive feedback were instrumental in shaping this research and completing my dissertation. They remained patient throughout robot troubles, last-minute conference deadlines, and a pandemic – indeed, this work wouldn't have happened without them. I also sincerely thank my dissertation committee member, Prof. Dr. Michael Prilla, for his insightful, detailed, and helpful feedback and advice.

My research collaborators have been essential in accomplishing this dissertation. Kirill, thank you for brainstorming new ideas, working on our fancy prototypes, and the long Slack chats to polish our papers. Stephanie, thank you for navigating the new environment of Gelsenkirchen and HRI research with me. Thank you, Til, Stephan, Jan, and Kevin, for supporting this research through your hard work on your university courses and bachelor/master theses. Younes, Valarie, and Astrid thank you for helping me stay sane during these last hectic months. With this, I sincerely thank all members of *Westphalian University of Applied Sciences* and *TU Dortmund University* for supporting my studies and career.

Jonas and Sarah, thank you for welcoming me to my second home at the *University of Duisburg-Essen* and involving me in your outstanding research. Uwe, thank you for your tireless support, the many questions answered, and the excellent collaboration. Roman, thank you for taking care of all technical issues – big and small – and for creating such beautiful illustrations. And although I did not work directly with Donald, David, and Marion, I am grateful for your support of the whole group in research and non-research activities. Thank you, Carina, Hagen, Nick, and Niklas, for being fantastic office mates, for sharing fun and occasional desperation. Thank you, Jonas, for always ensuring we had enough *Flaschenpost*, especially during hectic paper deadlines. Thank you, Marvin, for organizing and distributing the bachelor/master theses to our students. Jonathan, thank you for your support in all things machine learning and the opportunity to compare and share hiking experiences. Barbara, thank you for managing all the administrative issues in our group – without your tireless work, chaos might ensue. Last – but certainly not least – Alia, thank you for the years together, your support during difficult times, your delicious food, and being such a good friend.

My gratitude goes to my research collaborators, without whom this interdisciplinary work would not have been possible. Annalies, thank you for traveling with me around Germany to interview our target group and for introducing me to the interplay with social sciences. Felix, thank you for your ROS and technical support all this time and for being such a good discussion partner. Your viewpoints have been invaluable.

My research would not have been possible without the very people it hopes to address. To you, participants in home interviews, online talks, and study sessions, I will always be indebted. Thank you for spending your valuable free time to help further this work. I am also greatly indebted to the German Federal Ministry of Education and Research (BMBF) for their financial support through the *MobILe* and *DoF-Adaptiv* projects.

Lastly, I owe my friends and family immense appreciation for their support, understanding, and encouragement throughout these years. Lissy, thank you for being the best mother imaginable! Isabell, Steven, thank you for patiently listening to my – frequently hard-to-understand – research ramblings. Hendrik and Gregor, thank you for the sanity-saving hiking trips, and Doro and Nora for being the most understanding friends a sleep-deprived Ph.D. student could ask for. Finally, I thank my partner, Sarah, for her unconditional love and unbeatable belief in my abilities. Your encouragement and patience have been the cornerstone of my success.

So Long, and Thanks for All the Fish,

Max Pascher

TABLE OF CONTENTS

List of Figures

List of Tables

List of Acronyms

ADL	Activity of Daily Living
ADMC	Adaptive DoF Mapping Control
AI	Artificial Intelligence
AR	Augmented Reality
BMBF	Federal Ministry of Education and Research
CNN	Convolutional Neural Network
DFKI	German Research Center for Artificial Intelligence GmbH
DoF	Degree-of-Freedom
FWBI	Friedrich-Wilhelm-Bessel-Institut Forschungsgesellschaft mbH
HCAI	Human-centred AI
HCI	Human-Computer Interaction
HMD	Head-Mounted Display
HRC	Human-Robot Collaboration
HRI	Human-Robot Interaction
IoT	Internet of Things
KI	Künstliche Intelligenz
LED	Light Emitting Diode
MAR	Mobile Augmented Reality
MR	Mixed Reality
P	Publication
PRISMA	Preferred Reporting Items for Systematic Reviews
Q	Guiding Question
ROS	Robot Operating System
RQ	Research Question
SAR	Spatial Augmented Reality
TOR	Take-Over-Request
TCP	Tool Center Point
UI	User Interface
UNICEF	United Nations International Children's Emergency Fund
VR	Virtual Reality
WHO	World Health Organization
XR	Extended Reality

1
INTRODUCTION

Assistive technology is of fundamental importance for persons with permanent or temporary functional difficulties as it improves their functional ability, and enables and enhances their participation and inclusion in all domains of life.

– WHO and UNICEF [235]

Figure 1.1: ***Filling a glass of water, supported by an assistive robot***. *Robotic arms support users with motor impairments in day-to-day tasks like drinking. Effective implementation requires thoughtful design to enhance user benefits and support self-determined living. © Kevin Rupp, Frankfurt UAS*

1.1 Motivation and Research Gap

In 2023, the World Health Organization (WHO) reported that 15% of the global population lives with a disability [234]. In Germany alone, by the end of 2021, 7.8 million individuals were officially classified as *severely disabled* [206]. Notably, more than 58% of these cases were associated with physical disabilities, affecting a total of 4.5 million people. For many in this demographic, motor impairments have led to a significantly – and often permanently – reduced ability to move their extremities. This diminished mobility frequently leads to exclusion from social and professional contexts [235] and an increased reliance on caregivers for daily tasks, creating a near-constant need for their presence [142]. Furthermore, these concerns are amplified by the ageing population, who often encounter similar challenges to those with disabilities [204], and the growing demand for *ageing in place* solutions [152].

Assistive technologies serve as essential tools for individuals with motor impairments, frequently providing practical solutions to enhance their mobility and independence [51, 141, 161] (see Figure 1.1). These technologies, ranging from simple aids to advanced robotic systems, address specific needs and empower users to navigate daily tasks with greater autonomy [22, 45]. By reducing dependence on caregivers and facilitating participation in social and professional activities, assistive technologies contribute to substantially enhanced well-being, health, and overall eudaimonia [208]. While caregiver fatigue is well-documented [23, 114], care receivers also experience weariness when consistently surrounded by their assistants [53]. The opportunity to spend time without constant human presence, even for a few hours – and possibly facilitated through assistive technologies – can enhance the quality of life for individuals with motor impairments [30]. By lessening the constant need for caregivers, those once reliant on human assistance are empowered to regain their independence and achieve valuable alone-time. Recent advances in (semi-)autonomous technologies have made this level of support possible and prompted the onset of robotic device integration into selected aspects of our personal and professional lives. These innovations foster close-quarter collaborations with robots across diverse domains, spanning from industry assembly lines [28] to mobility aides [66] and patient care [182].

Focusing specifically on assistive robotic systems, Kyrarini et al. – in their comprehensive literature review – underscored the positive impact of these so-called *cobots* in supporting individuals with motor impairments in Activities of Daily Living (ADLs) [128]. Among these systems, assistive robotic arms emerge as a particularly valuable and versatile subset of collaborative technologies, capable of autonomously performing everyday pick-and-place operations [21]. Yet, new challenges arise when robots are assigned (semi-)autonomous tasks, potentially introducing additional stress for end users. As such, they need to be adequately addressed before and during the (pre-)ideation, design, and development process [176] – hereinafter denoted as *research process*. Notably, Pollak et al. [176] emphasise the diminished sense of control experienced by users when using a cobot's autonomous mode. Their study demonstrates that transitioning to manual mode not only enabled participants to regain control but also reduced stress levels noticeably. These findings are further corroborated by Kim et al., whose comparative study on control methods demonstrate markedly higher user satisfaction among the manual mode cohort [117].

In contrast to standardised industrial settings, care environments demand flexibility as cobots assist non-tech-savvy users in various, often highly situation-dependent, ways [54, 66, 71, 120, 210]. Operating robots in these contexts remains a significant challenge, as users need to consistently maintain oversight to operate the system effectively and safely. As emphasised by Stephanidis et al., *transparency, understandability* and *accountability* are foundational elements for achieving successful Human-Computer Interaction (HCI) [208]. Yet, a fundamental issue arises from the type of robots used, as multiple Degrees-of-Freedom (DoFs) require either complex multi-dimensional input devices or a division into different modes with two-dimensional joystick controls [140, 177]. The former is often impractical for individuals with motor impairments [44, 116], while the latter introduces time-consuming mode switches, frequently increasing task completion times [103]. Consequently, these commonly used control methods often prove unsuitable when designing assistive robotic solutions for people with motor impairments.

Addressing these issues, shared control systems – leveraging the advantages of Artificial Intelligence (AI) – can streamline and enhance robot operation accessibility in Human-Robot Interaction (HRI) / Human-Robot Collaboration (HRC) [2, 193]. Their success depends on well-designed communication of the robot's motion intent, ensuring users are consistently aware of and understand the level of support provided by the system [1]. Additionally, accommodating different users and their respective abilities may require tailored input devices or customised multi-modal feedback methods [105]. These challenges are compounded by the inherent difficulties in conducting research studies that require the physical interaction of robots and humans. Logistical complexities, high transportation costs, safety concerns related to robots and associated teething problems, recruitment challenges, and the limited availability of the target user groups collectively contribute to the intricacy of HRI research [12, 34, 120, 122, 147].

Research Gap

A notable research gap exists in the development and evaluation of interaction design concepts and multi-modal feedback methods for AI-enhanced assistive robotics, specifically when aiming for a user-centred design process to empower individuals with motor impairments in their day-to-day lives.

1.1.1 Motivating AI-enhanced Shareds Control for Assistive HRI

In a typical scenario motivating my research, a wheelchair-mounted assistive robotic arm, such as the *Kinova Jaco 2* [118], enables users to perform essential ADLs like drinking or eating (see Figure 1.2). This setup presents users with the challenge of operating six or more DoFs, necessitating complex input devices or cumbersome and potentially confusing mode switches. While manual control systems allow individuals to manoeuvre assistive robotic arms

(a) (b)

Figure 1.2: ***Usage of an assistive robotic arm***, *in a domestic setting. In* ***(a)****, the robotic arm is used to drink a glass of water;* ***(b)*** *shows the mounting to a wheelchair. The user controls each robot DoF sequentially via a joystick mounted on the armrest, with a button integrated into the wheelchair's headrest to alternate through the DoFs.*

with direct movement control, these methods are often time-intensive and, in current commercial offerings [14, 64, 118], limited to predefined movements along specific DoFs. In contrast, autonomous robots can manage ADLs for people with motor impairments but may introduce a new form of dependence, potentially compromising the desired sense of autonomy [20].

AI-based shared control systems, as detailed by Erdogan and Argall [62], represent a continuum of autonomy. This spectrum ranges from systems predominantly favouring manual control – and only subtly refining user inputs – to those where users mainly issue high-level commands for robotic execution. The gradient represents the nuanced balance of user involvement, from direct, moment-to-moment control to more abstract, directive interactions, illustrating the diverse approaches to augmenting human capabilities through robotics [208]. In uniting both approaches, shared control systems involve users directly in the control loop [179, 238] on an operational level [70].

These systems consider user inputs while also incorporating robotic decisions or suggestions to enhance system usability. This dual approach aims to maintain user independence by allowing personal input and autonomy while benefiting from the efficiency and precision of automated technology [109, 157]. Shared controls thus represent a balanced solution by merging human intuition with robotic efficiency to improve the usability of assistive technologies [15, 29, 68].

1.1.2 Motivating a Shared Control Framework in Assistive HRI

Shared control approaches are viewed with cautious optimism for their potential to enhance effective HRI, but several challenges currently impact – and potentially hinder – their widespread implementation. Addressing these obstacles is essential for advancing and practically implementing shared control systems for assistive technologies where user-centred design and adaptability are desirable [76, 111]. Researchers encounter challenges during the research and evaluation stages, including:

Challenge 1 – Exploration of Shared Control Systems: Developing shared control systems for assistive technologies requires extensive experimentation, fine-tuning, and balancing between user and robot control input [130].

Challenge 2 – Understanding Robot Behaviour: Despite extensive research into design solutions to communicate *robot motion intent*, clear insights into the entities, dimensions, and relations are still lacking. In assistive robotics, visualisation and feedback modalities must be precisely tailored to individual user needs and abilities, as there is no *one size fits all* solution [105].

Challenge 3 – Diverse Requirements for Input Devices: The suitability of input devices can vary greatly between users. Individual capabilities and needs may necessitate multi-modal input or require selecting from different input modalities [12].

Challenge 4 – Conducting In-person User Studies: The physical intermingling of robots and humans for research studies presents substantial challenges, including logistical and transportation complexities, safety concerns with robots in close proximity to humans, and the varying availability of target users [34].

Addressing these challenges presents difficulties for shared control researchers, as rigorous testing demands considerable resources and time, often constraining flexibility for the *to-be-tested* variables, such as the level of assistance the system provides, interaction design, intervention strategies, and use-case scenarios. Furthermore, the bulky, costly, and intricate nature of assistive robotic arms in associated research studies may hinder the involvement of the target group in the research process due to logistical and safety concerns. As one option to to facilitate a holistic and flexible approach to these challenges, my thesis proposes a modular and open framework for *in-silico* development and evaluation of shared control approaches, including options for updated suggestions, attention detection and guiding, as well as multi-modal control and feedback support. Adopting a simulation approach akin to those already successfully employed in industrial settings [144, 160, 217] allows for the exploration of different shared control applications while integrating various input devices and visualisation modalities, while addressing – at least partially - the challenges of integrating the target group in the research process.

Research Opportunities

Motivating AI-enhanced Shared Controls for Assistive HRI: Applications of AI-enhanced shared controls can combine the benefits of (semi-)autonomous actions with the flexibility of manual controls. This presents a significant research opportunity to fine-tune the optimal balance between human and algorithmic control inputs, enhance system legibility, and devise intervention strategies to improve the usability and accessibility of shared control-based assistive technologies.

Motivating a Shared Control Framework in Assistive HRI: AI-enhanced shared control applications require rigorous testing to determine the system's optimal parameters, which often entails considerable resource and time investments. A promising research opportunity involves the development of a testbed environment for *in-silico* research and evaluation of assistive robot shared control applications, serving as a beneficial intermediate stage prior to real-world testing. This approach can increase flexibility and resource efficiency for HRI researchers while facilitating the early involvement of target users in the research process.

1.2 Research Structure and Key Findings

This thesis addresses the research gap through a three-step process (STEP I–III). Each stage includes a central guiding question (Q) and highlights key findings from my research contributions. Any corresponding publication (P) is illustrated in Figure 1.3 and listed in *Related Publications* for each step.

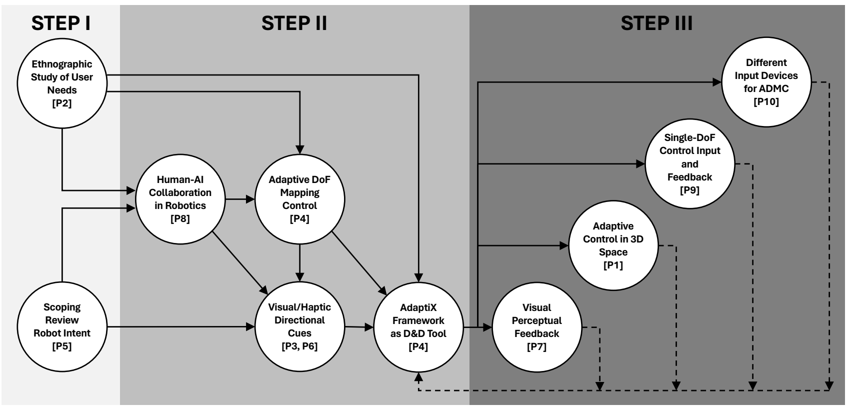


Figure 1.3: ***Overview of the thesis research structure****, including an architecture of the thesis's thematic connections across STEP I–III, along with the corresponding publications.*

In STEP I, recommendations are provided to understand the needs of target users with motor impairments regarding assistive technologies and the potential role of assistive robotic arms in addressing these needs. These recommendations stem directly from the user group and are derived through an ethnographic study. Further, the concept of *robot motion intent* within HRI is defined through a comprehensive scoping review, and its entities, properties, dimensions, and relationships are clarified to enhance communication and collaboration among researchers. Based on the findings in STEP I, STEP II delineates key factors influencing cooperation between humans and AI-enhanced robots, explicitly focusing on motion legibility, integration of user control input, and intervention possibilities. This step also presents a new shared control application with diverse visualisation options for DoF mappings and concludes with introducing the comprehensive *AdaptiX* framework. Lastly, STEP III reports empirical concept evaluations of *AdaptiX* in Virtual Reality (VR), Mixed Reality (MR), and Augmented Reality (AR) environments – consolidated under the term Extended Reality (XR) [95] – in remote and laboratory studies. During the evaluation process, the framework was improved based on the feedback received from each individual evaluation, following a user-centred design process [106].

The individual sections reported in this thesis are embedded in two research projects (*MobILe*: 16SV7866K and *DoF-Adaptiv*: 16SV8565), supported by the *German Federal Ministry of Education and Research (BMBF)*. Here, I collaborated with different partners from *TU Dortmund University, Westphalian University of Applied Sciences, Frankfurt University of Applied Sciences, German Research Center for Artificial Intelligence GmbH (DFKI), Friedrich-Wilhelm-Bessel-Institut Forschungsgesellschaft mbH (FWBI), HIDREX GmbH, pi4_robotics GmbH,* and *munevo GmbH*. The aim of both research projects was to use (semi-)autonomous robotic arms to support people with motor impairments by enhancing user capabilities and supporting self-determined living.

I switch from *I* to the scientific plural *we* when referring to papers conducted in collaboration with others. The contributing authors are clearly identified in the *Statement of Contributions* following the last chapter of this work.

1.2.1 STEP I: Analysing User Needs and Robot Intent Communication

STEP I delves into the early stages before the ideation process by clarifying user's needs from their own perspective and by delineating the various concepts associated with robot intent communication. The latter addresses the necessity for a mutual understanding between humans and their robotic assistants through a scoping review that clarifies various aspects of *robot motion intent* within HRI. The former comprises an ethnographic study of individuals with motor impairments, specifically their execution of ADLs. This study identified essential requirements for an assistive robotic arm tailored to day-to-day tasks. The step centres around the guiding question:

Q1 How can legible AI-enhanced assistive robots be effectively integrated in domestic care settings while accounting for subjective needs?

STEP I Key Findings

User-derived Recommendations: Current assistive technologies often adapt solutions initially designed for the general population to address the perceived needs of individuals with motor impairments, often without direct involvement from the target group. Our ethnographic study [P2] emphasises that structural, social, and collaborative concerns exist and that they need to be addressed throughout the research process. This approach yielded eleven recommendations for designing robotic drinking aids, informed by the expressed needs and observed lived experiences of our study participants. A significant finding from our research is the frequent desire expressed by participants to spend time without constant assistance, a need that can be met by developing effective and safe assistive technologies. Focusing specifically on a drinking aid, the positive adaptation of such technology would enhance access to hydration, a critical aspect often necessitating the presence of caregivers. Moreover, our user-centred approach underscores the benefits of involving the target group in the early research stages to enhance the likelihood of future user acceptance of assistive technology. Finally, our utilisation of *Google Cardboard* to simulate the use of a robotic arm offers an efficient and flexible solution, showcasing the effectiveness of VR technology to introduce HRI scenarios to users.

Intent Communication Model: Establishing effective communication of a robotic system's intent to users is essential for fostering collaboration and preventing task failures in HRI. Based on insights from our systematic literature review [P5], we devised a model for *robot motion intent* communication. This model highlights the primary entities of *robot*, *intent*, and *human* and identifies a communication flow among them resembling the classic HCI model introduced by Schomaker [191]. A central finding from our research is the significance of factors such as *attention*, *state*, and *instruction*, which often serve as essential pre- or post-cursors for explicit *motion* intent communication. Furthermore, we extracted several empirical findings that underscore the connection between intent information and intent location. These insights are important for developing more legible feedback methods to foster a shared understanding of robot behaviour. Our work supports researchers to better align their work with the suggested dimensions, thereby making it easier to assess and compare different studies (e.g., [137, 138, 192, 222]).

Related Publications: [P2] and [P5]

1.2.2 STEP II: Concept & Design of AI-enhanced Assistive Robotics

STEP II explores challenges linked to the collaboration between humans and AI-enhanced robots to extract empirical implications – covering the challenges of legibility, user control, and intervention – and to fine-tune user and system interactions. Based on these findings, a novel shared control approach – *Adaptive DoF Mapping Control (ADMC)* – is proposed, which utilises the multi-dimensional mapping of a robot's DoFs to allow for a simplified control by the user with a low-DoF input device. Subsequently, a set of several feedback methods – involving visual and haptic modalities – were conceptualised and developed accompanying *ADMC* to communicate suggested DoF mappings. The step concludes with introducing and implementing the *AdaptiX* framework. This framework includes a VR simulation environment and extensive customisation options, and as such, streamlines the integration of the target group into the research process, reduces operational overhead, and enhances overall efficiency.

The guiding question in STEP II is:

Q2 How can we engage in collaborative efforts with an AI-enhanced system by reducing input complexity?

STEP II Key Findings

Challenges in Human-AI Collaboration for Assistive Robotics: AI-enhanced systems can support users – especially when their interaction modalities are limited – in controlling an assistive robotic arm. In our work [P8], we identified three primary challenges associated with interacting with an AI-enhanced assistive robot and derived corresponding empirical implications. These insights were integral to the development of the *AdaptiX* framework [P4].

Shared Control Approach: Expanding on prior research and related work, we introduce *ADMC*, a context-aware shared control approach using multiple robot DoFs mapped onto a low-DoF standard input device (e.g., a joystick) [P4]. This eliminates the need for complex multi-dimensional input devices – which are often impractical for individuals with motor impairments – and reduces time-consuming mode switches during Cartesian robot control. Adopting the *ADMC* approach can facilitate the simplified and accessible operation of an assistive robotic arm for ADLs, thereby increasing independence from constant human care and company.

Feedback of Directional Cues: We devised visual feedback methods by leveraging insights from our research on *robot motion intent* communication [P5]. These visualisations create a safe and collaborative environment, allowing users to comprehend movement directions and implications based on DoF mappings [P6]. Additionally, we demonstrate the efficacy of vibrotactile haptic feedback in conveying directional cues from the robot [P3]. This offers an alternative to visual feedback, catering to individuals with visual impairments or addressing visual clutter from augmented information in a given scenario.

Comprehensive Framework: Ensuring effective addressing of the target group's concerns and well-designed intent communication is imperative for implementation of shared control systems in assistive robots. Our *AdaptiX* framework [P4] facilitates the development and assessment of robotic applications *in-silico*. It serves as a vital intermediary between concept ideation, development, and evaluation, offering HRI researchers enhanced flexibility and promoting efficient resource allocation (e.g., [83]). Notably, using a virtual model during the development stages simplifies the seamless integration of the target group.

Related Publications: [P3], [P4], [P5], [P6], and [P8]

1.2.3 STEP III: Evaluating Interaction Design & Robot Intent Communication

In STEP III, I present a visual approach to communicate the robot's world perception, ensuring object-aware navigation. Additionally, this phase involves the evaluation of our *AdaptiX* framework and the *ADMC* shared control approach. Using the XR framework introduced in STEP II, we conducted both remote and laboratory studies, employing a virtual robot in a simulation environment or a physical robot in the real world. The evaluations include systems with varying numbers of DoFs for user input and exploration of different input devices. The guiding question is:

Q3 How does the *AdaptiX* framework support researcher in developing and evaluating *ADMC* and other shared control approaches using low-DoF input devices?

STEP III Key Findings

Number of Input-DoFs: In our initial evaluation [P1], we employed a two-dimensional input for our shared control approach [P4], with *ADMC* utilising one axis for *Optimal Suggestion* and the other for *Adjusted Suggestion*. Participants raised concerns about the dynamically changing mapping of combined DoFs and the two-DoF input device. Subsequent studies [P9, P10] adopted a one-dimensional input for *ADMC*. Here, we demonstrated that adaptive controls significantly reduce task completion time, average number of necessary mode switches, and perceived workload compared to Cartesian control.

Suggesting an Updated DoF Mapping: Leveraging the benefits of adaptive controls, we conducted a comparison between two *ADMC* variants: *Continuous* and *Threshold*, distinguished by the time of suggestion communication to the user [P9] via legible and straightforward visualisations [P7]. We found no significant differences between *Continuous* and *Threshold*, which suggests that both discrete and continuous communication of movement suggestions enable users to efficiently utilise adaptive control methods. Qualitative interviews further supported these findings.

Exploring Input Devices: Expanding on the *Threshold* variant of *ADMC*, we employed *AdaptiX* to assess three distinct input devices for commanding a physical assistive robotic arm: a motion controller, assistive buttons, and a head-based approach [P10]. Although all participants successfully controlled the robotic arm with each input device to accomplish the project task, the study's results underscore the heightened effectiveness of hand-operated input methods compared to a head-based interaction approach.

Related Publications: [P1], [P4], [P7], [P9], and [P10]

1.3 Contribution

Assistive technologies empower individuals with disabilities by promoting self-determination and independent living [203], reducing dependence on human caregivers [10, 151], and enabling people to stay in their own homes [87] – a key consideration in an ageing population [3]. However, these systems are often designed *for* the target group with limited input *from* the community. Adopting a user-centred design process enhances user acceptance of upcoming assistive technologies by incorporating users' insights into their respective capabilities and needs [86, 98]. Building on this perspective, this thesis delves into the research process, presenting user-derived recommendations for an assistive drinking aid [P2] and a robot intent communication model [P5]. The model aids future researchers in aligning their work with suggested dimensions, facilitating the assessment and comparison of HRI studies.

We built on Goldau and Frese's two-dimensional shared control approach [84] and extended it into three-dimensional space, thereby increasing the potential DoFs and enabling a more precise representation of ADLs [P4]. It mitigates the need for complex – and often impractical for individuals with motor impairments – multi-dimensional input devices and time-consuming mode switches during Cartesian robot control through DoF mappings. For an legible behaviour, we developed visual and haptic feedback methods to communicate the resulting movement direction [P3, P6], including an arrow-based gizmo visualisation approach, in the *AdaptiX* framework.

The comprehensive *AdaptiX* framework, with its modular architecture and additional functionality, facilitates the development and evaluation of assistive robot control applications *in-silico* and in real-world settings, offering enhanced flexibility, promoting efficient resource allocation, and integrating the user group into the research process. Moreover, this thesis evaluates the *AdaptiX* framework and *ADMC* shared control approach. *AdaptiX* effectively enables the research of new interaction designs and feedback techniques *in-silico*, supporting real-time suggestions by user attention guiding. It also allows quick assessments of different input devices through its standardised

User Input Adapter, successfully serving as an interface between the physical robot and virtual communication via a XR Head-Mounted Display (HMD).

Research Contribution

This thesis contributes to AI-enhanced assistive robotics in domestic care settings through three key advancements. First, it provides an understanding of user needs regarding assistive technologies and delineates various concepts associated with *robot motion intent* communication. Second, the innovative shared control approach *ADMC* is presented, which generates DoF mappings based on the situational HRC task and communicates them as suggestions to the user. Third, the XR framework *AdaptiX* is introduced and evaluated for *in-silico* development and evaluation of multi-modal interaction designs and feedback methods for shared control applications. This work effectively narrows the research gap by facilitating the development and evaluation of shared control approaches while simplifying user integration into this process.

2
BACKGROUND AND DEFINITIONS

Chapter Two contextualises the thesis within the current HRI / HRC research landscape and defines foundational terminology essential for subsequent discussions. Key focal points include assistive robotics in domestic care, applications of shared control methodologies, and the implementation of multi-modal feedback strategies.

2.1 Human-Robot Interaction & Collaboration

The term *robot* encompasses a diverse array of (semi-)automated devices characterised by varying capabilities, technologies, and physical forms [88]. These versatile entities have the potential to contribute to our daily lives by providing assistance in workplaces, aiding with household tasks, and even accompanying us in public spaces [4, 17, 139]. Despite variances in DoFs and mobility among these cyber-physical systems, their diverse applications can augment human capabilities and improve efficiency [74]. Initially championed in industrial settings, robots performed strenuous tasks like manipulating and welding heavy components, typically within confined areas [102, 229]. And while they were originally viewed as mere tools to be operated remotely by human workers [187, 236], their evolution in functionality and purpose has been remarkable. The advent of lightweight materials [50, 85, 134] and the integration of safety sensors [55, 175, 225] enabled robots to become more adaptive to human presence, facilitating them to shut down or adjust operations when humans are in close proximity or when encountering resistance [78, 92, 188].

These advancements have led to the emergence of *cell-less* HRI [18], facilitating innovative applications like collaborative assembly tasks [196], crafting [173], or assisting people with disabilities in their daily activities [170]. A comprehensive review by Ajoudani et al. investigated various approaches to HRI, charting their evolution and emerging adoption over the past two decades [4]. They highlight the success of combining human cognitive skills – intelligence, flexibility, and responsiveness – with the precision and capacity for repetitive tasks inherent in robots. To delineate the spectrum of human-robot teaming, Matheson et al. categorised various types of *cell-less* HRI based on the degree of interaction proximity [143]. These include *coexistence* (sharing the same space at different times), *synchronised* (occupying the same space but not concurrently), *cooperation* (no spatial or temporal separation but working on separate tasks), and *collaboration* (jointly working on a task with interdependent actions).

These categorisations underline the importance of communication and transparent interaction in successful HRI. While advancements in human-aware navigation primarily focus on enabling robots to interpret and respond to human behaviour [126], it is equally imperative to facilitate humans' understanding of robotic conduct [124]. As highlighted by Matheson et al., the increasing physical intermingling of humans and robots accentuates the significance of effectively communicating *robot motion intent* for safe and efficient collaboration, constituting a foundational aspect of explainable robotics [143].

2.2 Assistive Robotics in Domestic Care

Optimising and streamlining collaborations becomes particularly important when designing for vulnerable user groups, such as with assistive technologies for people with motor impairments. Assistive robotics have the potential to substantially enhance independence and improve care by providing support and alleviating the burden on caregivers, thereby improving the quality of life for those requiring care [22, 36, 101, 117, 148, 214]. Exactly how assistive robotic systems can assist individuals with motor impairments has gained increasing attention in research. Notably, the *Robots for Humanity* project led by Chen et al. [43] and a seminal study by Fattal et al. [66] examined the feasibility and user acceptance of such technologies. While the ultimate aim is to fully (re-)integrate individuals with motor impairments into professional and social spheres, a recurring observation in their works is that current assistive technologies primarily focus on enabling the performance of ADLs [174]. These actions range from fundamental tasks such as eating and drinking to more complex activities like grooming and dressing [47].

Ongoing technological advancements are continually expanding the capabilities and enhancing the performance of cobots in the tasks they can perform for – and with – their users. In a study by Gallenberger et al., camera systems and machine learning were integrated into an autonomous robotic feeding system to identify food types and plan the picking and delivery process to the user's mouth [75]. Another approach, as detailed by Canal et al., used a learning-by-demonstration framework for feeding tasks [36]. These projects highlight robotic capacity to autonomously complete tasks with minimal user intervention, focusing on the technical aspects of developing assistive technology rather than fine control by users. This emphasises that implementing safe and user-friendly robotic solutions can fundamentally improve the quality of life for individuals requiring assistance [185]. Additionally, by assisting caregivers in their responsibilities and even facilitating certain tasks to be accomplished without human assistance, the overall quality and accessibility of care are enhanced [125]. The resulting increase in independence is particularly beneficial for individuals with motor impairments, meeting the community's desire for extended periods of alone-time and privacy [168].

In their research, Drolshagen et al. discovered that individuals with disabilities generally adapt well to working alongside cobots, even in close proximity [59]. Overall, robotic assistance tends to be positively received by people with motor impairments, especially when their specific needs are considered in the design process [67], and when sufficient oversight is provided to ensure a sense of security [24]. As such, effective communication of the robot's motion intent emerges as a crucial factor for achieving high acceptance among end users. These findings are corroborated by Beaudoin et al., who investigated the long-term usage of the *Kinova Jaco* – a robotic arm representing a notable advancement in assistive technology [19]. Their study covered various themes, including improvements in daily task capabilities, satisfaction levels with the *Jaco* system, psychological impacts, and the broader implications for users and their caregivers. According to Beaudoin et al., nearly all participants reported increased autonomy in certain aspects of life and noted positive psycho-social effects. Notably, a marked success was the improved ability of participants to drink independently using the *Kinova Jaco*, thereby reducing their dependence on human assistance while simultaneously increasing their well-being and health through continuous access to beverages [35, 49, 97].

2.3 Shared Control Applications for Robots

The appropriate level of autonomy in assistive robots is a common point of contention. Highly autonomous systems [218] that reduce user interaction to mere oversight have been found to elicit stress [176] and feelings of distrust [242] among users. Conversely, at the other end of the spectrum, manual controls with only minor adjustments to the user's input [201] can be challenging, or even impossible, for users with certain types or degrees of impairments to operate [43, 112]. Shared control presents a middle ground by integrating manual user operation through standard input devices with algorithmic assistance from computer software to adjust the resulting motion. This approach addresses concerns associated with purely autonomous systems and manual controls [1]. In shared control, there is a collaborative effort between the user and the robot on the operational level, empowering individuals with motor impairments to actively participate in their care. Consequently, such methods can enhance the sense of independence and improve ease of use compared to entirely manual controls [70]. By maintaining a balance between autonomy and user involvement, shared control systems can provide a more acceptable and comfortable experience for individuals relying on assistive technologies [90, 181, 220].

A distinct approach is the shared control system proposed by Goldau and Frese [84]. This system integrates the cardinal DoFs of a robotic arm based on the situational task and aligns them with a low-DoF input device. The process involves attaching a camera to the robotic arm's gripper and using a Convolutional Neural Network (CNN) trained on ADLs performed by individuals without motor impairments [84], akin to the learning-by-demonstration method used in autonomous robots [36]. The CNN then provides a set of re-mapped DoFs, ranked according to their predicted effectiveness for the given situation, allowing users to execute a variety of movements as required. Furthermore, this CNN-based approach offers extensibility, as it can be trained to distinguish between many different situations, enhancing its practicality for everyday use. In their proof-of-concept study, which involved a two-dimensional simulation environment featuring a robotic gripper representation and a target object, Goldau and Frese [84] observed faster task execution with the proposed system than manual controls. However, users perceived the shared control approach as more complex, expressing their preference for a more extensive training phase – even in this low-DoF environment. Their findings highlight the need for more intuitive and responsive interaction feedback when controlling a robot.

My research in AI-enhanced assistive robotics builds on Goldau and Frese's approach, but extends it from two dimensions to three-dimensional space. This extension increases the potential DoFs, enabling a more precise representation of ADLs. By incorporating additional functionality, visualisations, and a Robot Operating System (ROS) integration, my work facilitates the development and evaluation of innovative interaction and control methods based on a shared control approach.

2.4 Multi-Modal Feedback Methods

Safe and effective HRC relies on a seamless communication of *robot motion intent* [143]. The subsequent sections offer an overview of multi-modal feedback methods designed to convey directions and guide the user's attention effectively .

2.4.1 Directional Visualisations

In recent years, AR technology has increasingly been applied in HRC contexts [13, 56], with much prior research focusing primarily on HMDs, Mobile Augmented Reality (MAR), and Spatial Augmented Reality (SAR) to visualise *robot motion intent* [93, 184, 228]. Rosen et al. showed that AR could offer significant improvements over traditional desktop interfaces in visualising intended robot motions [184]. Their study demonstrated that a AR/MR HMD allows humans to determine where the robot is going to move more quickly and accurately compared to existing two-dimensional baselines. However, existing literature predominantly addresses visualising motion intent for autonomous robotic systems [9, 38, 52, 93, 209, 230]. These studies explore how AR can effectively communicate the planned trajectory or behaviour of robots, thereby enhancing predictability and safety in shared environments. In their comprehensive literature review, Suzuki et al. examined the interplay between AR and robotics in greater detail, highlighting the potential of AR-based visualisations for conveying information such as movement trajectories or the internal state of the robot [212]. Their work underscores the potential of AR technologies in enhancing communication between humans and robots, which is crucial for safe and efficient HRC. However, their review does not delve deeply into the specific categorisations or intricacies associated with intent, suggesting a need for further research in understanding and classifying its multifaceted nature. Moreover, limited attention has been given to communicating suggestions for robot intentions and associated control modalities. This approach not only displays the robot's current or planned actions but also provides users with intuitive interfaces and feedback mechanisms to understand and influence these actions.

In my research, I am using different XR feedback technologies, including VR, MR, and AR. Within this research scope, we implement visual feedback by simulating AR in a VR environment and utilising directional cues registered in three-dimensional space. This method allows users to comprehend various movement directions for both actual control and the suggested DoFs combinations. A primary strategy involves the use of arrows, a straightforward and universally understood visualisation technique as demonstrated in prior works [197, 200, 228]. The overarching goal is to provide users with an intuitive and efficient means to interpret and interact with the suggested movements and controls of the robotic system.

2.4.2 Vibrotactile Haptic Feedback

Research on vibrotactile signals as a feedback modality highlights their efficacy in directing attention, guiding actions, and conveying patterns [94, 96, 231]. Barralon et al. conducted a study on pattern recognition using a vibrotactile belt equipped with eight actuators, where participants were tasked with selecting the correct corresponding visual representation [16]. Lee and Starner proposed *BuzzWear*, a wearable tactile display with three vibration actuators designed for notification purposes by modulating intensity, pattern, direction, and starting point [131]. Their findings show that after 40 minutes of training, subjects could already distinguish between 24 patterns with up to 99% accuracy. Vibrotactile haptic feedback finds application in guidance contexts as well. Lehtinen et al. conducted a study using a vibrotactile glove to assist in a visual search task on a flat plane displayed on a wall [132]. Their findings showcased that the impact of visual complexity can be significantly mitigated through enhanced spatial precision in the guidance provided.

However, while certainly successful for specific tasks, a prevalent challenge in tactile displays is their limited resolution. To address this constraint, researchers have turned to tactile illusions for simulating smooth movement patterns [46], such as *Phantom Sensations* [6, 169], *Apparent Tactile Motion* [33, 119, 195], and the *Cutaneous Rabbit* illusion [79, 149, 180, 205]. Tan et al. performed a study using a 3 x 3 tactile display, employing the *Cutaneous Rabbit* sensation to communicate eight two-dimensional directional cues. They achieved successful cue recognition, demonstrating the attainable spatial resolution when using vibrotactile haptic feedback [213]. While prior research has predominantly concentrated on two-dimensional directional cues (e.g., [213]) or enabled users to feel directions as they approach with their hand (e.g., [94]), our objective is to convey three-dimensional directional cues. We extend the work established by Tan et al. [213] for communicating two-dimensional directions and expand upon it by combining their base with pulse or intensity mapping to also communicate the gradient. This novel approach leverages vibrotactile feedback for more nuanced and spatially aware interactions, potentially improving user navigation and understanding of robotic movement or environmental cues.

2.4.3 Feedback Modalities for User Attention Guidance

In shared control systems, directing user attention to the robot's assistance is imperative to prevent potential hazards such as collisions [172]. This becomes especially important when either party manoeuvres the robot in a manner that might exacerbate the situation or when something out of the ordinary happens. Various feedback modalities have been proposed to guide user attention as a supplementary feedback mechanism alongside AR [26, 40, 159]. Incorporating multi-modal feedback – including auditory, visual, and tactile/haptic signals – improves system predictability and user response, thus facilitating prompt and informed decision-making.

Similar strategies are employed in autonomous vehicles, which issue Take-Over-Requests (TORs) for the human operator to regain control in complex situations [8, 239]. This TOR prompts the driver to assume manual control of the vehicle to prevent a collision or to drive in areas the vehicle cannot handle autonomously. Auditory, visual, and tactile/haptic modalities are commonly used for TORs [240] – either as a single sensory input [172] or a combination of multiple variants [171]. Simulation studies, along with research on reaction times to different sensory stimuli, indicate that multi-modal feedback results in the lowest possible reaction times in shared control systems [31, 57, 123]. By integrating these feedback mechanisms, existing assistive robotics can enhance safety and efficiency. Many systems including robots are already equipped with screens, speakers, and vibration motors, facilitating the addition of robust, multi-modal feedback and – thereby – improving the overall effectiveness of shared control systems. Based on these insights, I propose using multi-modal feedback as an effective means to convey an update to suggestions for DoF mapping and to garner user attention.

3

AI-ENHANCED ASSISTIVE HUMAN-ROBOT COLLABORATION

The availability of context and the use of context in interactive applications offer new possibilities to tailor applications and systems on-the-fly *to the current situation. However, context influences and often fundamentally changes interactive systems.*

– A. Schmidt [190]

Chapter Three delineates the contributions of ten research papers centred on designing and developing interaction and feedback modalities for assistive HRC. In the initial stage – STEP I – the discussion revolves around two papers: [P2], which analyses user needs and concerns regarding assistive robotic arms, and [P5], which covers a comprehensive literature review on how robots communicate their intent. STEP II details the concept and design of AI-enhanced assistive robotics, beginning with the development of a shared control approach by mapping DoFs [P4, P8]. It further examines different approaches for communicating these DoF mappings, including visualisations [P6] and vibrotactile haptic feedback [P3]. Further, a transitional XR framework is introduced, serving as an innovative *in-silico* testbed environment which can be used throughout the research process and for evaluating shared control applications in user studies [P4]. Publication [P4] contributes to STEP II and STEP III, with the former concentrating on the framework's development and the latter exploring its practicality through remote [P1] and laboratory studies [P9, P10], employing a virtual robot in a simulation environment or a physical robot in the real world. Due to revisions between these user studies, *AdaptiX* was evaluated in different versions. Lastly, STEP III details the development and evaluation of visualisations communicating the robot's environment perception to the user, particularly in terms of detected obstacles [P7]. Each step includes specific research questions, an introductory summary, and detailed insights from the corresponding papers.

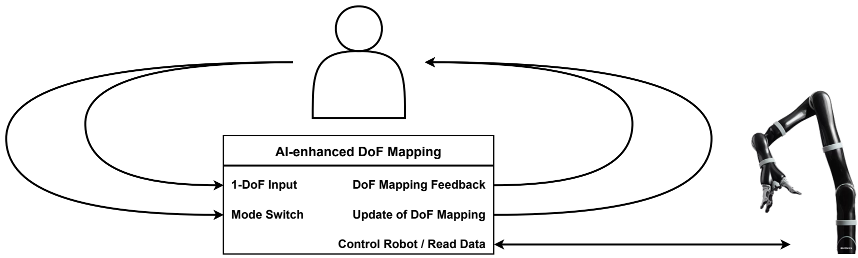


Figure 3.1: ***Overview of the AI-enhanced DoF mapping concept,*** *including the important entities for collaboration: the user, the control system, and the robotic arm.*

The system proposed in this thesis adopts an AI-enhanced DoF mapping concept to enable users to control a high-DoF robotic arm with a single-DoF control input (see Figure 3.1). By generating a combination of several DoFs, the system suggests a mapping for multi-dimensional movements of the robotic arm. This mapping, along with any updates to it, is conveyed to the user through multi-modal feedback. By manipulating a single-DoF input device, users can control the robot along the generated DoF mapping. Additionally, they have the option to perform a mode switch to activate another combination of DoFs. By combining the user input and the activated DoF mapping, the system regulates the robotic arm and ensures the intended robot movement by interpreting its data.

3.1 STEP I: Analysing User Needs and Robot Intent Communication

STEP I explores user needs and concerns towards an AI-enhanced assistive robotic arm, as well as the factors influencing effective robot intent communication. Centred on supporting individuals with motor impairments in ADLs, the first research question (RQ) arises from the recurrent challenges faced by the target demographic, compounded by the diverse spectrum of impairment types and severity levels:

RQ1.1 What are the essential needs of individuals with motor impairments for an assistive robotic arm?

To address RQ1.1, I present the findings of an ethnographic study, resulting in recommendations for the embedding of assistive robotic arms in the users' day-to-day life [P2]. Through interviews and *in-situ* observations conducted in participants' homes, we obtained a comprehensive understanding of user needs regarding their care, thereby giving the community a voice in the research process.

As robots gain greater autonomy, their interactions with humans in shared spaces become more frequent. This necessitates the development of a mutual understanding between humans and their robotic counterparts, as highlighted in research question:

RQ1.2 How can robots effectively communicate their intended movement, and what comprises this intent?

In response to RQ1.2, I propose a *robot motion intent* model based on a scoping review [P5]. Our review seeks to clarify the definition, properties, and relationships of *robot motion intent* within the field of HRI. Furthermore, the study offers an extensive overview of the various types of visualisations, modalities frequently used in communicating *robot motion intent,* along with derived empirical implications, and suggests numerous avenues for future research.

3.1.1 Ethnographic Study of User Needs

Assistive technologies are increasingly recognised as a meaningful addition in domestic care settings, capable of reducing the need for constant human support and providing opportunities for individuals with physical impairments to regain some level of independence [168]. However, studies conducted by Klein [121] and Merkel and Kucharski [153] highlight cases of non-acceptance and non-use. They advocate for a shift towards devices that are better aligned with the needs of the target group. Similarly, Vines et al. recommend involving future users in the developing progress to enhance product acceptance [227].

In our ethnographic study [P2], we explored user perspectives regarding potential systems assisting with drinking and eating tasks. Through interviews and *in-situ* observations involving 15 users with motor impairments, we gained a situational perspective and derived a set of requirements towards assistive technologies. The interviews were structured into four sections, covering participants' living situations, attitudes towards eating and drinking, required levels of assistance, and preferences for an ideal robotic aid. To introduce participants to the concept and to simulate a close-contact setting, we employed a *Google Cardboard* [89] with a stereoscopic video featuring our laboratory robot setup. This lightweight solution proved essential since interviews took place in participants' homes. The video depicted a robotic arm delivering a glass of water to the user's face, with the original sounds of the robotic aid increasing the authenticity of the experience. Focusing on participants' consumption of food and drinks with the assistance of their caregivers, the *in-situ* observations recorded the relative location of the assistant, methods employed, and communication between both parties. Based on our analysis, we derived a set of eleven recommendations for a robotic drinking aid. These recommendations are categorised into three groups addressing structural, social, and collaborative concerns, as illustrated in Figure 3.2.

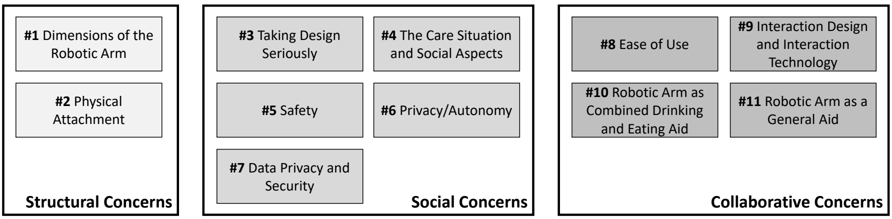


Figure 3.2: ***User derived set of recommendations*** *that support researchers and practitioners in designing assistive robots, categorised into three groups addressing structural, social, and collaborative concerns.*

Consistent with the findings of Fattal et al. [66] and others [153, 216], our results indicate structural concerns expressed by our participants. Consequently, our findings underscore the necessity for recommendations pertaining to the physical characteristics of the robotic arm and attachment site. Participants further emphasised the importance of limiting the total number of assistive systems in their homes. Consequently the ideal versatile robotic aid should perform multiple tasks, supporting many different ADLs. A crucial insight revealed participants' desire to spend time without their assistants, which requires them to complete ADLs on their own with robotic support. This autonomy would promote independent and self-determined living [30], while simultaneously alleviating caregivers and addressing the well-documented caregiver fatigue [114]. Participants particularly valued the inclusion in the design process of a device developed specifically for them. The VR environment of the *Google Cardboard* allowed us to virtually *bring* the robot to the participants and mitigated logistic challenges. This involvement fosters a sense of ownership and ensures that the resulting recommendations are finely attuned to their unique needs. Beyond the individual impact, these recommendations represent a crucial step in bridging the gap between technological design and the nuanced contexts of the target group and – as such – increase acceptance of future assistive technology within the community.

3.1.2 Scoping Review: *Robot Motion Intent*

As robots increasingly operate within shared spaces alongside humans, their growing autonomy – especially in close-contact interactions – underscores the need for a mutual understanding. While robotic research addresses this challenge from sensory and path planning perspectives, as seen in human-aware navigation [126], the field of HRI focuses on enhancing human comprehension of robot behaviour [25, 184, 228]. However, the intricacies of human communication often get lost in this context, requiring an understanding of robotic behaviour from its own frame of reference. Yet, progress in this area is impeded by the lack of a clear definition of *robot motion intent*. The usage of this concept is often ambiguous, encompassing various aspects without consistent definition or specification by researchers. Instead, similar underlying ideas have been explored in the literature under terms such as conveying the inner state of a robot [215], communicating a spatial perception of the outside world [27], or expressing a forthcoming/planned movement, either directly or indirectly [39, 146].

To investigate essential and current themes in robot intent communication, we conducted a scoping review [P5], following a multi-step process in accordance with the *Preferred Reporting Items for Systematic Reviews (PRISMA)* [166] guidelines. The analysis revealed that several papers either presented, combined, or empirically compared multiple intents. Consequently, we systematically extracted all individual intents from the paper corpus (n= 77), resulting in a total of 172 unique intents. Analysing the identified intents, we mapped the primary entities *robot*, *intent*, and *human*, and identified a communication flow among them resembling the HCI model proposed by Schomaker [191]. Reflecting on all entities, our analysis of the intents revolved around: 1) why they were communicated (*goal*), 2) who communicated them (*robot*), 3) what they communicated (*intent*), 4) to whom they were communicated (*human*), and 5) in which circumstances they were communicated (*context*). Dimensions, categories, and properties emerged from the data through an open coding process of the extracted answers. Specifically, we identified the kind of robot, location of intent, type of intent, information of intent, and role of human as our dimensions with the resulting intent communication model shown in Figure 3.3.

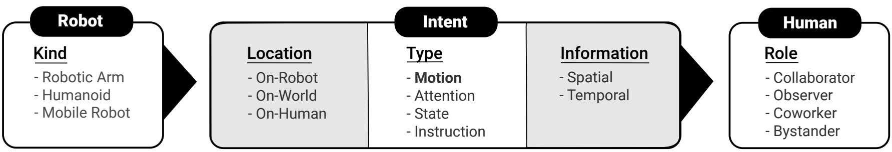


Figure 3.3: ***Overview of the intent communication model from robot to human.*** *The three entities (i.e., robot, intent, human) and their dimensions are derived from our literature corpus [P5]. The flow of communication parallels the human-computer interaction model from Schomaker [191].*

By delineating the dimensions of intent information, we found that the spatial property plays a significant role. Information registered in space establishes a direct connection between real-world objects and displayed information, whereas information unregistered in space lacks this immediate link, requiring an additional mental step to establish the connection. Consequently, conveying information unregistered in space may be less intuitive, prompting researchers to explore various combinations of information to mitigate this challenge. Findings addressing the spatial property indicate that both discrete and continuous information are effective for communicating intent, with the combination of these two forms proving to be the most practical method in conveying robot intent. Additionally, for the intent location dimension, we found that most of the review's corpus leans towards presenting intent information as close as possible to the robot's target to ensure a comprehensive situational understanding by the user. Most importantly, our scoping review emphasises the need for a broader perspective on *robot motion intent*, revealing that it encompasses intent types that may initially appear unrelated to motion. Our analysis identified *attention*, *state*, and *instruction* as crucial elements, often serving as necessary pre- or post-cursors for effective communication of explicit motion intent.

3.1.3 Summary & Next Steps

STEP I embodies two aspects, with the first focusing on gathering information and analysing user needs regarding an assistive robotic arm (RQ1.1), while the second delves into the examination of prior usage of *robot motion intent* in the literature (RQ1.2).

Ensuring the effectiveness of assistive technologies relies on usability for the target group. End users, being the most knowledgeable about their capabilities and needs, play a key role in determining the optimal interaction methods with robotic devices. Hence, considering the user's preferred collaboration mode, whether high-level (e.g., command-based) or manual (e.g., joystick-, button-based), becomes imperative. Equally essential is transparent communication of the current mode to users and their human assistants. By extending our interviews with a stereoscopic video of our laboratory robot setting, we observed that

VR simulation solutions can prevent excessive workload of the possibly vulnerable users during the research process and accelerate prototyping [219]. Furthermore, partially focusing on the research process in simulation – as seen here by using *Google Cardboard* – helps alleviate the challenges posed by the bulky, expensive, and intricate nature of assistive robotic arms. Presenting new interaction and control options becomes much less time-consuming while simultaneously excluding potentially dangerous close-contact situations with users before glitches are managed [P2].

Our scoping review reinforces the demand for a more holistic understanding of *robot motion intent*, which encompasses various intent categories initially seeming unrelated to motion [P5]. However, our analysis uncovers that *attention*, *state*, and *instruction* consistently serve as essential prerequisites either before or after conveying explicit *motion intent*. Tailoring approaches for different types of intent, intent information, and intent location provides an opportunity for multi-modal feedback. This not only informs the user about the robot's intended movement but also offers a chance for timely intervention.

3.2 STEP II: Concept & Design of AI-enhanced Assistive Robotics

In STEP II – guided by the findings of STEP I – I explore shared control applications, analyse challenges associated with human-AI collaboration and derive empirical implications from this. Based on these insights, I introduce the concept of our *ADMC* shared control approach, which generates and suggests DoF mappings to the user. Subsequently, visualisations and a vibrotactile haptic communication approach for these DoF mappings are conceptualised. The step concludes with the introduction and implementation of the *AdaptiX* framework, which facilitates both the development and evaluation of shared control applications.

Several challenges complicate the effective development of shared control approaches, potentially hindering progress if not adequately addressed. Highly autonomous systems [218] reduce the amount of required user interaction but may induce stress [176] and feelings of distrust in users [242]. On the other hand, manual controls [201] can be challenging, or even impossible, due to user motor impairments [43, 112]. Developing successful shared control systems for assistive technologies necessitates extensive experimentation, fine-tuning, and a delicate balance between user and robot control input [130] to reconcile the advantages and disadvantages of the autonomy spectrum extremes. Additionally, there is a noticeable gap in understanding optimal system strategies for specific situations and the diverse user profiles within the target user group, leading to the research questions:

RQ2.1 What considerations are essential when designing shared control applications?

To address RQ2.1, insights from two research projects are presented, as detailed in [P8] and [P4]. The former delineates three main challenges when interacting with an AI-enhanced systems, while the latter introduces our novel shared control approach – *ADMC*.

RQ2.2 How can an intended movement direction by DoF mappings be communicated?

To communicate the recommended movement direction to the user (RQ2.2), I examine various visualisation concepts from our work [P6]. Additionally, I introduce a feedback concept that goes beyond visualisations by utilising vibrotactile haptic feedback, as explored in our *HaptiX* concept and subsequent evaluation [P3].

RQ2.3 How can logistical burdens be reduced and the target group be better involved in the research and evaluation process?

To include the target group during the research and evaluation process while simultaneously reducing logistical burdens associated with *in-home* robot installations and laboratory visits (RQ2.3), I propose a comprehensive XR testbed environment [P4] for the *in-silico* development and evaluation of AI-enhanced shared control approaches and multi-modal feedback methods [P4]. In assistive robotics, the customisation of visualisation, feedback modalities and – maybe most importantly – the selection of suitable input devices [12] is paramount to address individual needs and abilities, recognising the absence of a *one size fits all* solution [105]. *AdaptiX*, featuring a VR simulation environment and extensive customisation options, not only facilitates the integration of the target group into the research process but also reduces operational overhead and enhances overall efficiency. The open source *AdaptiX* framework presented in this thesis is available for use by future researchers.

3.2.1 Human-AI Collaboration in Assistive Robotics

Current HRI research highlights a significant challenge faced by developers: optimising the autonomy level of assistive robots [130]. Striking a balance is crucial, as purely autonomous systems [218] – where users primarily issue high-level commands for the robot to execute – may diminish user interaction and trust. On the other end of the spectrum, manual controls [201], with only minimal alterations to the user's input, could prove impractical for users with specific impairments [117, 176, 242]. Various approaches are currently in use across different settings to achieve a balance, including time-optimal methods [103], blended mode switching [65], shared control templates [178], and body-machine interfaces [107]. These shared control approaches – combining manual input with algorithmic assistance – represent a promising research direction. However, drawing from the relevant literature identified in [P4, P8], we delineated three main challenges when interacting with an AI-enhanced assistive robot [P8]. Addressing these challenges is crucial for establishing the viability of shared control solutions.

AI Legibility: While objectively the robot system is designed to act in the user's best interest, user trust is not guaranteed. Establishing trust necessitates transparency and legibility that users can comprehend. Additionally, users should be able to intervene in the robot's control in the event of errors or incorrect suggestions for interaction. Effective communication of intent also involves capturing or guiding the user's attention, which, in turn, may require employing multi-modal stimuli tailored to the situation and user capabilities.

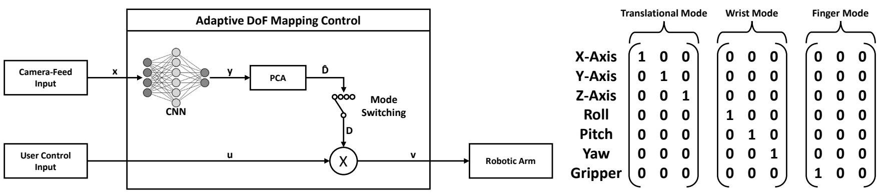
AI User Control: Prioritising user control presents challenges in robot interaction, where the complexity and DoF limitations of available input devices can pose usability issues. To achieve an optimal shared control balance, we propose starting with a minimised set of user interactions and increasing that on demand, depending on the individual capabilities. While optimal ways of accomplishing a goal may require complex intervention from the robot, such interventions may be challenging for users to understand and, therefore, trust. A low DoF of the user input further allows a more extensive selection of specific input devices to be used to control the robot.

AI Intervention: The aim is to keep users in the loop so that they can intervene appropriately whenever the AI reaches its limits. However, avoiding the imposition of sole decision-making on the user is crucial to mitigate cognitive demand and temporal delays. Optimal efficiency is achieved through the implementation of a *four-eye principle*, where the AI operates with implicit user consent until intervention becomes necessary to fulfil the task's goal.

Previous research (e.g., [174]) has demonstrated that pick-and-place tasks are ubiquitously necessary to perform ADLs. Consequently, it is important that shared control applications are implemented first for these straightforward tasks before more complex sequences are examined. If users encounter difficulties in comprehending shared controls for pick-and-place tasks, it is plausible that more intricate tasks could lead to additional frustration. Therefore, we designed and optimised our initial version of *ADMC* to perform pick-and-place tasks with an assistive robotic arm. Our work highlights the benefits and importance of sensible interaction design, which addresses these challenges and requires both a deep understanding of and interconnection with the AI technology.

3.2.2 The *Adaptive DoF Mapping Control* Concept

The original adaptive concept presented by Goldau and Frese involves integrating the primary DoFs of a robotic arm based on the current context and aligning them with a low-DoF input device [84]. This alignment is achieved by attaching a colour-and-depth camera to the robotic arm's gripper and training a CNN using individuals without motor impairments to perform ADLs, akin to the learning-by-demonstration approach used in autonomous robots demonstrated by Canal et al. [36]. In a proof-of-concept study, Goldau and Frese compared the control of a simulated two-dimensional robot using manual controls and CNN-based controls. While their approach showed the successful usage of DoF mappings, it lacks integration into realistic real-world scenarios.



(a) (b)

Figure 3.4: ***Concept of ADMC using a CNN***. ***(a)*** *Control pipeline for proposed adaptive shared control and* ***(b)*** *matrix representation of DoF mappings: columns represent input-DoFs, rows represent output-DoFs, and subsets represent modes [P4]. Two columns were added to represent zero movement mappings in Finger Mode.*

Building on Goldau and Frese's methodology [84], our work [P4] extends their approach from two dimensions to three-dimensional space. This expansion increases the number of potential DoFs, allowing for a more accurate representation of ADLs within our framework. In our adaptive DoF mapping concept – denoted as *ADMC* – the objective is to offer a set of DoF mappings arranged by their effectiveness in executing the pick-and-place task used in our experiments. These DoFs mappings (see Figure 3.4) are suggested by a rule engine (e.g., a CNN or script-based approach). The concept of *usefulness* assumes that maximising the cardinal DoFs of the robot aligned with an input DoF while progressing toward the next goal is the most advantageous approach.

This optimal DoF mapping suggestion is considered the preferred choice, primarily due to a substantial reduction in the need for mode switches when multiple DoFs are consolidated into a single motion. Combining more DoFs – provided it is suitable for the given context – minimises the necessity for mode switches. Consequently, the DoF mappings are organised based on the number of DoFs they combine. Additionally, alongside the *Optimal Suggestion*, a second suggestion is presented, representing an orthogonal variation of the first suggestion. This second option offers the highest degree of variability in spatial direction while keeping the number of combined DoFs unchanged. Users may find this secondary suggestion valuable for adjusting their position while maintaining a sensible orientation toward the next goal. The following DoF mappings were employed:

Optimal Suggestion: Combining translation, rotation, and finger movement (opening and closing) into one suggestion, causing the gripper to move towards the target, pick it up, or release it on the intended surface.

Adjusted Suggestion: An orthogonal suggestion based on *Optimal Suggestion* but excluding the finger movement. Allows the users to adjust the gripper's position while still being correctly orientated.

Translation Suggestion: A pure translation towards the next target, disregarding any rotation.

Rotation Suggestion: A pure rotation towards the next target disregarding any translation.

Gripper Suggestion: Opening or closing of the gripper's fingers.

3.2.3 Concepts to Visualise AI-generated Movements

The objective of achieving a high level of AI legibility revolves around enhancing the understanding of how the AI reassigns input mapping or adjusts the movement trajectory of the robot. In our survey on *robot motion intent* approaches [P5], we observed that, for conveying location information such as movement direction, head-mounted technology such as AR HMDs are effective in visually representing proposed robot movement [48]. Although this research has explored *robot motion intent*, there needs to be more insight into what works best in various situations and for various user demographics. Customising the visualisation and feedback modality is paramount, as there is no *one size fits all* solution [105]. To address this issue, we proposed design concepts spanning a spectrum between two extremes – indicative and explanatory [P6]. Indicative feedback focuses on essential information only, providing a swift and straightforward solution suitable for proficient robot users. In contrast, explanatory feedback entails showing movements in great detail, offering extensive information that particularly benefit novice users.

DoF-Indicator: In this concept, Light Emitting Diodes (LEDs) are attached to the robot's axis and joints communicate active (LED lights up) and inactive (LED does not light up) DoFs. Alternatively, LEDs could be mounted on a bar in front of the user and referring to each joint by the corresponding number (e.g., joint 1 – 7). Users of the system must derive the resulting movement direction based on the active DoFs. Consequently, this form of visualisation is likely better suited for experienced users.

DoF-Combination-Indicator: Here, DoF mappings are communicated through a simplified representation of the robot itself, capable of executing movements in only two DoFs simultaneously, such as rotating and extending the arm. This approach reduces the complexity of the robot and enhances the user's comprehension of the intended movement. As an AR representation, it either can be displayed separately in the corner of the AR screen or overlays the actual robot, further decreasing the robot's complexity.

Gizmo Visualisation: With this visualisation approach, gizmos – arrows, planes and point clouds – convey the robot's current movement capabilities. Planes represent a two-dimensional representation of the intended movement possibility (e.g., x/y or x/z plane). Point clouds extend this by a third dimension, enabling a visualisation of a (complex) three-dimensional space surrounding the robot. Alternatively, arrows indicate the respective movement directions based on the DoF mapping and even provide a forward and backward direction in terms of controlling the robot. Arrows may vary in shape, being either straight or curved, depending on the complexity of the DoF mapping.

Demonstration: Current movement possibilities are demonstrated either through the physical robot itself or an AR ghost representation. In both cases, rapid movement signifies the intended motion.

Arrows – as one of the simplest gizmo visualisations – are commonly employed to represent movement directions, both in HRI and everyday applications. Previous studies have demonstrated their effectiveness in visualising the *robot motion intent* of a *Baxter* robotic arm in AR [186], as well as a mobile robot's trajectory in SAR [39, 104, 145]. Findings by Zein et al. show that an arrow-based MR visualisation leads to a significantly lower workload compared to either auto-completed trajectories without a visualisation or traditional teleoperation [241]. This emphasises the suitability of arrow-based visualisations to communicated the intended movement directions of the *ADMC* approach.

Beyond Visualisation for AI-generated Movements

Human perception of objects in their environment predominantly relies on the sense of sight. However, situations may arise where this ability can be impaired or entirely unavailable. These circumstances may include objects being obscured by other items or User Interface (UI) elements (i.e., visual clutter) or being positioned outside the individual's field of view. Additionally, optical perception may be limited or impossible for vision-impaired individuals. In addressing this, Burke et al. demonstrated that the haptic modality can partially compensate for the absence of visual information and, in certain instances, outperform audio-based cues [31]. This type of feedback can be particularly valuable when the visual channel is overwhelmed by distracting information [42, 115]. It can also be effectively combined with other sensory modalities.

Building on these insights, our research [P3] investigated various design approaches for conveying three-dimensional directional cues using vibrotactile feedback. The study involved the development of two conditions based on the *Cutaneous Rabbit* [80] illusion and one condition based on *Apparent Tactile Motion* [33] to communicate two-dimensional directions. The gradient of the overall three-dimensional direction was subsequently encoded using methods such as the number of discrete vibration pulses, vibration intensity, or a combination of both. Our findings demonstrate that users can understand three-dimensional directional cues and associate them with forthcoming movements. This approach can be employed to map AI-generated movement suggestions onto vibration input on the hand to improve accessibility. Variations in the intensity of the actuators can indicate the degree of directional change, allowing users to better visualise the generated trajectory.

3.2.4 The *AdaptiX* Framework as a Research and Evaluation Tool

AdaptiX [P4] facilitates the development and evaluation of shared control applications in a high-resolution transitional MR environment. The XR framework incorporates a VR simulation environment featuring a virtual robotic arm (e.g., a *Kinova Jaco 2*) and offers extensive customisation options. This *in-silico* approach streamlines the research process while simultaneously reducing overhead and increasing efficiency. An overview of the framework's architecture is presented in Figure 3.5.

In addition to Cartesian robot control, our study incorporates our proposed *ADMC* concept as the standard shared control approach. *ADMC* operates on suggestions generated by a rule engine (e.g., a CNN or script-based approach) for user control. The script-based approach is particularly valuable as

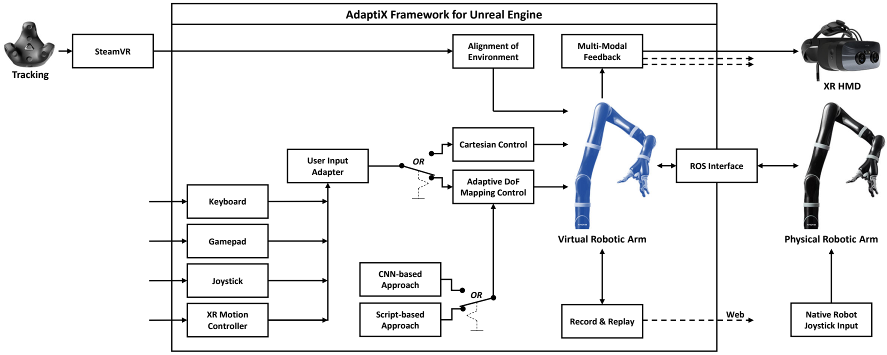


Figure 3.5: ***Overview of the AdaptiX architecture****, illustrating each component, their directional communication, and the crossover from and to the framework [P4]. The user input is either used for Cartesian Control or Adaptive DoF Mapping Control (ADMC). For ADMC, either a CNN-based or script-based rule engine can be selected.*

it helps mitigate potential biases that may arise from more generic methods like CNN-based controls, which are currently still limited in scope [41, 77, 133]. Integrated directly into the *Unreal Engine* [61], our *ADMC* concept allows researchers and developers to fully customise control methods, system behaviour, and feedback techniques using *C++* or *Blueprints*.

AdaptiX supports various pre-implemented input devices and offers an adaptor class for streamlining the development and implementation of additional input devices. This adaptability allows researchers and developers to effortlessly incorporate their own innovative ideas and concepts. The integration of a ROS interface within *AdaptiX* facilitates seamless connections to non-simulated physical robotic arms, thereby enabling bidirectional interactions and data exchange through a *DigitalTwin* and *PhysicalTwin* approach. Straightforward trajectory programming is made possible by manually guiding the Tool Center Point (TCP) of a simulated or physical robotic arm to a desired location and recording both its position and orientation for future replay. The system further provides customisation options by allowing adjustments to specific details such as camera positions and background scenes, thus creating a highly customisable environment. *AdaptiX* enables the exploitation of the entire continuum of MR. This extends the use of the transitional framework to new scenarios and environments – including the real world. Consequently, the virtual and real environments of the robotic arm are aligned, allowing researchers to seamlessly switch between the user controlling the real and virtual robot. The level of MR can be adjusted in various steps (cf. the *virtuality continuum* of Milgram and Kishino [156]). The MR environment setups include:

1. the completely real environment with the real robotic arm,
2. the real environment extended with visual cues,
3. the real environment into which the virtual robot is transferred and displayed (with and without visual cues),
4. the virtual environment into which the real robot is transferred and displayed (with and without visual cues),
5. the completely virtual environment with the virtual robotic arm.

Integrating the modular and extendable *AdaptiX* framework provides a comprehensive foundation for developing novel interaction designs and feedback methods for shared control applications. *AdaptiX* offers advantages in both remote and in-person studies, eliminating the need for a physical robotic device during initial ideation and prototyping stages, thereby enhancing flexibility, accessibility, and efficiency. Additionally, it streamlines the research process by obviating the need for researchers to start from scratch when implementing their individual solutions.

3.2.5 Summary & Next Steps

STEP II outlines the conceptual development of AI-enhanced assistive robotics, which includes the exploration of human-AI collaboration [P8], the application of shared control using *ADMC* [P4], and the development of visual [P6] as well as vibrotactile haptic [P3] feedback methods to convey *robot motion intent*. These efforts result in a XR framework for the research process [P4] of shared control applications in HRI.

Considering the primary challenges associated with collaboration between humans and AI-enhanced systems (RQ2.1), we devised and developed a shared control approach (*ADMC*) that priorities user control while facilitating seamless intervention strategies and provides legible DoF mapping suggestions through multi-modal feedback methods (RQ2.2). The objective of the *ADMC* approach is to present a set of DoF mappings ordered based on their effectiveness in accomplishing a pick-and-place task, which are integral to many ADLs. In this context, a *useful* mapping maximises the cardinal DoFs of the robot assigned to a low input-DoF. This shared control application enables users to control a complex robotic arm with a reduced input DoF. Based on an initial training of the CNN, the system is not limited to cardinal DoFs or pre-determined motions of an autonomous system. However, it retains the capability to accurately represent and execute those, while also adapting to a wider range of movements and behaviours. Therefore, the system provides a DoF mapping that adapts to the current environmental and situational HRC task goal.

One significant aspect of predicting robot behaviour is understanding its motion intent and comprehending how it *conceptualises* its actions [P5]. To address this, we focused on visual solutions that communicate the robots AI-generated motion intent to a human collaborator. Our approach entails a design exploration employing various visualisation techniques to optimise user understanding of DoF mapping effect, ideally resulting in increased safety and fostering end user acceptance. We offer different sets of visualisations tailored for both novice and experienced users.

Furthermore, we introduce a solution to evaluate novel interaction techniques and feedback methods with the target group without physically transporting the robot to the users' homes or study participants to the laboratory. This approach mitigates additional logistical burdens and ensures the involvement of the target group into multiple steps of the development process (RQ2.3). Additionally, future researchers stand to benefit from using *AdaptiX*. They can leverage its capabilities across the entire spectrum of MR and choose to develop a fully immersive VR environment, basic AR visual cues or even a desktop application using a pure screen space. With the simulated and real-world environments of the robotic arm perfectly aligned, nearly seamless switching between controlling the real and virtual robot is possible and can be adjusted during run time to provide optimal interaction solutions.

3.3 STEP III: Evaluating Interaction Design & Robot Intent Communication

The final step of this thesis encompasses an empirical concept evaluation of perceptual feedback and *ADMC* within AI-enhanced assistive HRI. The initial paper [P7] focuses on communicating robot perception during autonomous tasks within a care environment. Especially in this context, promptly relaying any identified objects and obstacles to the user is crucial for enhancing trust in the assistive technology system. Furthermore, the exploration and assessment of the *ADMC* concept are detailed across three papers: two VR simulation studies – one remote [P1] and one in-lab [P9] – and one MR in-lab study involving the control of a physical assistive robotic arm [P10]. Revisions made between these user studies prompted evaluations of various updated versions of *AdaptiX*.

As such, STEP III explores concepts of AI-enhanced assistive HRI, evaluated within the *AdaptiX* framework along these research questions:

RQ3.1 How effectively can an autonomous robotic system communicate its perception, and how well can users recognise perception errors?

I addressed RQ3.1 by creating a scenario in which the gripper of an autonomous robotic arm navigates through a set of obstacles, placed on a table [P7]. I compared three visualisation concepts to assess the user's ability to recognise obstacles as perceived by the robot.

RQ3.2 How efficient is the *ADMC* approach in three-dimensional space with a two-DoF input device for moving an assistive robotic arm?

To address RQ3.2, we assessed the *ADMC* approach in a remote VR study, wherein participants performed pick-and-place tasks, comparing an adaptive approach with a non-adaptive control method [P1]. Participants were tasked with controlling the robotic arm along the suggested DoF mapping with one joystick-axis and the alternative DoF mapping with the second axis.

RQ3.3 What is the efficiency of *ADMC* when refining the interaction design and use a one-DoF input device for moving the robotic arm?

Building on the experiences and results of our initial user study [P1], we optimised our control concept to an one-dimensional input (RQ3.3). Additionally, we evaluated different time points for communicating DoF mapping suggestions to users in a VR laboratory study [P9], as well as different input devices for controlling a physical robot [P10].

3.3.1 Visual Perception Feedback for AI-enhanced Assistive Robotics

As a precursor to the *AdaptiX* framework [P4], we developed a three-dimensional testbed environment to facilitate studies for (semi-)autonomous HRI in close-contact scenarios [P7]. The reported studies [P7] explored various visualisation approaches to effectively convey the robotic arm's perception, especially information regarding detected objects within its physical environment. This perception communication is essential as any failure in object detection may harm the user, leading to unintended incidents such as knocking over items or causing damage during interaction. We applied this concept to a breakfast scenario, where a robot assists in tasks like picking up a bottle and pouring water into a glass.

Here, we investigated three different SAR visualisations for robot perception [P7]. Although primarily confined to two dimensions, research in the context of motion intent has demonstrated that SAR can be adapted to cover a dynamic workspace that includes multiple surface areas [9, 38]. In our scenario, this adaptation is relevant for interacting with objects both on a table and retrieving objects from shelves. SAR has the capacity to augment larger areas of the surroundings, potentially extending beyond the user's physical field of view. Unlike HMDs, SAR can also be viewed by secondary users. However, the achievable field of view depends on the mounting position of the projection system. While SAR may expand the visible augmentation area, it still faces challenges in effectively communicating information about objects that are off-screen, a challenge that we addressed in our research [P7] by investigating the effectiveness on-screen and off-screen components of the visualisations. In two web-based remote studies – one with the target group and one with the general public – we compared the efficiency of the visualisations when the robot fails to recognise an object, such as when the object leaves the sensor coverage, resulting in the deactivation of the corresponding visualisation. Both quantitative and qualitative findings underscored the significance of an easily comprehensible visualisation (e.g., *Line* [P7]). We use this approach for conveying the motion intent of the assistive robotic arm, employing straightforward gizmo visualisations, such as arrows [P6], in our *AdaptiX* framework [P4].

3.3.2 Visualising Adaptive DoF Mappings in 3D Space

In our initial study [P1], we used the *AdaptiX* framework to investigate the proposed *ADMC* control method alongside associated visual cues for different DoF mappings. The study aimed to assess the performance of the novel adaptive control method – adapted from Goldau and Frese [84] – within a three-dimensional environment in comparison to the standard mode-switch approach featuring cardinal DoF mappings. Additionally, the study explored the impact of variations in the appearance of visual cues on the performance of the adaptive control method. Due to the at-the-time ongoing *COVID-19* pandemic [233], the research was conducted remotely within a VR environment created by the *AdaptiX* framework. Participants without specific backgrounds were recruited, provided they had access to the necessary hardware for an immersive experience, such as an *Meta Quest 2* [154] VR HMD.

During the study, participants were tasked with repeatedly executing a simple pick-and-place operation by controlling a virtual *Kinova Jaco 2* using one of three control types. These control types included the *Classic* visualisation, a method based on *Double Arrow* cues employing two arrows attached to the gripper's fingers, and a visually simplified version called *Single Arrow*, which utilised only one arrow positioned in the middle of the gripper. Comparative results revealed that adaptive controls required significantly fewer mode switches than the classic control methods. However, there were no notable improvements in task completion time or perceived workload. Participants in the study also expressed concerns about the dynamically changing mapping of combined DoFs and the use of a two-DoF input device.

3.3.3 Single-DoF Control Input and Multi-Modal Feedback

In a subsequent study [P9], we assessed two new adaptive control methods for an assistive robotic arm, one of which incorporated a multi-modal approach for guiding the user's attention. The study was conducted in a laboratory setting to corroborate or challenge the initial investigation findings [P1] regarding participants' interaction with the assistive robotic arm using the *AdaptiX* framework. In this adaptive system, continuous calculations determined the optimal mapping of DoFs for task completion while in motion. These calculations were presented to users as alternative control options during the task. Users could cycle through suggestions by pressing a button on the input device or continue with the current DoFs. The study compared two variations, namely *Continuous* and *Threshold*, which differed in the timing of when suggestions were presented to the user. These were contrasted against a non-adaptive *Classic* control method, examining effects on task completion time, mode switches, perceived workload, and user opinions. Results indicated comparable performance between *Continuous* and *Threshold* in terms of quantitative measures and qualitative insights, suggesting both methods effectively conveyed directional cues to users.

3.3.4 Different Input Devices for *ADMC*

A third study [P10] highlights the MR capabilities of the *AdaptiX* framework and its integration potential with various input devices. Here, we employed the *Varjo XR-3* [224] XR HMD to explore an interaction design and feedback technique akin to the *Threshold* approach [P9]. By incorporating the XR HMD, the prototype creates an AR environment for the user, augmenting the physical setup with visual cues. Unlike the previous study [P9] involving a virtual pick-and-place task, this setup incorporates physical objects, a physical drop area, and a physical robotic arm with AR cues delivered through the headset. Participants compared three assistive input techniques: 1) a head-based control method utilising head deflection on the pitch axis for continuous input and on the roll axis for mode-switching, 2) a gamepad input using the *Xbox Adaptive Controller* [155] extended with *Logitech Adaptive Gaming Kit* [135] buttons for discrete input, and 3) the control-stick of a *Joy-Con* [162] motion controller, serving as a baseline in comparison to our previous approach [P9].

Both of our selected input methods – *Joy-Con* and *Logitech Adaptive Gaming Kit* buttons – show promise as effective approaches for controlling a robotic arm within a shared control application. Our research findings indicate that both hand-operated input methods, whether providing discrete or continuous input data, offer several advantages: first, they reduce the perceived workload of the user, and second, they enhance user perceptions of *perceived usefulness*, *perceived ease of use*, *emotions*, and *comfort* during interaction with the robotic arm. This suggests that hand-operated input methods can improve the user experience and usability of shared control applications for robotic arms.

3.3.5 Summary & Future Work

To address RQ3.1, our research on perceptional feedback visualisations [P7] demonstrated the advantages of straightforward and easy-to-understand visualisations. We applied this approach to communicate motion intent in subsequent research studies [P1, P9, P10]. The *Line* visualisation, preferred in both studies [P7], has not been integrated as a perceptional feedback visualisation into *AdaptiX* yet, but it is planned for future versions.

Our remote study [P1] using *AdaptiX* proved to be effective in evaluating new interaction designs and feedback techniques. A notable advantage of this approach is that the physical robotic device does not need to be physically present during these initial studies when testing and assessing critical design components. We showed that adaptive mappings of the robot's DoFs – *ADMC* – can lead to a significantly lower number of mode switches compared to standard control methods. However, our study could not conclusively show improvements in task completion time or reduced cognitive load. Also, challenges concerning the understanding of DoF mappings resulting from the two-dimensional input device were raised during the study (RQ3.2).

The integrated multi-modal feedback [P9] is a crucial feature of *AdaptiX*, capable of supporting real-time suggestions by guiding user attention. Regarding RQ3.3, we examined possible effects for our single-DoF controlled *ADMC* compared to *Classic* on task completion time, number of necessary mode switches, perceived workload, and subjective user experience. Although some participants found the combined visual-auditory-haptic multi-modal feedback to be "irritating" [P9], it effectively conveyed updated suggestions. In contrast to our previous study [P1], we show the significant efficiency to adaptive control methods compared to non-adaptive approaches. Future studies may continue exploring diverse input and feedback modalities along with corresponding user contexts using *AdaptiX* as a virtual simulation framework to enabling the seamless setup of further user research.

By integrating virtual cues into a real-world setting [P10], the research moved closer to reality on the MR-continuum than the previous two case studies [P1, P9]. Here, *AdaptiX* proved to be a successful and user-friendly interface bridging the gap between physical robot control and communication via an XR HMD. One key advantage of *AdaptiX* is its ability to quickly evaluate different input devices' efficiency in controlling the robotic arm within the context of adaptive DoF mappings. The standardised *User Input Adapter* offers researchers the flexibility in choosing from various technologies, supporting continuous, discrete, and absolute user input methods. Moreover, its modular nature allows for further customisation to meet specific research needs.

I invite the research community to further enhance the *AdaptiX* framework according to their specific contexts and needs. Researchers are encouraged to create custom levels or scenarios and integrate new interfaces into the framework. This collaborative approach can contribute to the continued development and versatility of *AdaptiX*, expanding its potential applications in the field of AI-enhanced assistive HRC.

4

DISCUSSION

In *Chapter Three*, I presented ten research papers and delineated their contributions to the development and evaluation of a shared control approach in AI-enhanced assistive robotics. Building on this foundation, *Chapter Four* contextualises the significance of this work within the broader landscape of HRI.

The chapter commences by reflecting on the overarching research approach, along the three guiding questions: **Q1**) how can legible AI-enhanced assistive robots be effectively integrated in domestic care settings while accounting for subjective needs, **Q2**) how can collaborative efforts with an AI-enhanced system facilitated by reducing input complexity, and **Q3**) how does the *AdaptiX* framework support researcher in developing and evaluating *ADMC* and other shared control approaches using low-DoF input devices? Subsequently, the discussion explores two fundamental challenges within HRI that are addressed by this work: first, enhancing the likelihood of successful human-centred design through strategic integration of the target group, and second, establishing legible control systems, which is particularly vital with a target group that already relies on others for a significant portion of their daily activities.

4.1 Critical Reflection of the Research Approach

Predefined key elements have shaped the overarching research approach and design choices outlined in this thesis, thereby influencing its ultimate outcomes. This includes the central focus on designing legible AI-enhanced control approaches for assistive robotic arms for people with motor impairments in a user-centred design approach, focusing specifically on pick-and-place tasks-based ADLs. To address **Q1**, STEP I reports an ethnographic study with the target demographic to explore enhancing daily routines with assistive systems. Drawing on insights from prior studies [189, 226] highlighting low acceptance rates and underutilisation of assistive devices, this thesis prioritises designing based on expressed user needs. While the *in-situ* home visits offered a more realistic depiction of the lived realities of the target demographic than laboratory settings, it is important to note the limited geographic distribution of study participants, potentially impacting generalisability of the reported insights. Furthermore, understanding the robot's behaviour is crucial for successful HRC, especially for non-tech-savvy users unfamiliar with complex robotic arms and close interactions [124, P5]. By facilitating effective communication of *robot motion intent*, I sought to increase understanding and predictability towards the system and create legible AI–enhanced assistive robots for our target user group. To communicate directional cues – based on the *ADMC*'s DoF mapping – we chose XR HMDs, as they significantly improve users' understanding of *robot motion intent* [P5]. Even though current HMD-technology is not able to provide an immersive XR experience efficiently (due to, e.g., size, weight, battery), research can and should use these concepts [37, 223]. Bulky XR HMDs are particularly unsuitable for individuals with motor impairments, especially given the varied – and in some cases progressively deteriorating – physical capabilities. As such, current XR HMDs present a *one size fits all* approach that does not address these individual circumstances [11, 12, 105]. Consequently, the proposed *ADMC* interaction design concept and multi-modal feedback modalities were not evaluated with the target demographic in STEP III, leading to results that should be interpreted relative to each other rather than in absolute terms. Nevertheless, evaluating the *ADMC* concept in a XR environment still provides valuable insights that can be extrapolated to the target group. Study participants remained seated, simulating the position of being in a wheelchair. They also used low-DoF input modalities akin to those used by individuals with motor impairments to control their assistive devices. Looking ahead, future research should prioritise including intended users in individual case studies to gain a more comprehensive understanding of potential challenges and validate or contest the insights gained in [P9] and [P10].

In addressing **Q2**, the proposed shared control *ADMC* approach introduced in STEP II solely focuses on the concept of combining a robot's DoFs to generate distinct DoF mappings [P4, P8]. These mappings, based on an underlying rule engine (e.g., CNN or script-based approach), are communicated to the user as three-dimensional directional cues [P6]. By reducing input complexity through this DoF mapping, users can control a high-DoF robotic arm via a low-DoF input device. This approach builds on Goldau and Frese's proposed combination of specific robot DoFs [84], but extends it from two-dimensional to three-dimensional space and incorporating environmental and situational awareness, facilitating its use for pick-and-place tasks-based ADLs. In contrast, existing literature has primarily focused on goal-oriented shared control applications [7, 91, 178, 237]. In such systems, the goal-oriented approach involves detecting or determining the intended movement goal, computing a corresponding trajectory, and incorporating user input to keep the user in the loop. Contrary to this, the *ADMC* concept does not work with a fixed goal location. Instead, it utilises its inherent rule engine, the arm's posture, and a colour-and-depth camera feeds of the robot's gripper surrounding. This approach enables a more flexible and situationally aware shared control system, integrating human cognitive skills (i.e., intelligence, flexibility, responsiveness) with the robot's precision for successful HRI [4].

To evaluate the interaction design of our *ADMC* concept **(Q3)**, we used a task-specific script [P1, P9, P10] as the underlying rule engine to provide DoF mappings. To ensure a more deterministic study setting and mitigate potential biases, we decided against a more holistic but currently limited method like a CNN-based control [41, 77, 133]. However, it is important to acknowledge that our task-specific script proves effective solely within a controlled experimental environment, where the positions and rotations of all relevant objects and obstacles are factored in. Nevertheless, from a user perspective, the system still proposes situationally aware DoF mappings and provides opportunities to adjust them, whether based on AI or a predefined task-specific script. Thus, I recognise that our studies may not comprehensively assess the quality of *ADMC* when employing a CNN as the rule engine.

Due to the then-ongoing *COVID-19* pandemic, we conducted a remote VR study [P1] to evaluate the initial iteration of our *ADMC* approach. Despite being unsupervised due to the pandemic's constraints – and influenced by the additional participants' effort to download and install the study environment on their hardware – I consider our study setup reasonable under the given circumstances and capable of yielding reliable results. Additionally, we conducted supervised laboratory studies in [P9] and in [P10] to evaluate our *ADMC* approach, confirming significant improvements in adaptive control methods over non-adaptive ones. In our second user study [P9], the lack of measurable differences between different time-based communication methods for feedforward recommendations – *Continuous* and *Threshold* – suggests that both discrete and continuous communication of movement suggestions enable efficient usage of adaptive control methods. While participants overall expressed positivity towards adaptive control methods, individual preferences varied considerably between both approaches. Some participants favoured the higher level of control afforded by *Continuous*, while others preferred the comfort provided by *Threshold*. Consequently, future development of adaptive control methods should – in accordance with Burkolter et al. [32] – incorporate personalisation options to enhance comfort, end user acceptance, and accommodate any future changes of capabilities.

Customisation would be particularly advantageous for *Threshold*-based controls, as several participants expressed irritation with the haptic and audio signals. Allowing users to adjust modalities, signal intensity, and even the threshold itself may enhance usability while retaining the benefits of an adaptive method. Furthermore, when evaluating different input modalities [P10], we observed a significant increase in perceived workload [100] during the head-based interaction. This was substantiated during post-interviews, with participants reporting mental demand and difficulties aligning their actions with the suggested arrow directions. We also noticed that the *Varjo XR-3* HMD might be too bulky and heavy to be used for head-based interaction – especially for precise interactions that take a longer time. Moreover, we observed that the specific initial placement of objects were perceived as disadvantageous compared to others, as the robot is fixed in place and has to perform unlegible movements for novice users to reach the target. However, further experimental studies are needed to determine which factors shape personal preferences and how customisation or crossover methods can deliver the best results.

4.2 Design of Human-centred AI

Despite the increasing prevalence of automation facilitated by AI across various domains, such as autonomous driving [108], content curation [81], and financial decision-making [82], it is imperative to recognise the crucial role humans play in these systems. AI's long-term success and effectiveness is contingent on acknowledging that human agency is critical in its design, implementation, and use. As such, the concept of Human-centred AI (HCAI) focuses on creating AI systems that amplify and augment rather than substitute human abilities [199], which is conceivably particularly valued by a user group that already relies on assistance for many aspects of their lives. HCAI seeks to preserve human control in a way that ensures AI meets human needs while simultaneously allowing users to increase autonomy, retain agency, and remain in the loop regarding their care.

Research by Stephanidis et al. [208], and Ozmen Garibay et al. [165] represents a significant step towards human-centred design in assistive technologies. They emphasise involving users as collaborators rather than mere recipients of technology, recognising their unique insights into their needs, capabilities, and lived circumstances. At the core of this approach lies a collaborative, interdisciplinary methodology that prioritises the direct involvement of users and co-design processes. Echoing their sentiments, the WHO recommends viewing users as active collaborators in the delivery of assistive technology services rather than passive recipients [235]. Similarly, the WHO emphasises that individuals who use assistive technology generally possess valuable insights into their unique needs and circumstances. Adopting this approach ensures that assistive technologies not only promote accessibility and the enhancement of physical capabilities but also fundamentally improve mental well-being and self-determined living. Additionally, Stephanidis and Salvendy's *Design for All* approach in the field of HCI integrates human-centred design principles with accessibility and assistive technologies, including *Universal Design* principles for physical products and constructed spaces [207]. This approach prioritises users as central to the interaction design process, particularly focusing on individuals with disabilities. Following these principles, our user-centred co-design process, beginning with our ethnographic study [P2], allows for a deeper understanding of our target demographic's nuanced needs, concerns, and daily challenges.

Further, when aiming to adhere to the foundational principles of Stephanidis and Salvendy [207], it is imperative to assess the practical implications of adopting human-centred approaches in the design, development, and deployment of assistive technologies [167, 202]. This involves critically examining how such methodologies have already been implemented in real-world scenarios and their impact on user satisfaction and overall effectiveness. As such, evaluating case studies where collaborative, interdisciplinary methodologies have led to innovative solutions can offer valuable insights into best practices and potential pitfalls [129, 164]. Allowing an easy evaluation of case studies, the *AdaptiX* framework [P4] supports HRI research by facilitating an evaluation of assistive robots' interaction design *in-silico* to ensure the effectiveness and acceptance of these robotic aids already in the early research stages and iterative prototype creation by shifting the burden away from study participants to the researcher while reducing the total burden. User feedback gathered in these early evaluations is essential to assess the system's requirements to ensure they meet the user's needs. When then progressing to physical prototyping, the shared control approach, interaction design, and feedback methods can be directly transferred into real-world scenarios, mitigating the need and researcher's effort of starting from scratch.

Moreover, the role of emerging technologies such as AI and the Internet of Things (IoT) in enhancing the adaptability and personalisation of assistive devices offers another promising area for exploration [73, 158, 163], for example, by using wearable and environmental sensors to provide professional assistance services [72]. However, the integration of AI into assistive technologies also raises ethical considerations around privacy, data security, and the potential for increasing the digital divide [60, 63, 150]. In our *ADMC* shared control approach [P4], the posture data of the robotic arm as well as – when using the CNN as underlying rule engine – the feed of a colour-and-depth camera are integrated. This data is used to suggests DoF mappings based on the environmental and situational contexts, focusing solely on objects and obstacles to provide useful mapping suggestions. Recordings within this CNN approach are not stored and only used as live input, mitigating the threat of privacy issues. Examining these ethical questions is vital for ensuring that technological advancements align with the fundamental values of equity, inclusively, and human dignity [110, 221]. And, while the integration of human-centred design principles marks a significant advancement in the field of assistive technologies, it is only the beginning [99, 164]. The ongoing dialogue between researchers, practitioners, users, and policymakers must continue to address both challenges and opportunities brought by technological innovation. Future research should focus on developing scalable, sustainable models for co-design that prioritise the diverse needs and capabilities of all users, ultimately leading to more accessible, effective, and empowering assistive technologies.

4.3 The Human in the Loop

A prevalent challenge in assistive robotics is identifying effective methodologies and technologies for controlling such robots [232]. Devices in this category often have a large number of DoFs, rendering them complex in terms of maneuverability [127]. For instance, an assistive robotic arm equipped with a basic gripping mechanism can execute movements within a three-dimensional space, including translational motions along Cartesian coordinates and rotational movements such as yaw, pitch, and roll [5, 211], typically comprising between five to seven DoFs [194]. Conventional control interfaces, like joysticks, are limited to managing two DoFs. To control a high-DoF robotic device with a lower-DoF input device, a strategy called mode switching is employed [136]. However, this approach requires users to select a specific mode, temporarily disregarding other DoFs. Although high-DoF input devices exist, their accessibility remains limited for individuals with motor disabilities [58, 113, 183]. In their exploration of HRI mechanisms utilising conventional button-based mode switching systems, Herlant et al. showed that upwards of one-sixth of the total operational time was spent on mode alteration [103]. Their research substantiated the premise that the implementation of automatic mode switching enhances user satisfaction [103], particularly within deterministic simulation environments and when objectives are clearly defined [220].

For supporting users, shared control applications merge manual user input and algorithmic assistance, addressing the limitations of fully autonomous or manual systems [1]. They promote a cooperative interaction between the user and the robot, significantly aiding individuals with motor impairments in participating more actively during day-to-day tasks. Consequently, this strategy enhances independence and usability compared to purely manual controls [70]. Providing a comprehensive overview, Flemisch et al. proposed the relationship between shared control, shared and cooperative guidance and control, and human–machine cooperation, with increasing autonomy of the system [69, 70]. As shown, human–machine cooperation and shared and cooperative guidance support on a strategic (e.g., navigation) and tactical (e.g., guidance) level, requiring reduced user interaction and therefore leading to elicit stress [176] and feelings of distrust [242] among users. Although the system is designed to act in the user's best interest, addressing this issue requires the user to build trust, necessitating transparency and legibility that users can comprehend [P8]. By facilitating legible interactions and controls, users can intervene if the shared control application makes a mistake or gives inappropriate suggestions for interaction. Communicating robot intent further requires directing the user's attention, eventually necessitating multimodal stimuli, depending on the situation and the user's capabilities [P5]. By maintaining a balance between autonomy and user involvement, shared control systems – functioning on the operational level [70] and keeping the user directly in the loop – can provide a more acceptable and comfortable experience for individuals relying on assistive technologies [90, 181, 220], something that is particularly important when designing for people that already have to rely on others for heightened support.

Based on these findings, our *ADMC* shared control approach aims to increase the users' sense of *being in control* – as confirmed by participants in [P9], by including user input directly into the outcome. *ADMC* suggests mappings of several robot DoFs to the user based on effectiveness and usefulness in executing the current task and situational environment [P4]. The concept of *use-*

fulness assumes that maximising the cardinal DoFs of the robot aligned with an input DoF while progressing toward the next goal is the most advantageous approach. The user controls the robot along the resulting DoF mapping movement direction with a classic two-dimensional input device (e.g., a joystick), with one input axis controlling robot movement and the second performing a mode switch action to choose from different DoF mappings. Evaluations demonstrate that our *ADMC* approach significantly reduces task completion time [P9], the average number of necessary mode switches [P1, P9], and the perceived workload [P9] compared to a non-adaptive input method. Moreover, qualitative insights revealed that non-adaptive *Classic* control method could still be a valuable addition in specific situations when *ADMC* suggestions did not match their expectations, confirming the the users' preferences for more manual shared control approaches rather than autonomous aids.

5
CONCLUSION

The old computing was about what computers could do; the new computing is about what users can do. Successful technologies are those that are in harmony with users' needs. They must support relationships and activities that enrich the users' experiences.

– B. Shneiderman [198]

Figure 5.1: ***Grasping an object from a shelf by using ADMC***. *The interaction and feedback methods for our AI-enhanced assistive robot were evaluated with the target group. Participants were tasked with picking up various objects from a shelf and placing them into a basket on a table in front of them. The ADMC approach was assessed using three different input devices: head-based controls, assistive buttons, and a joystick. © Matthias Kraus, masapido Filmproduktion*

5.1 Summary of Research Contributions

Assistive technologies can have a transformative impact on individuals with motor impairments by fostering autonomy, facilitating independent living, and reducing the reliance on human caregivers. With an ageing population presenting additional substantial challenges, the potential benefits of these technologies are becoming increasingly apparent. However, the development of such technologies often neglects the perspectives and insights of the very individuals they aim to assist. Adopting a user-centred design approach can substantially improve the acceptance and effectiveness of new assistive technologies by incorporating direct feedback from users regarding their specific needs and capabilities. This includes striking a balance between autonomous actions by assistive aids and potentially demanding manual controls, thus harmonising the respective advantages of both extremes of the assistive technology autonomy spectrum. This shared control approach ensures that users are optimally supported while retaining control over their care to the greatest extent possible.

In this context, my thesis presents an interaction design approach for AI-enhanced assistive HRI. As such, my work began with the preliminary stages before the research process, mainly focusing on two key perspectives: first, presenting recommendations for an assistive robotic drinking aid based on extensive user feedback and *in-situ* observations, and second, introducing a robot intent communication model. Both the user-derived recommendations and the intent communication model are intended for use by future researchers in designing and developing accessible, acceptable, and legible robotic aids. Stepping into the initial stages of the research process, this work extends the concept of shared control from two dimensions [84] into three-dimensional space. Our *ADMC* concept increases the DoFs available for controlling a robotic arm, resulting in a more precise representation of ADLs. This approach addresses the limitations posed by complex multi-dimensional input devices and time-consuming mode switches for individuals with motor impairments [140, 177].

To communicate the resulting movement direction, visual and haptic feedback methods – including a gizmo visualisation technique – have been developed and integrated into the transitional XR framework *AdaptiX*. With its modular design and additional functionalities, *AdaptiX* streamlines the design, development, and evaluation of assistive robot control applications in both virtual simulations and real-world settings. Emphasising flexibility, *AdaptiX* facilitates optimal resource usage and fosters enhanced involvement of the target user group in the research process, thus promoting effective collaboration and user-centred design. Furthermore, this thesis evaluated the iteratively developed framework and its included *ADMC* approach through several empirical user studies. *AdaptiX* emerged as a valuable tool for the development and evaluation of novel interaction designs and feedback mechanisms in virtual environments, facilitating real-time user feedback through attention guidance. Moreover, it enables swift assessment of various input devices through its standardised *User Input Adapter*, effectively bridging the gap between physical robots and virtual communication via XR HMDs.

5.2 Future Work

The research presented in this thesis can serve as a foundational framework for future researchers and practitioners to expand the discussed concepts by integrating novel approaches in AI-enhanced assistive robotics. Consequently, forthcoming studies should investigate the topics outlined below:

5.2.1 Communicating Updated DoF Mapping Suggestions

In our user study [P9], we evaluated whether updated DoF mapping suggestions should be continuously communicated or provided after reaching a specific difference between the currently activated and suggested mapping. Both methods did not exhibit any significant quantitative or qualitative differences, underscoring the necessity for individualisation options to accommodate different user capabilities and needs. In this study, only the DoF mapping considered to be the most *useful* was directly presented to the user. The other modes (i.e., adjusted, translation, rotation, and gripper) had to be assessed and selected sequentially through mode switches. Future work could compare different *at-a-glance* DoF mapping visualisations to assist users in selecting the most helpful mode for their tasks without having to switch sequentially through all different modes. This might also improve comparability of different visualisation approaches. Likewise, different input methods to directly select a specific mode can be designed, developed, and evaluated to mitigate the sequential *one-button-click* approach and probably further decrease task completion time and perceived workload of the users.

5.2.2 *ADMC* Approach in Industrial Settings

The shared control approach introduced in this thesis – *ADMC* – was designed to assist individuals with motor impairments in carrying out ADLs assisted by a robotic arm. Consequently, the *AdaptiX* framework was primarily conceived and optimised to facilitate the development and evaluation of DoF mapping suggestions for pick-and-place tasks, as they are part of many ADLs. However, the application possibilities of AI-enhanced robotic arms extend beyond domestic care to other domains, including the industrial and manufacturing sectors, where robots have long been utilised to support workers. Recent advancements in robotics have led to the emergence of *cell-less* HRI [18], facilitating innovative and close-contact collaborations, sharing the same physical space. As such, workers may use a robotic arm as a *third hand* and could encounter issues because of limited capacity for input modalities – akin to the previous target group of people with motor impairments. A focused ethnographic study to identify key aspects for effective and seamless HRC in the workplace should be conducted as an initial step to gather specific user needs. Based on the respective findings, an adapted approach of specific DoF mappings can be designed and developed to better suit the evaluated industrial tasks. The industrial environment may also impose certain restrictions on a multi-modal feedback approach, such as increased ambient light and noise, which need to be considered during the research process. Additionally, and in contrast to the previous target group, workers in industrial settings may be able to use different kinds of input/feedback technology and different types of robotic arms, requiring further evaluation steps and situation-specific customisation options.

5.2.3 Enhancement of DoF Mapping Suggestions

ADMC represents a concept and initial implementation of a shared control approach utilising DoF mappings based on an underlying rule engine (e.g., a CNN or script-based approach). The CNN's training dataset was derived from a learning-by-demonstration approach, using participant movement data during pick-and-place tasks. Conversely, a script-based approach was employed for a predictable behaviour of the model and therefore more controlled user studies, constraining the use case to specific scenarios and tasks. In comprehensive real-world contexts, numerous additional factors must be considered to suggest the appropriate DoF mapping for the current environment, scenario, task, and state. Future research can focus on enhancing the underlying rule engine by refining training strategies, integrating environmental and contextual data, and analysing human movements across a broader range of daily tasks. Such a comprehensive AI-enhanced model would enable the robotic arm to operate effectively in a wide variety of scenarios – therefore increasing its universality – and achieve more precise control in tasks like grasping objects from a shelf (see Figure 5.1). Furthermore, such a model may be used in different variants of shared control approaches, each tailored to create a mapping based on the output of the AI. Moving forward, the underlying research can be expanded by the integration of multiple different – and potentially more efficient – assistive robots, able to perform an increased and diversified array of tasks. Additionally, various innovative interaction designs, feedback methods, and intervention strategies could be investigated when interacting with AI-enhanced robots.

5.3 Concluding Remarks

Assistive technologies are instrumental in improving the independence and quality of life for individuals with motor impairments. Central to the success of these innovations is the active involvement of the target group in the research process. Their firsthand insights into their contexts and lived experiences are invaluable, ensuring that solutions are effectively tailored to truly meet their needs. Finding operational and technological solutions to facilitate end user inclusion in this process ensures that they can contribute effectively and impactfully. By adopting accessible design principles and leveraging inclusive technologies – like the *in-silico* tool *AdaptiX* – researchers can create environments where all stakeholder voices are heard and respected.

Furthermore, placing emphasis on meaningful shared control and improving the legibility of assistive technologies is essential to mitigate dependency and foster greater user acceptance. Empowering users to maintain control over their devices and care routines not only encourages autonomy but also reinforces their sense of agency. By adopting user-centred approaches and incorporating customisable features, solutions can be tailored to meet individual needs, thereby fostering trust, acceptance, and – consequently – independence.

Looking ahead, future research should focus on refining and expanding upon the concepts introduced in this thesis. This includes further exploration of communication methods for updated DoF mapping suggestions, extending the application of the *ADMC* approach to other settings, and enhancing the underlying *AdaptiX* rule engine to accommodate a broader range of tasks and scenarios. By addressing the discussed challenges and building on the foundation laid out in this work, the field of AI-enhanced assistive robotics can continue improving the quality of life for individuals with physical impairments and contribute to a future where their full inclusion in social and professional spheres becomes a reality.

Statement of Contributions

The research within this thesis would not have been possible without my supervisor, colleagues, and the students I supervised. The table below separates my contribution from others' to the included papers.

Table A.1: Clarification of my own and others' contributions to the projects included in this thesis.

Paper	My Contribution	Co-author(s) Contribution
[P1]	I developed an initial version of the *Unreal Engine* framework, co-developed the concept and methodology, provided resources, supported in the study design, conducted the validation, and have been involved in writing and editing the paper.	K. Kronhardt and S. Rübner – in the context of their scientific specialisation course – extended the given *Unreal Engine* framework by a script-based adaptive control approach, developed an arrow-based visualisation, and extended the framework for a remote study. They conducted the study, performed data curation, and created the first draft of the paper. F. F. Goldau and U. Frese proposed the initial idea of DoF mapping. J. Gerken provided feedback on the paper.

[P2]	I was the leading author, conducted the interviews and in-situ observations, did the thematic analyses and the main part of the paper writing.	A. Baumeister conducted the interviews, in-situ observations, and thematic analysis together with me, supported the study design, and contributed to the paper writing. S. Schneegass, B. Klein, and J. Gerken provided feedback on the paper.
[P3]	I was the leading author, developed the concept, assisted in developing the prototype, designed the study, supported the on-site data gathering, and did the data analysis.	T. Franzen built – in context of his master's thesis – the prototype and conducted the study. K. Kronhardt supported the study design. U. Gruenefeld assisted in the data analysis and writing the paper. S. Schneegass and J. Gerken gave feedback on the paper.
[P4]	I was the leading author, co-developed the concepts, developed the software architecture, and built the comprehensive framework.	F. F. Goldau co-developed the concepts and framework modules. K. Kronhardt was involved in the initial phases of the framework and co-developed the concept. U. Frese and J. Gerken gave feedback on the paper.
[P5]	I was the leading author, conducted the literature review, and analysed the data to derive the intent communication model and empirical findings.	U. Gruenefeld created two of the figures and supported the paper writing. U. Gruenefeld, S. Schneegass, and J. Gerken were involved in the analysis phase to discuss derived entities, dimensions, and properties of the model. Additionally, they gave feedback on the paper.

[P6]	I was the leading author and co-developed the visualisation concepts.	K. Kronhardt and T. Franzen co-developed the visualisation concepts in brainstorming sessions with me. J. Gerken provided feedback on the paper.
[P7]	I was the leading author, developed the initial prototype, conducted both remote studies and interviews, and analysed the data.	K. Kronhardt added a *study mode* to alternate conditions to the prototype. T. Franzen assisted in the initial data analysis. U. Gruenefeld supported designing the study, analysis of data, and writing of the paper. S. Schneegass and J. Gerken gave feedback on the paper.
[P8]	I was the leading author and coordinator.	K. Kronhardt supported with the conceptual ideas within this paper and the draft creation. J. Freienstein supported the manuscript's draft creation and refinement. J. Gerken co-authored the paper and gave feedback.
[P9]	I was the leading author, designed the study, supported the onsite data gathering, and analysed the data.	K. Kronhardt – in context of his master's thesis – built upon the given prototype, conducted the study, and contributed to the data analysis, supported the writing, and provided the figures in the paper. F. F. Goldau and U. Frese proposed the initial idea of DoF mapping. J. Gerken provided feedback on the paper.

[P10]	I was the leading author, developed the overall concept, co-developed the prototype within the framework environment, set the study design, and analysed the study data.	K. Zinta – in context of his bachelor's thesis – co-developed the prototype and conducted the user study. J. Gerken providing feedback during the ideation phase.

My Publications

[P1] Kronhardt, K., Rübner, S., Pascher, M., Goldau, F. F., Frese, U., and Gerken, J. (2022). "Adapt or Perish? Exploring the Effectiveness of Adaptive DoF Control Interaction Methods for Assistive Robot Arms." In: *Technologies* 10.1. DOI: 10.3390/technologies10010030.

[P2] Pascher, M., Baumeister, A., Schneegass, S., Klein, B., and Gerken, J. (2021). "Recommendations for the Development of a Robotic Drinking and Eating Aid - An Ethnographic Study." In: *Human-Computer Interaction – INTERACT 2021*. Ed. by C. Ardito, R. Lanzilotti, A. Malizia, H. Petrie, A. Piccinno, G. Desolda, and K. Inkpen. Springer International Publishing, pp. 331–351. DOI: 10.1007/978-3-030-85623-6_21.

[P3] Pascher, M., Franzen, T., Kronhardt, K., Gruenefeld, U., Schneegass, S., and Gerken, J. (2023a). "HaptiX: Vibrotactile Haptic Feedback for Communication of 3D Directional Cues." In: *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI EA '23. Association for Computing Machinery (ACM). DOI: 10.1145/3544549.3585601.

[P4] Pascher, M., Goldau, F. F., Kronhardt, K., Frese, U., and Gerken, J. (2024a). "AdaptiX – A Transitional XR Framework for Development and Evaluation of Shared Control Applications in Assistive Robotics." In: *Proc. ACM Hum.-Comput. Interact.* 8.EICS. DOI: 10.1145/3660243.

[P5] Pascher, M., Gruenefeld, U., Schneegass, S., and Gerken, J. (2023b). "How to Communicate Robot Motion Intent: A Scoping Review." In: *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. CHI '23. Association for Computing Machinery (ACM). DOI: 10.1145/3544548.3580857.

[P6] Pascher, M., Kronhardt, K., Franzen, T., and Gerken, J. (2022a). "Adaptive DoF: Concepts to Visualize AI-Generated Movements in Human-Robot Collaboration." In: *Proceedings of the 2022 International Conference on Advanced Visual Interfaces*. AVI 2022. Association for Computing Machinery (ACM). DOI: 10.1145/3531073.3534479.

[P7] Pascher, M., Kronhardt, K., Franzen, T., Gruenefeld, U., Schneegass, S., and Gerken, J. (2022b). "My Caregiver the Cobot: Comparing Visualization Techniques to Effectively Communicate Cobot Perception to People with Physical Impairments." In: *Sensors* 22.3. DOI: 10.3390/s22030755.

[P8] Pascher, M., Kronhardt, K., Freienstein, J., and Gerken, J. (2024b). "Exploring AI-enhanced Shared Control for an Assistive Robotic Arm." In: *Engineering Interactive Computer Systems – EICS 2023 International Workshops and Doctoral Consortium*. Ed. by M. Harrison, C. Martinie, N. Micallef, P. Palanque, A. Schmidt, M. Winckler, E. Yigitbas, and L. Zaina. Springer Nature Switzerland, pp. 1–14. DOI: 10.1007/978-3-031-59235-5_10.

[P9] Pascher, M., Kronhardt, K., Goldau, F. F., Frese, U., and Gerken, J. (2023c). "In Time and Space: Towards Usable Adaptive Control for Assistive Robotic Arms." In: *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. Institute of Electrical and Electronics Engineers (IEEE), pp. 2300–2307. DOI: 10.1109/RO-MAN57019.2023.10309381.

[P10] Pascher, M., Zinta, K., and Gerken, J. (2024c). "Exploring of Discrete and Continuous Input Control for AI-enhanced Assistive Robotic Arms." In: *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. HRI '24 Companion. Association for Computing Machinery (ACM). DOI: 10.1145/3610978.3640626.

References

[1] Abbink, D. A., Carlson, T., Mulder, M., Winter, J. C. F. de, Aminravan, F., Gibo, T. L., and Boer, E. R. (2018). “A Topology of Shared Control Systems—Finding Common Ground in Diversity.” In: *IEEE Transactions on Human-Machine Systems* 48.5, pp. 509–525. DOI: 10.1109/thms.2018.2791570.

[2] Abi-Farraj, F., Osa, T., Peters, N. P. J., Neumann, G., and Giordano, P. R. (2017). “A learning-based shared control architecture for interactive task execution.” In: *2017 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 329–335. DOI: 10.1109/ICRA.2017.7989042.

[3] Agree, E. M. and Freedman, V. A. (2011). “A Quality-of-Life Scale for Assistive Technology: Results of a Pilot Study of Aging and Technology.” In: *Physical Therapy* 91.12, pp. 1780–1788. DOI: 10.2522/ptj.20100375.

[4] Ajoudani, A., Zanchettin, A. M., Ivaldi, S., Albu-Schäffer, A., Kosuge, K., and Khatib, O. (2017). “Progress and prospects of the human–robot collaboration.” In: *Autonomous Robots* 42.5, pp. 957–975. DOI: 10.1007/s10514-017-9677-2.

[5] Ali, A. S., Madariaga, M. G., McGeary, D. C., Pruehsner, W. R., and Enderle, J. D. (2007). “The assistive robotic arm.” In: *2007 IEEE 33rd Annual Northeast Bioengineering Conference*, pp. 291–292. DOI: 10.1109/NEBC.2007.4413392.

[6] Alles, D. S. (1970). “Information Transmission by Phantom Sensations.” In: *IEEE Transactions on Man-Machine Systems* 11.1, pp. 85–91. DOI: 10.1109/TMMS.1970.299967.

[7] Amirshirzad, N., Kumru, A., and Oztop, E. (2019). “Human Adaptation to Human–Robot Shared Control.” In: *IEEE Transactions on Human-Machine Systems* 49.2, pp. 126–136. DOI: 10.1109/THMS.2018.2884719.

[8] Anbarasan, S., Masood, K., Zoppi, M., Frémont, V., and Molfino, R. M. (2021). “A Qualitative Analysis to Develop Handover Strategies for Freight Handling Autonomous Vehicle.” In: *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, pp. 2582–2588. DOI: 10.1109/ITSC48978.2021.9565043.

[9] Andersen, R. S., Madsen, O., Moeslund, T. B., and Amor, H. B. (2016). "Projecting robot intentions into human environments." In: *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, pp. 294–301. DOI: 10.1109/ROMAN.2016.7745145.

[10] Anderson, W. L. and Wiener, J. M. (2013). "The Impact of Assistive Technologies on Formal and Informal Home Care." In: *The Gerontologist* 55.3, pp. 422–433. DOI: 10.1093/geront/gnt165.

[11] Arévalo Arboleda, S., Becker, M., and Gerken, J. (2022). "Does One Size Fit All? A Case Study to Discuss Findings of an Augmented Hands-Free Robot Teleoperation Concept for People with and without Motor Disabilities." In: *Technologies* 10.4. DOI: 10.3390/technologies10010004.

[12] Arévalo Arboleda, S., Pascher, M., Baumeister, A., Klein, B., and Gerken, J. (2021a). "Reflecting upon Participatory Design in Human-Robot Collaboration for People with Motor Disabilities: Challenges and Lessons Learned from Three Multiyear Projects." In: *The 14th PErvasive Technologies Related to Assistive Environments Conference - PETRA 2021*. ACM. DOI: 10.1145/3453892.3458044.

[13] Arévalo Arboleda, S., Ruecker, F., Dierks, T., and Gerken, J. (2021b). "Assisting Manipulation and Grasping in Robot Teleoperation with Augmented Reality Visual Cues." In: *CHI Conference on Human Factors in Computing Systems (CHI '21)*. ACM. DOI: 10.1145/3411764.3445398.

[14] Assistive Innovations B.V. (2024). iARM - Robotic arms for humans. Published online: https://www.assistive-innovations.com/en/robotic-arms/iarm (last visited on July 19, 2024).

[15] Ballesteros, J., Urdiales, C., Velasco, A. B. M., and Ramos-Jiménez, G. (2017). "A Biomimetical Dynamic Window Approach to Navigation for Collaborative Control." In: *IEEE Transactions on Human-Machine Systems* 47.6, pp. 1123–1133. DOI: 10.1109/THMS.2017.2700633.

[16] Barralon, P., Ng, G., Dumont, G. A., Schwarz, S. K. W., and Ansermino, J. M. (2009). "Design of Rhythm-Based Vibrotactile Stimuli Around the Waist: Evaluation of Two Encoding Parameters." In: *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 39.5, pp. 1062–1073. DOI: 10.1109/TSMCA.2009.2025026.

[17] Bauer, A., Wollherr, D., and Buss, M. (2008). “Human–Robot Collaboration: A Survey.” In: *International Journal of Humanoid Robotics* 05.01, pp. 47–66. DOI: 10.1142/S0219843608001303.

[18] Bauer, W., Bender, M., Braun, M., and Rally, Peter und Scholtz, Oliver (2016). Lightweight Robots in Manual Assembly – Best to Start Simply: Examining Companies' Initial Experiences with Lightweight Robots. Ed. by Fraunhofer Institute for Industrial Engineering IAO.

[19] Beaudoin, M., Lettre, J., Routhier, F., Archambault, P. S., Lemay, M., and Gélinas, I. (2019). “Long-term use of the JACO robotic arm: a case series.” In: *Disability and Rehabilitation: Assistive Technology* 14.3, pp. 267–275. DOI: 10.1080/17483107.2018.1428692.

[20] Beaudoin, M., Routhier, F., Lettre, J., Archambault, P., and Lemay, M. (2015). “Effects and impacts of a robotic arm used by individuals with upper limb motor impairment: A scoping review.” In: *Annals of Physical and Rehabilitation Medicine* 58, e98. DOI: 10.1016/j.rehab.2015.07.239.

[21] Beaudoin, M., Lettre, J., Routhier, F., Archambault, P. S., Lemay, M., and Gélinas, I. (2018). “Impacts of robotic arm use on individuals with upper extremity disabilities: A scoping review.” In: *Canadian Journal of Occupational Therapy* 85.5. PMID: 30866682, pp. 397–407. DOI: 10.1177/0008417418820878.

[22] Bemelmans, R., Gelderblom, G. J., Jonker, P., and Witte, L. de (2012). “Socially Assistive Robots in Elderly Care: A Systematic Review into Effects and Effectiveness.” In: *Journal of the American Medical Directors Association* 13.2, pp. 114–120. DOI: 10.1016/j.jamda.2010.10.002.

[23] Beydoun, J., Nasrallah, L., Sabrah, T., and Caboral-Stevens, M. (2019). “Towards a Definition of Caregiver Fatigue: A Concept Analysis.” In: *Advances in Nursing Science* 42.4, pp. 297–306. DOI: 10.1097/ans.0000000000000262.

[24] Boada, J. P., Maestre, B. R., and Genís, C. T. (2021). “The ethical issues of social assistive robotics: A critical literature review.” In: *Technology in Society* 67, p. 101726. DOI: https://doi.org/10.1016/j.techsoc.2021.101726.

[25] Bodden, C., Rakita, D., Mutlu, B., and Gleicher, M. (2018). “A flexible optimization-based method for synthesizing intent-expressive robot arm motion.” In: *The International Journal of Robotics Research* 37.11, pp. 1376–1394. DOI: 10.1177/0278364918792295.

[26] Bolano, G., Roennau, A., and Dillmann, R. (2018). “Transparent Robot Behavior by Adding Intuitive Visual and Acoustic Feedback to Motion Replanning.” In: *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, pp. 1075–1080. DOI: 10.1109/ROMAN.2018.8525671.

[27] Bonci, A., Cen Cheng, P. D., Indri, M., Nabissi, G., and Sibona, F. (2021). “Human-Robot Perception in Industrial Environments: A Survey.” In: *Sensors* 21.5. DOI: 10.3390/s21051571.

[28] Bragança, S., Costa, E., Castellucci, I., and Arezes, P. M. (2019). “A Brief Overview of the Use of Collaborative Robots in Industry 4.0: Human Role and Safety.” In: *Occupational and Environmental Safety and Health*. Ed. by P. M. Arezes, J. S. Baptista, M. P. Barroso, P. Carneiro, P. Cordeiro, N. Costa, R. B. Melo, A. S. Miguel, and G. Perestrelo. Springer International Publishing, pp. 641–650. DOI: 10.1007/978-3-030-14730-3_68.

[29] Brooks, C. and Szafir, D. (2020). “Visualization of Intended Assistance for Acceptance of Shared Control.” In: *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 11425–11430. DOI: 10.1109/IROS45743.2020.9340964.

[30] Brose, S. W., Weber, D. J., Salatin, B. A., Grindle, G. G., Wang, H., Vazquez, J. J., and Cooper, R. A. (2010). “The Role of Assistive Robotics in the Lives of Persons with Disability.” In: *American Journal of Physical Medicine &; Rehabilitation* 89.6, pp. 509–521. DOI: 10.1097/phm.0b013e3181cf569b.

[31] Burke, J., Prewett, M., Gray, A., Yang, L., Stilson, F., Coovert, M., Elliot, L., and Redden, E. (2006). “Comparing the effects of visual-auditory and visual-tactile feedback on user performance: a meta-analysis.” In: *Proceedings of the 8th international conference on Multimodal interfaces (ICMI’06)*. Association for Computing Machinery (ACM), pp. 108–117. DOI: 10.1145/1180995.1181017.

[32] Burkolter, D., Weyers, B., Kluge, A., and Luther, W. (2014). "Customization of user interfaces to reduce errors and enhance user acceptance." In: *Applied Ergonomics* 45.2, Part B, pp. 346–353. DOI: `https://doi.org/10.1016/j.apergo.2013.04.017`.

[33] Burtt, H. E. (1917). "Tactual illusions of movement." In: *Journal of Experimental Psychology* 2, pp. 371–385.

[34] Bustamante, S., Peters, J., Schölkopf, B., Grosse-Wentrup, M., and Jayaram, V. (2021). "ArmSym: A Virtual Human–Robot Interaction Laboratory for Assistive Robotics." In: *IEEE Transactions on Human-Machine Systems* 51.6, pp. 568–577. DOI: `10.1109/THMS.2021.3106865`.

[35] Campeau-Lecours, A., Lamontagne, H., Latour, S., Fauteux, P., Maheu, V., Boucher, F., Deguire, C., and L'Ecuyer, L.-J. C. (2017). "Kinova Modular Robot Arms for Service Robotics Applications." In: *International Journal of Robotics Applications and Technologies* 5.2, pp. 49–71. DOI: `10.4018/ijrat.2017070104`.

[36] Canal, G., Alenyà, G., and Torras, C. (2016). "Personalization Framework for Adaptive Robotic Feeding Assistance." In: *Social Robotics*. Vol. 9979. Springer International Publishing, pp. 22–31. DOI: `10.1007/978-3-319-47437-3_3`.

[37] Carmigniani, J., Furht, B., Anisetti, M., Ceravolo, P., Damiani, E., and Ivkovic, M. (2010). "Augmented reality technologies, systems and applications." In: *Multimedia Tools and Applications* 51.1, pp. 341–377. DOI: `10.1007/s11042-010-0660-6`.

[38] Chadalavada, R. T., Andreasson, H., Krug, R., and Lilienthal, A. J. (2015). "That's on my mind! robot to human intention communication through on-board projection on shared floor space." In: *2015 European Conference on Mobile Robots (ECMR)*. Institute of Electrical and Electronics Engineers (IEEE), pp. 1–6. DOI: `10.1109/ECMR.2015.7403771`.

[39] Chadalavada, R. T., Andreasson, H., Schindler, M., Palm, R., and Lilienthal, A. J. (2020). "Bi-directional navigation intent communication using spatial augmented reality and eye-tracking glasses for improved safety in human–robot interaction." In: *Robotics and Computer-Integrated Manufacturing* 61, p. 101830. DOI: `10.1016/j.rcim.2019.101830`.

[40] Chakraborti, T., Sreedharan, S., Kulkarni, A., and Kambhampati, S. (2018). "Projection-Aware Task Planning and Execution for Human-in-the-Loop Operation of Robots in a Mixed-Reality Workspace." In: *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, pp. 4476–4482. DOI: 10.1109/IROS.2018.8593830.

[41] Charouh, Z., Ezzouhri, A., Ghogho, M., and Guennoun, Z. (2022). "A Resource-Efficient CNN-Based Method for Moving Vehicle Detection." In: *Sensors* 22.3, p. 1193. DOI: 10.3390/s22031193.

[42] Chen, T., Wu, Y.-S., and Zhu, K. (2018a). "Investigating Different Modalities of Directional Cues for Multi-Task Visual-Searching Scenario in Virtual Reality." In: *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology*. VRST '18. Association for Computing Machinery (ACM). DOI: 10.1145/3281505.3281516.

[43] Chen, T. L., Ciocarlie, M., Cousins, S., Grice, P. M., Hawkins, K., Hsiao, K., Kemp, C. C., King, C.-H., Lazewatsky, D. A., Leeper, A. E., Nguyen, H., Paepcke, A., Pantofaru, C., Smart, W. D., and Takayama, L. (2013). "Robots for humanity: using assistive robotics to empower people with disabilities." In: *IEEE Robotics & Automation Magazine* 20.1, pp. 30–39. DOI: 10.1109/MRA.2012.2229950.

[44] Chen, X., Zhao, B., Wang, Y., Xu, S., and Gao, X. (2018b). "Control of a 7-DOF Robotic Arm System With an SSVEP-Based BCI." In: *International Journal of Neural Systems* 28.08, p. 1850018. DOI: 10.1142/s0129065718500181.

[45] Cherubini, A., Oriolo, G., Macrí, F., Aloise, F., Cincotti, F., and Mattia, D. (2008). "A multimode navigation system for an assistive robotics project." In: *Autonomous Robots* 25.4, pp. 383–404. DOI: 10.1007/s10514-008-9102-y.

[46] Cholewiak, R. W. and Collins, A. A. (2000). "The generation of vibrotactile patterns on a linear array: influences of body site, time, and presentation mode." In: *Perception & Psychophysics* 62.6, pp. 1220–1235. DOI: 10.3758/bf03212124.

[47] Chung, C.-S., Wang, H., and Cooper, R. A. (2013). "Functional assessment and performance evaluation for assistive robotic manipulators: Literature review." In: *The Journal of Spinal Cord Medicine* 36.4, pp. 273–289. DOI: 10.1179/2045772313Y.0000000132.

[48] Cleaver, A., Tang, D. V., Chen, V., Short, E. S., and Sinapov, J. (2021). "Dynamic Path Visualization for Human-Robot Collaboration." In: *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*. HRI '21 Companion. Association for Computing Machinery (ACM), pp. 339–343. DOI: 10.1145/3434074.3447188.

[49] Collett, J., Dawes, H., Cavey, A., Meaney, A., Sackley, C., Wade, D., and Howells, K. (2011). "Hydration and independence in activities of daily living in people with multiple sclerosis: a pilot investigation." In: *Disability and Rehabilitation* 33.19-20, pp. 1822–1825. DOI: 10.3109/09638288.2010.549286.

[50] Comari, S., Di Leva, R., Carricato, M., Badini, S., Carapia, A., Collepalumbo, G., Gentili, A., Mazzotti, C., Staglianò, K., and Rea, D. (2022). "Mobile cobots for autonomous raw-material feeding of automatic packaging machines." In: *Journal of Manufacturing Systems* 64, pp. 211–224. DOI: 10.1016/j.jmsy.2022.06.007.

[51] Connors, K., Mahony, L., and Morgan, P. (2019). "Variation in assistive technology use in Motor Neuron Disease according to clinical phenotypes and ALS Functional Rating Scale – Revised Score: A prospective observational study." In: *NeuroRehabilitation* 44.2, pp. 303–313. DOI: 10.3233/nre-182511.

[52] Coovert, M. D., Lee, T., Shindev, I., and Sun, Y. (2014). "Spatial augmented reality as a method for a mobile robot to communicate intended movement." In: *Computers in Human Behavior* 34, pp. 241–248. DOI: 10.1016/j.chb.2014.02.001.

[53] Cox, E. O. and Dooley, A. C. (1996). "Care-Receivers' Perception of Their Role in the Care Process." In: *Journal of Gerontological Social Work* 26.1–2, pp. 133–152. DOI: 10.1300/j083v26n01_09.

[54] Cresswell, K., Cunningham-Burley, S., and Sheikh, A. (2018). "Health Care Robotics: Qualitative Exploration of Key Challenges and Future Directions." In: *Journal of Medical Internet Research* 20.7, e10410. DOI: 10.2196/10410.

[55] Czubenko, M. and Kowalczuk, Z. (2021). "A Simple Neural Network for Collision Detection of Collaborative Robots." In: *Sensors* 21.12, p. 4235. DOI: 10.3390/s21124235.

[56] Dianatfar, M., Latokartano, J., and Lanz, M. (2021). "Review on existing VR/AR solutions in human–robot collaboration." In: *Procedia CIRP* 97. 8th CIRP Conference of Assembly Technology and Systems, pp. 407–411. DOI: `https://doi.org/10.1016/j.procir.2020.05.259`.

[57] Diederich, A. and Colonius, H. (2004). "Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time." In: *Perception & psychophysics* 66.8, pp. 1388–1404.

[58] Donati, M., Fanucci, L., Iacopetti, F., and Vecchio, A. (2015). "A modular suite to support human-smartphone interaction for people with motor skill impairments." In: *2015 10th International Conference on Design & Technology of Integrated Systems in Nanoscale Era (DTIS)*, pp. 1–5. DOI: `10.1109/DTIS.2015.7127385`.

[59] Drolshagen, S., Pfingsthorn, M., Gliesche, P., and Hein, A. (2021). "Acceptance of industrial collaborative robots by people with disabilities in sheltered workshops." In: *Frontiers in Robotics and AI*, p. 173.

[60] Elendu, C., Amaechi, D. C., Elendu, T. C., Jingwa, K. A., Okoye, O. K., John Okah, M., Ladele, J. A., Farah, A. H., and Alimi, H. A. (2023). "Ethical implications of AI and robotics in healthcare: A review." In: *Medicine* 102.50, e36671. DOI: `10.1097/md.0000000000036671`.

[61] Epic Games (2024). Unreal Engine. Published online: `https://www.unrealengine.com` (last visited on July 19, 2024). Version 4.27.2.

[62] Erdogan, A. and Argall, B. D. (2017). "The effect of robotic wheelchair control paradigm and interface on user performance, effort and preference: An experimental assessment." In: *Robotics and Autonomous Systems* 94, pp. 282–297. DOI: `https://doi.org/10.1016/j.robot.2017.04.013`.

[63] Ettari, S. V., Roden, E., Ahuja, V., and Samadani, U. (2022). "Oculogica: An Eye-Catching Innovation in Health Care and The Privacy Implications Of Artificial Intelligence and Machine Learning in Diagnostics For The Human Brain." In: *SMU Science and Technology Law Review* 25.1, p. 23. DOI: `10.25172/smustlr.25.1.3`.

[64] EXXOMOVE (2024). BATEO - Armmobility. Published online: `https://www.exxomove.de/armmobilitaet` (last visited on July 19, 2024).

[65] Ezeh, C., Trautman, P., Devigne, L., Bureau, V., Babel, M., and Carlson, T. (2017). "Probabilistic vs linear blending approaches to shared control for wheelchair driving." In: *2017 International Conference on Rehabilitation Robotics (ICORR)*. Institute of Electrical and Electronics Engineers (IEEE), pp. 835–840. DOI: 10.1109/ICORR.2017.8009352.

[66] Fattal, C., Leynaert, V., Laffont, I., Baillet, A., Enjalbert, M., and Leroux, C. (2018). "SAM, an Assistive Robotic Device Dedicated to Helping Persons with Quadriplegia: Usability Study." In: *International Journal of Social Robotics* 11.1, pp. 89–103. DOI: 10.1007/s12369-018-0482-7.

[67] Federici, S. (2018). Assistive technology assessment handbook. CRC Press, Taylor & Francis Group.

[68] Fitzsimons, K., Kalinowska, A., Dewald, J. P., and Murphey, T. D. (2020). "Task-based hybrid shared control for training through forceful interaction." In: *The International Journal of Robotics Research* 39.9, pp. 1138–1154. DOI: 10.1177/0278364920933654.

[69] Flemisch, F., Abbink, D., Itoh, M., Pacaux-Lemoine, M.-P., and Weßel, G. (2016). "Shared control is the sharp end of cooperation: Towards a common framework of joint action, shared control and human machine cooperation." In: *IFAC-PapersOnLine* 49.19, pp. 72–77. DOI: 10.1016/j.ifacol.2016.10.464.

[70] Flemisch, F., Abbink, D. A., Itoh, M., Pacaux-Lemoine, M.-P., and Weßel, G. (2019). "Joining the blunt and the pointy end of the spear: towards a common framework of joint action, human–machine cooperation, cooperative guidance and control, shared, traded and supervisory control." In: *Cognition, Technology & Work* 21.4, pp. 555–568. DOI: 10.1007/s10111-019-00576-1.

[71] Fosch-Villaronga, E. and Özcan, B. (2019). "The Progressive Intertwinement Between Design, Human Needs and the Regulation of Care Technology: The Case of Lower-Limb Exoskeletons." In: *International Journal of Social Robotics* 12.4, pp. 959–972. DOI: 10.1007/s12369-019-00537-8.

[72] Fouad, H., Hassanein, A. S., Soliman, A. M., and Al-Feel, H. (2020). "Analyzing patient health information based on IoT sensor with AI for improving patient assistance in the future direction." In: *Measurement* 159, p. 107757. DOI: 10.1016/j.measurement.2020.107757.

[73] Freitas, M. P. de, Piai, V. A., Farias, R. H., Fernandes, A. M. R., Moraes Rossetto, A. G. de, and Leithardt, V. R. Q. (2022). "Artificial Intelligence of Things Applied to Assistive Technology: A Systematic Literature Review." In: *Sensors* 22.21, p. 8531. DOI: 10.3390/s22218531.

[74] Galin, R. R. and Meshcheryakov, R. V. (2020). "Human-Robot Interaction Efficiency and Human-Robot Collaboration." In: *Robotics: Industry 4.0 Issues & New Intelligent Control Paradigms*. Ed. by A. G. Kravets. Springer International Publishing, pp. 55–63. DOI: 10.1007/978-3-030-37841-7_5.

[75] Gallenberger, D., Bhattacharjee, T., Kim, Y., and Srinivasa, S. S. (2019). "Transfer Depends on Acquisition: Analyzing Manipulation Strategies for Robotic Feeding." In: *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 267–276. DOI: 10.1109/HRI.2019.8673309.

[76] Gandhi, V., Prasad, G., Coyle, D., Behera, L., and McGinnity, T. M. (2014). "EEG-Based Mobile Robot Control Through an Adaptive Brain–Robot Interface." In: *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 44.9, pp. 1278–1285. DOI: 10.1109/TSMC.2014.2313317.

[77] Garcia, P. P., Santos, T. G., Machado, M. A., and Mendes, N. (2023). "Deep Learning Framework for Controlling Work Sequence in Collaborative Human–Robot Assembly Processes." In: *Sensors* 23.1, p. 553. DOI: 10.3390/s23010553.

[78] Gbouna, Z. V., Pang, G., Yang, G., Hou, Z., Lv, H., Yu, Z., and Pang, Z. (2021). "User-Interactive Robot Skin With Large-Area Scalability for Safer and Natural Human-Robot Collaboration in Future Telehealthcare." In: *IEEE Journal of Biomedical and Health Informatics* 25.12, pp. 4276–4288. DOI: 10.1109/JBHI.2021.3082563.

[79] Geldard, F. A. (1975). Sensory saltation : metastability in the perceptual world / Frank A. Geldard. Lawrence Erlbaum Associates ; distributed by the Halsted Press Division of Wiley Hillsdale, N.J. : New York, 133 p. :

[80] Geldard, F. A. and Sherrick, C. E. (1972). "The Cutaneous "Rabbit": A Perceptual Illusion." In: *Science* 178.4057, pp. 178–179. DOI: 10.1126/

science.178.4057.178. eprint: https://www.science.org/doi/pdf/10.1126/science.178.4057.178.

[81] Gillespie, T. (2020). "Content moderation, AI, and the question of scale." In: *Big Data & Society* 7.2, p. 205395172094323. DOI: 10.1177/2053951720943234.

[82] Goel, M., Tomar, P. K., Vinjamuri, L. P., Swamy Reddy, G., Al-Taee, M., and Alazzam, M. B. (2023). "Using AI for Predictive Analytics in Financial Management." In: *2023 3rd International Conference on Advance Computing and Innovative Technologies in Engineering (ICACITE)*, pp. 963–967. DOI: 10.1109/ICACITE57410.2023.10182711.

[83] Goldau, F., Shivashankar, Y., Baumeister, A., Drescher, L., Tolle, P., and Frese, U. (2023). "DORMADL - Dataset of Human-Operated Robot Arm Motion in Activities of Daily Living." In: *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 11396–11403. DOI: 10.1109/IROS55552.2023.10341459.

[84] Goldau, F. F. and Frese, U. (2021). "Learning to Map Degrees of Freedom for Assistive User Control: Towards an Adaptive DoF-Mapping Control for Assistive Robots." In: *The 14th PErvasive Technologies Related to Assistive Environments Conference*. PETRA 2021. Association for Computing Machinery (ACM), pp. 132–139. DOI: 10.1145/3453892.3453895.

[85] Golluccio, G., Gillini, G., Marino, A., and Antonelli, G. (2021). "Robot Dynamics Identification: A Reproducible Comparison With Experiments on the Kinova Jaco." In: *IEEE Robotics & Automation Magazine* 28.3, pp. 128–140. DOI: 10.1109/MRA.2020.3004149.

[86] Gómez-Sebastià, I., Garcia-Gasulla, D., Barruè, C., Vázquez-Salceda, J., and Cortés, U. (2011). "ALIVE Meets SHARE-it: An Agent-Oriented Solution to Model Organisational and Normative Requirements in Assistive Technologies." In: *Electronic Healthcare*. Springer Berlin Heidelberg, pp. 319–326. DOI: 10.1007/978-3-642-23635-8_40.

[87] Goodacre, K., McCreadie, C., Flanagan, S., and Lansley, P. (2008). "Enabling Older People to Stay at Home: The Costs of Substituting and Supplementing Care with Assistive Technology." In: *British Journal of Occupational Therapy* 71.4, pp. 130–140. DOI: 10.1177/030802260807100402.

[88] Goodrich, M. A. and Schultz, A. C. (2008). "Human–Robot Interaction: A Survey." In: *Foundations and Trends® in Human–Computer Interaction* 1.3, pp. 203–275. DOI: 10.1561/1100000005.

[89] Google LLC (2024). Google Cardboard – Google VR. Published online: https://arvr.google.com/cardboard/ (last visited on July 19, 2024).

[90] Gopinath, D., Jain, S., and Argall, B. D. (2017). "Human-in-the-Loop Optimization of Shared Autonomy in Assistive Robotics." In: *IEEE Robotics and Automation Letters* 2.1, pp. 247–254. DOI: 10.1109/LRA.2016.2593928.

[91] Gottardi, A., Tortora, S., Tosello, E., and Menegatti, E. (2022). "Shared Control in Robot Teleoperation With Improved Potential Fields." In: *IEEE Transactions on Human-Machine Systems* 52.3, pp. 410–422. DOI: 10.1109/THMS.2022.3155716.

[92] Grau, A., Indri, M., Lo Bello, L., and Sauter, T. (2021). "Robots in Industry: The Past, Present, and Future of a Growing Collaboration With Humans." In: *IEEE Industrial Electronics Magazine* 15.1, pp. 50–61. DOI: 10.1109/MIE.2020.3008136.

[93] Gruenefeld, U., Prädel, L., Illing, J., Stratmann, T., Drolshagen, S., and Pfingsthorn, M. (2020). "Mind the ARm: Realtime Visualization of Robot Motion Intent in Head-Mounted Augmented Reality." In: *Proceedings of the Conference on Mensch Und Computer*. MuC '20. Association for Computing Machinery (ACM), pp. 259–266. DOI: 10.1145/3404983.3405509.

[94] Grushko, S., Vysocký, A., Heczko, D., and Bobovský, Z. (2021). "Intuitive Spatial Tactile Feedback for Better Awareness about Robot Trajectory during Human-Robot Collaboration." In: *Sensors (Basel, Switzerland)* 21.17. DOI: 10.3390/s21175748.

[95] Guan, J., Irizawa, J., and Morris, A. (2022). "Extended Reality and Internet of Things for Hyper-Connected Metaverse Environments." In: *2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW)*, pp. 163–168. DOI: 10.1109/VRW55335.2022.00043.

[96] Günther, S., Müller, F., Funk, M., Kirchner, J., Dezfuli, N., and Mühlhäuser, M. (2018). "TactileGlove: TactileGlove: Assistive Spatial

Guidance in 3D Space through Vibrotactile Navigation." In: *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference*. ACM, pp. 273–280. DOI: 10.1145/3197768.3197785.

[97] Haas, U. (2012). Pflege von Menschen mit Querschnittlähmung Probleme, Bedürfnisse, Ressourcen und Interventionen. Huber.

[98] Hamidi, F., Poneres, K., Massey, A., and Hurst, A. (2018). "Who Should Have Access to my Pointing Data? Privacy Tradeoffs of Adaptive Assistive Technologies." In: *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility*. ASSETS '18. Association for Computing Machinery (ACM), pp. 203–216. DOI: 10.1145/3234695.3239331.

[99] Harris, N. (2017). "The Design and Development of Assistive Technology." In: *IEEE Potentials* 36.1, pp. 24–28. DOI: 10.1109/MPOT.2016.2615107.

[100] Hart, S. G. (2006). "Nasa-Task Load Index (NASA-TLX); 20 Years Later." In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50.9, pp. 904–908. DOI: 10.1177/154193120605000909.

[101] Hashimoto, K., Saito, F., Yamamoto, T., and Ikeda, K. (2013). "A Field Study of the Human Support Robot in the Home Environment." In: *2013 IEEE Workshop on Advanced Robotics and Its Social Impacts (ARSO)*. Institute of Electrical and Electronics Engineers (IEEE), pp. 143–150. DOI: 10.1109/ARSO.2013.6705520.

[102] Hentout, A., Aouache, M., Maoudj, A., and Akli, I. (2019). "Human–robot interaction in industrial collaborative robotics: a literature review of the decade 2008–2017." In: *Advanced Robotics* 33.15-16, pp. 764–799. DOI: 10.1080/01691864.2019.1636714.

[103] Herlant, L. V., Holladay, R. M., and Srinivasa, S. S. (2016). "Assistive Teleoperation of Robot Arms via Automatic Time-Optimal Mode Switching." In: *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*. HRI '16. Institute of Electrical and Electronics Engineers (IEEE), pp. 35–42. DOI: 10.1109/HRI.2016.7451731.

[104] Hetherington, N. J., Croft, E. A., and Loos, H. M. V. der (2021). "Hey Robot, Which Way Are You Going? Nonverbal Motion Legibility Cues

for Human-Robot Spatial Interaction." In: *IEEE Robotics and Automation Letters* 6.3, pp. 5010–5015. DOI: 10.1109/lra.2021.3068708.

[105] Holloway, C. (2019). "Disability Interaction (DIX): A Manifesto." In: *Interactions* 26.2, pp. 44–49. DOI: 10.1145/3310322.

[106] *ISO 9241-210:* Ergonomics of human-system interaction - Part 210: Human-centred design for interactive systems (2019). International Standard. Beuth Verlag.

[107] Jain, S., Farshchiansadegh, A., Broad, A., Abdollahi, F., Mussa-Ivaldi, F., and Argall, B. (2015). "Assistive robotic manipulation through shared autonomy and a Body-Machine Interface." In: *2015 IEEE International Conference on Rehabilitation Robotics (ICORR)*. Institute of Electrical and Electronics Engineers (IEEE), pp. 526–531. DOI: 10.1109/ICORR.2015.7281253.

[108] Jamshidi, F., Zhang, L., and Nezhadalinaei, F. (2021). "Autonomous Driving Systems: Developing an Approach based on A* and Double Q-Learning." In: *2021 7th International Conference on Web Research (ICWR)*, pp. 82–85. DOI: 10.1109/ICWR51868.2021.9443139.

[109] Javdani, S., Srinivasa, S., and Bagnell, A. (2015). "Shared Autonomy via Hindsight Optimization." In: *Robotics: Science and Systems XI*. RSS2015. Robotics: Science and Systems Foundation. DOI: 10.15607/rss.2015.xi.032.

[110] Joamets, K. and Chochia, A. (2021). "Access to Artificial Intelligence for Persons with Disabilities: Legal and Ethical Questions Concerning the Application of Trustworthy AI." In: *Acta Baltica Historiae et Philosophiae Scientiarum* 9.1, pp. 51–66. DOI: 10.11590/abhps.2021.1.04.

[111] Kadouche, R., Mokhtari, M., Giroux, S., and Abdulrazak, B. (2008). "Semantic Matching Framework for Personalizing ubiquitous Environment dedicated to people with special needs." In: *2008 Third International Conference on Pervasive Computing and Applications*. Vol. 2, pp. 582–587. DOI: 10.1109/ICPCA.2008.4783679.

[112] Kallinen, K. (2017). "The Effects of Transparency and Task Type on Trust, Stress, Quality of Work, and Co-Worker Preference During Human-Autonomous System Collaborative Work." In: *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on*

Human-Robot Interaction. HRI '17. Association for Computing Machinery (ACM), pp. 153–154. DOI: 10.1145/3029798.3038386.

[113] Kane, S. K., Jayant, C., Wobbrock, J. O., and Ladner, R. E. (2009). "Freedom to roam: a study of mobile device adoption and accessibility for people with visual and motor disabilities." In: *Proceedings of the 11th International ACM SIGACCESS Conference on Computers and Accessibility*. Assets '09. Association for Computing Machinery, pp. 115–122. DOI: 10.1145/1639642.1639663.

[114] Kang, S.-G., Song, S.-W., Kim, S.-H., Kang, Y.-J., Kim, Y.-R., and Eun, Y. (2020). "Fatigue and Mental Status of Caregivers of Severely Chronically Ill Patients." In: *Pain Research and Management* 2020, pp. 1–6. DOI: 10.1155/2020/6372857.

[115] Kaul, O. B. and Rohs, M. (2016). "HapticHead: 3D Guidance and Target Acquisition through a Vibrotactile Grid." In: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*.

[116] Keates, S., Trewin, S., and Elliott, J. P. (2006). "Including Users with Motor Impairments in Design." In: *Encyclopedia of Human Computer Interaction*. IGI Global, pp. 317–323. DOI: 10.4018/978-1-59140-562-7.ch049.

[117] Kim, D.-J., Hazlett-Knudsen, R., Culver-Godfrey, H., Rucks, G., Cunningham, T., Portee, D., Bricout, J., Wang, Z., and Behal, A. (2012). "How Autonomy Impacts Performance and Satisfaction: Results From a Study With Spinal Cord Injured Subjects Using an Assistive Robot." In: *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 42.1, pp. 2–14. DOI: 10.1109/TSMCA.2011.2159589.

[118] Kinova inc. (2021). Jaco assistive robot - User guide. Published online: https://assistive.kinovarobotics.com/uploads/EN-UG-007-Jaco-user-guide-R05.pdf (last visited on July 19, 2024). EN-UG-007-r05-202111.

[119] Kirman, J. H. (1974). "Tactile apparent movement: The effects of interstimulus onset interval and stimulus duration." In: *Perception & Psychophysics* 15.1, pp. 1–6. DOI: 10.3758/BF03205819.

[120] Klebbe, R., Klüber, K., Dahms, R., and Onnasch, L. (2022). "Caregivers' Perspectives on Human–Robot Collaboration in Inpatient Elderly Care Settings." In: *Machines* 11.1, p. 34. DOI: 10.3390/machines11010034.

[121] Klein, B. (2020). Hilfsmittel, Assistive Technologien und Robotik: Selbstständigkeit und Lebensqualität im Alter erhalten. Kohlhammer Verlag.

[122] Klimaszewski, J. and Władziński, M. (2021). "Human Body Parts Proximity Measurement Using Distributed Tactile Robotic Skin." In: *Sensors* 21.6. DOI: 10.3390/s21062138.

[123] Kosinski, R. J. (2008). "A literature review on reaction time." In: *Clemson University* 10.1, pp. 337–344.

[124] Kragic, D., Gustafson, J., Karaoguz, H., Jensfelt, P., and Krug, R. (2018). "Interactive, Collaborative Robots: Challenges and Opportunities." In: *Proceedings of the 27th International Joint Conference on Artificial Intelligence*. IJCAI'18. AAAI Press, pp. 18–25.

[125] Krishnan, R. H. and Pugazhenthi, S. (2013). "Mobility assistive devices and self-transfer robotic systems for elderly, a review." In: *Intelligent Service Robotics* 7.1, pp. 37–49. DOI: 10.1007/s11370-013-0142-6.

[126] Kruse, T., Pandey, A. K., Alami, R., and Kirsch, A. (2013). "Human-aware robot navigation: A survey." In: *Robotics and Autonomous Systems* 61.12, pp. 1726–1743. DOI: 10.1016/j.robot.2013.05.007.

[127] Kruthika, K., Kiran Kumar, B., and Lakshminarayanan, S. (2016). "Design and development of a robotic arm." In: *2016 International Conference on Circuits, Controls, Communications and Computing (I4C)*, pp. 1–4. DOI: 10.1109/CIMCA.2016.8053274.

[128] Kyrarini, M., Lygerakis, F., Rajavenkatanarayanan, A., Sevastopoulos, C., Nambiappan, H. R., Chaitanya, K. K., Babu, A. R., Mathew, J., and Makedon, F. (2021). "A survey of robots in healthcare." In: *Technologies* 9.1, p. 8. DOI: 10.3390/technologies9010008.

[129] Layman, F., Crowle, C., and Ravenscroft, J. (2012). "Communication Access for a Student with Multiple Disabilities: An Interdisciplinary Collaborative Approach." In: *Computers Helping People with Special*

Needs. Springer Berlin Heidelberg, pp. 353–360. DOI: 10.1007/978-3-642-31534-3_53.

[130] Lebrasseur, A., Lettre, J., Routhier, F., Archambault, P. S., and Campeau-Lecours, A. (2019). "Assistive robotic arm: Evaluation of the performance of intelligent algorithms." In: *Assistive Technology* 33.2, pp. 95–104. DOI: 10.1080/10400435.2019.1601649.

[131] Lee, S. C. and Starner, T. (2010). "BuzzWear: Alert Perception in Wearable Tactile Displays on the Wrist." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '10. Association for Computing Machinery (ACM), pp. 433–442. DOI: 10.1145/1753326.1753392.

[132] Lehtinen, V., Oulasvirta, A., Salovaara, A., and Nurmi, P. (2012). "Dynamic Tactile Guidance for Visual Search Tasks." In: *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology*. Association for Computing Machinery (ACM), pp. 445–452.

[133] Li, T.-H. S., Kuo, P.-H., Tsai, T.-N., and Luan, P.-C. (2019). "CNN and LSTM Based Facial Expression Analysis Model for a Humanoid Robot." In: *IEEE Access* 7, pp. 93998–94011. DOI: 10.1109/ACCESS.2019.2928364.

[134] Linsinger, M., Stecken, J., Kutschinski, J., and Kuhlenkötter, B. (2019). "Situational task change of lightweight robots in hybrid assembly systems." In: *Procedia CIRP* 81, pp. 81–86. DOI: 10.1016/j.procir.2019.03.015.

[135] Logitech (2024). Adaptive Gaming Kit. Published online: https://www.logitechg.com/en-us/products/gamepads/adaptive-gaming-kit-accessories.943-000318.html (last visited on July 19, 2024).

[136] Losey, D. P., Jeon, H. J., Li, M., Srinivasan, K., Mandlekar, A., Garg, A., Bohg, J., and Sadigh, D. (2021). "Learning latent actions to control assistive robots." In: *Autonomous Robots* 46.1, pp. 115–147. DOI: 10.1007/s10514-021-10005-w.

[137] Lunding, R. S., Lystbæk, M. N., Feuchtner, T., and Grønbæk, K. (2023). "AR-supported Human-Robot Collaboration: Facilitating Workspace Awareness and Parallelized Assembly Tasks." In: *2023 IEEE Interna-*

tional Symposium on Mixed and Augmented Reality (ISMAR), pp. 1064–1073. DOI: 10.1109/ISMAR59233.2023.00123.

[138] Lunding, R. S., Lunding, M. S., Feuchtner, T., Petersen, M. G., Grønbæk, K., and Suzuki, R. (2024). "RoboVisAR: Immersive Authoring of Condition-based AR Robot Visualisations." In: *Proceedings of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. HRI '24. Association for Computing Machinery, pp. 462–471. DOI: 10.1145/3610977.3634972.

[139] Mahdi, H., Akgun, S. A., Saleh, S., and Dautenhahn, K. (2022). "A survey on the design and evolution of social robots — Past, present and future." In: *Robotics and Autonomous Systems* 156, p. 104193. DOI: https://doi.org/10.1016/j.robot.2022.104193.

[140] Maheu, V., Archambault, P. S., Frappier, J., and Routhier, F. (2011). "Evaluation of the JACO robotic arm: Clinico-economic study for powered wheelchair users with upper-extremity disabilities." In: *2011 IEEE International Conference on Rehabilitation Robotics*, pp. 1–5. DOI: 10.1109/ICORR.2011.5975397.

[141] Mann, W. C. (2005). Smart Technology for Aging, Disability, and Independence: The State of the Science. Wiley. DOI: 10.1002/0471743941.

[142] Martinsen, B., Harder, I., and Biering-Sorensen, F. (2008). "The meaning of assisted feeding for people living with spinal cord injury: a phenomenological study." In: *Journal of Advanced Nursing* 62.5, pp. 533–540. DOI: 10.1111/j.1365-2648.2008.04637.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1365-2648.2008.04637.x.

[143] Matheson, E., Minto, R., Zampieri, E. G. G., Faccio, M., and Rosati, G. (2019). "Human–Robot Collaboration in Manufacturing Applications: A Review." In: *Robotics* 8.4. DOI: 10.3390/robotics8040100.

[144] Matsas, E. and Vosniakos, G.-C. (2015). "Design of a virtual reality training system for human–robot collaboration in manufacturing tasks." In: *International Journal on Interactive Design and Manufacturing (IJIDeM)* 11.2, pp. 139–153. DOI: 10.1007/s12008-015-0259-2.

[145] Matsumaru, T. (2006). "Mobile Robot with Preliminary-announcement and Display Function of Forthcoming Motion using Projection Equip-

ment." In: *ROMAN 2006 - The 15th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 443–450. DOI: 10.1109/ROMAN.2006.314368.

[146] Matsumaru, T. (2007). "Mobile Robot with Preliminary-announcement and Indication Function of Forthcoming Operation using Flat-panel Display." In: *Proceedings 2007 IEEE International Conference on Robotics and Automation*. Institute of Electrical and Electronics Engineers (IEEE), pp. 1774–1781. DOI: 10.1109/ROBOT.2007.363579.

[147] Mavrogiannis, C., Baldini, F., Wang, A., Zhao, D., Trautman, P., Steinfeld, A., and Oh, J. (2023). "Core Challenges of Social Robot Navigation: A Survey." In: *J. Hum.-Robot Interact.* 12.3. DOI: 10.1145/3583741.

[148] McColl, D., Louie, W.-Y. G., and Nejat, G. (2013). "Brian 2.1: A Socially Assistive Robot for the Elderly and Cognitively Impaired." In: *IEEE Robotics & Automation Magazine* 20.1, pp. 74–83. DOI: 10.1109/MRA.2012.2229939.

[149] McDaniel, T., Goldberg, M., Villanueva, D., Viswanathan, L. N., and Panchanathan, S. (2011). "Motor learning using a kinematic-vibrotactile mapping targeting fundamental movements." In: *Proceedings of the 19th ACM international conference on Multimedia*. ACM.

[150] McDonald, N., Massey, A. K., and Hamidi, F. (2021). "AI-Enhanced Adaptive Assistive Technologies: Methods for AI Design Justice." In: *IEEE Data Eng. Bull.* 44, pp. 3–13.

[151] McMillen, A.-M. and Söderberg, S. (2002). "Disabled Persons' Experience of Dependence on Assistive Devices." In: *Scandinavian Journal of Occupational Therapy* 9.4, pp. 176–183. DOI: 10.1080/110381202260501208.

[152] McMurray, J., Strudwick, G., Forchuk, C., Morse, A., Lachance, J., Baskaran, A., Allison, L., and Booth, R. (2017). "The Importance of Trust in the Adoption and Use of Intelligent Assistive Technology by Older Adults to Support Aging in Place: Scoping Review Protocol." In: *JMIR Res Protoc* 6.11, e218. DOI: 10.2196/resprot.8772.

[153] Merkel, S. and Kucharski, A. (2018). "Participatory Design in Gerontechnology: A Systematic Literature Review." In: *The Gerontologist* 59.1, e16–e25. DOI: 10.1093/geront/gny034.

[154] Meta (2024). Meta Quest 2. Published online: https://www.meta.com/de/quest/products/quest-2/ (last visited on July 19, 2024).

[155] Microsoft (2024). Xbox Adaptive Controller. Published online: https://www.microsoft.com/en-us/d/xbox-adaptive-controller/8nsdbhz1n3d8 (last visited on July 19, 2024).

[156] Milgram, P. and Kishino, F. (1994). "A taxonomy of mixed reality visual displays." In: *IEICE Transactions on Information and Systems* 77.12, pp. 1321–1329.

[157] Muelling, K., Venkatraman, A., Valois, J.-S., Downey, J. E., Weiss, J., Javdani, S., Hebert, M., Schwartz, A. B., Collinger, J. L., and Bagnell, J. A. (2017). "Autonomy infused teleoperation with application to brain computer interface controlled manipulation." In: *Autonomous Robots* 41.6, pp. 1401–1422. DOI: 10.1007/s10514-017-9622-4.

[158] Mulfari, D., Minnolo, A. L., and Puliafito, A. (2017). "Wearable Devices and IoT as Enablers of Assistive Technologies." In: *2017 10th International Conference on Developments in eSystems Engineering (DeSE)*, pp. 14–19. DOI: 10.1109/DeSE.2017.51.

[159] Mullen, J., Mosier, J., Chakrabarti, S., Chen, A., White, T., and Losey, D. (2021). "Communicating Inferred Goals With Passive Augmented Reality and Active Haptic Feedback." In: *IEEE Robotics and Automation Letters* 6.4, pp. 8522–8529. DOI: 10.1109/LRA.2021.3111055.

[160] Müller, S. L., Schröder, S., Jeschke, S., and Richert, A. (2017). "Design of a Robotic Workmate." In: *Lecture Notes in Computer Science*. Springer International Publishing, pp. 447–456. DOI: 10.1007/978-3-319-58463-8_37.

[161] Muthu, P., Tan, Y., Latha, S., Dhanalakshmi, S., Lai, K. W., and Wu, X. (2023). "Discernment on assistive technology for the care and support requirements of older adults and differently-abled individuals." In: *Frontiers in Public Health* 10. DOI: 10.3389/fpubh.2022.1030656.

[162] Nintendo (2024). Nintendo Joy-Con. Published online: https://www.nintendo.de/Hardware/Zubehor/Zubehor-1183144.html (last visited on July 19, 2024).

[163] Oladele, D. A., Markus, E. D., and Abu-Mahfouz, A. M. (2021). "Adaptability of Assistive Mobility Devices and the Role of the Internet of Medical Things: Comprehensive Review." In: *JMIR Rehabilitation and Assistive Technologies* 8.4, e29610. DOI: 10.2196/29610.

[164] Ortiz-Escobar, L. M., Chavarria, M. A., Schönenberger, K., Hurst, S., Stein, M. A., Mugeere, A., and Rivas Velarde, M. (2023). "Assessing the implementation of user-centred design standards on assistive technology for persons with visual impairments: a systematic review." In: *Frontiers in Rehabilitation Sciences* 4. DOI: 10.3389/fresc.2023.1238158.

[165] Ozmen Garibay, O., Winslow, B., Andolina, S., Antona, M., Bodenschatz, A., Coursaris, C., Falco, G., Fiore, S. M., Garibay, I., Grieman, K., Havens, J. C., Jirotka, M., Kacorri, H., Karwowski, W., Kider, J., Konstan, J., Koon, S., Lopez-Gonzalez, M., Maifeld-Carucci, I., McGregor, S., Salvendy, G., Shneiderman, B., Stephanidis, C., Strobel, C., Ten Holter, C., and Xu, W. (2023). "Six Human-Centered Artificial Intelligence Grand Challenges." In: *International Journal of Human–Computer Interaction* 39.3, pp. 391–437. DOI: 10.1080/10447318.2022.2153320.

[166] Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., McGuinness, L. A., Stewart, L. A., Thomas, J., Tricco, A. C., Welch, V. A., Whiting, P., and Moher, D. (2021). "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews." In: *Systematic Reviews* 10.1. DOI: 10.1186/s13643-021-01626-4.

[167] Panchanathan, S., Chakraborty, S., McDaniel, T., and Tadayon, R. (2016). "Person-Centered Multimedia Computing: A New Paradigm Inspired by Assistive and Rehabilitative Applications." In: *IEEE MultiMedia* 23.3, pp. 12–19. DOI: 10.1109/MMUL.2016.51.

[168] Park, D., Hoshi, Y., Mahajan, H. P., Kim, H. K., Erickson, Z., Rogers, W. A., and Kemp, C. C. (2020). "Active robot-assisted feeding with a general-purpose mobile manipulator: Design, evaluation, and lessons learned." In: *Robotics and Autonomous Systems* 124, p. 103344. DOI: 10.1016/j.robot.2019.103344.

[169] Park, G. and Choi, S. (2018). “Tactile Information Transmission by 2D Stationary Phantom Sensations.” In: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. CHI ’18. Association for Computing Machinery (ACM), pp. 1–12. DOI: 10.1145/3173574.3173832.

[170] Pascher, M., Schneegass, S., and Gerken, J. (2019). “SwipeBuddy: A Teleoperated Tablet and Ebook-Reader Holder for a Hands-Free Interaction.” In: *Human-Computer Interaction – INTERACT 2019*. Ed. by D. Lamas, F. Loizides, L. Nacke, H. Petrie, M. Winckler, and P. Zaphiris. Vol. 11749. Springer, pp. 568–571. DOI: 10.1007/978-3-030-29390-1_39.

[171] Petermeijer, S., Bazilinskyy, P., Bengler, K., and de Winter, J. (2017a). “Take-over again: Investigating multimodal and directional TORs to get the driver back into the loop.” In: *Applied Ergonomics* 62, pp. 204–215. DOI: https://doi.org/10.1016/j.apergo.2017.02.023.

[172] Petermeijer, S., Doubek, F., and Winter, J. de (2017b). “Driver response times to auditory, visual, and tactile take-over requests: A simulator study with 101 participants.” In: *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pp. 1505–1510. DOI: 10.1109/SMC.2017.8122827.

[173] Peternel, L., Tsagarakis, N., Caldwell, D., and Ajoudani, A. (2018). “Robot adaptation to human physical fatigue in human—robot co-manipulation.” In: *Auton. Robots* 42.5, pp. 1011–1021. DOI: 10.1007/s10514-017-9678-1.

[174] Petrich, L., Jin, J., Dehghan, M., and Jagersand, M. (2022). “A Quantitative Analysis of Activities of Daily Living: Insights into Improving Functional Independence with Assistive Robotics.” In: *2022 International Conference on Robotics and Automation (ICRA)*, pp. 6999–7006. DOI: 10.1109/ICRA46639.2022.9811960.

[175] Phan, T.-P., Chao, P. C.-P., Cai, J.-J., Wang, Y.-J., Wang, S.-C., and Wong, K. (2018). “A novel 6-DOF force/torque sensor for COBOTs and its calibration method.” In: *2018 IEEE International Conference on Applied System Invention (ICASI)*, pp. 1228–1231. DOI: 10.1109/ICASI.2018.8394511.

[176] Pollak, A., Paliga, M., Pulopulos, M. M., Kozusznik, B., and Kozusznik, M. W. (2020). "Stress in manual and autonomous modes of collaboration with a cobot." In: *Computers in Human Behavior* 112, p. 106469. DOI: https://doi.org/10.1016/j.chb.2020.106469.

[177] 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." In: *Computers in Industry* 129, p. 103446. DOI: https://doi.org/10.1016/j.compind.2021.103446.

[178] Quere, G., Hagengruber, A., Iskandar, M., Bustamante, S., Leidner, D., Stulp, F., and Vogel, J. (2020). "Shared Control Templates for Assistive Robotics." In: *2020 IEEE International Conference on Robotics and Automation (ICRA)*. Institute of Electrical and Electronics Engineers (IEEE), pp. 1956–1962. DOI: 10.1109/ICRA40945.2020.9197041.

[179] Rahal, R., Matarese, G., Gabiccini, M., Artoni, A., Prattichizzo, D., Giordano, P. R., and Pacchierotti, C. (2020). "Caring About the Human Operator: Haptic Shared Control for Enhanced User Comfort in Robotic Telemanipulation." In: *IEEE Transactions on Haptics* 13.1, pp. 197–203. DOI: 10.1109/TOH.2020.2969662.

[180] Raisamo, J., Raisamo, R., and Surakka, V. (2009). "Evaluating the effect of temporal parameters for vibrotactile saltatory patterns." In: *Proceedings of the 2009 international conference on Multimodal interfaces*. ACM. DOI: 10.1145/1647314.1647381.

[181] Reddy, S., Dragan, A., and Levine, S. (2018). "Shared Autonomy via Deep Reinforcement Learning." In: *Robotics: Science and Systems XIV*. RSS2018. Robotics: Science and Systems Foundation. DOI: 10.15607/rss.2018.xiv.005.

[182] Robinson, F. and Nejat, G. (2022). "An analysis of design recommendations for socially assistive robot helpers for effective human-robot interactions in senior care." In: *Journal of Rehabilitation and Assistive Technologies Engineering* 9. DOI: 10.1177/20556683221101389.

[183] Rohmer, E., Pinheiro, P., Raizer, K., Olivi, L., and Cardozo, E. (2015). "A novel platform supporting multiple control strategies for assistive robots." In: *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, pp. 763–769. DOI: 10.1109/ROMAN.2015.7333663.

[184] Rosen, E., Whitney, D., Phillips, E., Chien, G., Tompkin, J., Konidaris, G., and Tellex, S. (2019). “Communicating and controlling robot arm motion intent through mixed-reality head-mounted displays.” In: *The International Journal of Robotics Research* 38.12-13, pp. 1513–1526. DOI: 10.1177/0278364919842925.

[185] Rozas Llontop, D. A., Cornejo, J., Palomares, R., and Cornejo-Aguilar, J. A. (2020). “Mechatronics Design and Simulation of Anthropomorphic Robotic Arm mounted on Wheelchair for Supporting Patients with Spastic Cerebral Palsy.” In: *2020 IEEE International Conference on Engineering Veracruz (ICEV)*, pp. 1–5. DOI: 10.1109/ICEV50249.2020.9289665.

[186] Ruffaldi, E., Brizzi, F., Tecchia, F., and Bacinelli, S. (2016). “Third point of view augmented reality for robot intentions visualization.” In: *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*. Springer, pp. 471–478.

[187] Safaric, R., Hedrih, I., Klobucar, R., and Sorgo, B. (2003). “Remote controlled robot arm.” In: *IEEE International Conference on Industrial Technology, 2003*. Vol. 2, 1202–1207 Vol.2. DOI: 10.1109/ICIT.2003.1290836.

[188] Sánchez-Calleja, I., Martínez-Gutiérrez, A., Ferrero-Guillén, R., Díez-González, J., and Perez, H. (2023). “Contact System Method for the Precise Interaction Between Cobots and Mobile Robots in Smart Manufacturing.” In: *International Journal of Precision Engineering and Manufacturing*. DOI: 10.1007/s12541-023-00907-3.

[189] Scherer, M. J. (2005). Living in the state of stuck: How assistive technology impacts the lives of people with disabilities. Brookline Books.

[190] Schmidt, A. (2005). “Interactive context-aware systems interacting with ambient intelligence.” In: *Ambient intelligence*. Ed. by G. Riva, F. Davide, F. Vatalaro, and M. Alcañiz. IOS Press. Chap. 9, pp. 159–178.

[191] Schomaker, L. (1995). A taxonomy of multimodal interaction in the human information processing system.

[192] Schreiter, T., Morillo-Mendez, L., Chadalavada, R. T., Rudenko, A., Billing, E., Magnusson, M., Arras, K. O., and Lilienthal, A. J. (2023). “Advantages of Multimodal versus Verbal-Only Robot-to-Human Com-

munication with an Anthropomorphic Robotic Mock Driver." In: *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, pp. 293–300. DOI: 10.1109/RO-MAN57019.2023.10309629.

[193] Selvaggio, M., Cognetti, M., Nikolaidis, S., Ivaldi, S., and Siciliano, B. (2021). "Autonomy in Physical Human-Robot Interaction: A Brief Survey." In: *IEEE Robotics and Automation Letters* 6.4, pp. 7989–7996. DOI: 10.1109/LRA.2021.3100603.

[194] Sharma, A. and Noel, M. M. (2012). "Design of a low-cost five-finger anthropomorphic robotic arm with nine degrees of freedom." In: *Robotics and Computer-Integrated Manufacturing* 28.4, pp. 551–558. DOI: 10.1016/j.rcim.2012.01.001.

[195] Sherrick, C. E. and Rogers, R. (1966). "Apparent haptic movement." In: *Perception & Psychophysics* 1, pp. 175–180.

[196] Sherwani, F., Asad, M. M., and Ibrahim, B. (2020). "Collaborative Robots and Industrial Revolution 4.0 (IR 4.0)." In: *2020 International Conference on Emerging Trends in Smart Technologies (ICETST)*, pp. 1–5. DOI: 10.1109/ICETST49965.2020.9080724.

[197] Shindev, I., Sun, Y., Coovert, M., Pavlova, J., and Lee, T. (2012). "Exploration of Intention Expression for Robots." In: *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*. HRI '12. Association for Computing Machinery (ACM), pp. 247–248. DOI: 10.1145/2157689.2157778.

[198] Shneiderman, B. (2002). Leonardo's Laptop: Human Needs and the New Computing Technologies. MIT Press.

[199] Shneiderman, B. (2022). "Human-centered AI: ensuring human control while increasing automation." In: *Proceedings of the 5th Workshop on Human Factors in Hypertext*. HUMAN '22. Association for Computing Machinery. DOI: 10.1145/3538882.3542790.

[200] Shrestha, M. C., Kobayashi, A., Onishi, T., Uno, E., Yanagawa, H., Yokoyama, Y., Kamezaki, M., Schmitz, A., and Sugano, S. (2016). "Intent communication in navigation through the use of light and screen indicators." In: *2016 11th ACM/IEEE International Conference on Human-*

Robot Interaction (HRI), pp. 523–524. DOI: 10.1109/HRI.2016.7451837.

[201] Sijs, J., Liefhebber, F., and Romer, G. W. R. (2007). "Combined Position & Force Control for a robotic manipulator." In: *2007 IEEE 10th International Conference on Rehabilitation Robotics*. Institute of Electrical and Electronics Engineers (IEEE). DOI: 10.1109/icorr.2007.4428414.

[202] Smith, E. M., Gowran, R. J., Mannan, H., Donnelly, B., Alvarez, L., Bell, D., Contepomi, S., Ennion (Wegner), L., Hoogerwerf, E.-J., Howe, T., Jan, Y.-K., Kagwiza, J., Layton, N., Ledgerd, R., MacLachlan, M., Oggero, G., Pettersson, C., Pousada, T., Scheffler, E., and Wu, S. (2018). "Enabling appropriate personnel skill-mix for progressive realization of equitable access to assistive technology." In: *Disability and Rehabilitation: Assistive Technology* 13.5, pp. 445–453. DOI: 10.1080/17483107.2018.1470683.

[203] Smith, K. H. (2018). "Review of AbleData: Tools & Technologies to Enhance Life." In: *Journal of Consumer Health on the Internet* 22.2, pp. 188–194. DOI: 10.1080/15398285.2018.1451156.

[204] Sousa Leite, E. de, Peixoto Rodrigues, T., Andrade Duarte de Farias, M. d. C., Silva Paredes Moreira, M. A., Gouveia Dias Bittencourt, G. K., Bezerra de Oliveira, F., Albino Simpson, C., and Oliveira Silva, A. (2016). "Influence of Assistive Technology for the maintenance of the Functionality of Elderly People: an integrative review." In: *International Archives of Medicine*. DOI: 10.3823/1892.

[205] Spelmezan, D., Jacobs, M., Hilgers, A., and Borchers, J. (2009). "Tactile Motion Instructions for Physical Activities." In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. CHI '09. Association for Computing Machinery (ACM), pp. 2243–2252. DOI: 10.1145/1518701.1519044.

[206] Statistisches Bundesamt (Destatis) (2022). Disability Facts and Figures – Brief Report 2021. Published online: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Behinderte-Menschen/Publikationen/Downloads-Behinderte-Menschen/sozial-schwerbehinderte-kb-5227101219004.html (last visited on July 19, 2024).

[207] Stephanidis, C. and Salvendy, G. (1998). "Toward an Information Society for All: An International Research and Development Agenda." In: *International Journal of Human-Computer Interaction* 10.2, pp. 107–134. DOI: 10.1207/s15327590ijhc1002_2.

[208] Stephanidis, C., Salvendy, G., Antona, M., Chen, J. Y. C., Dong, J., Duffy, V. G., Fang, X., Fidopiastis, C., Fragomeni, G., Fu, L. P., Guo, Y., Harris, D., Ioannou, A., Jeong, K.-a. (, Konomi, S., Krömker, H., Kurosu, M., Lewis, J. R., Marcus, A., Meiselwitz, G., Moallem, A., Mori, H., Fui-Hoon Nah, F., Ntoa, S., Rau, P.-L. P., Schmorrow, D., Siau, K., Streitz, N., Wang, W., Yamamoto, S., Zaphiris, P., and Zhou, J. (2019). "Seven HCI Grand Challenges." In: *International Journal of Human–Computer Interaction* 35.14, pp. 1229–1269. DOI: 10.1080/10447318.2019.1619259.

[209] Stulp, F., Grizou, J., Busch, B., and Lopes, M. (2015). "Facilitating intention prediction for humans by optimizing robot motions." In: *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Institute of Electrical and Electronics Engineers (IEEE), pp. 1249–1255. DOI: 10.1109/IROS.2015.7353529.

[210] Su, H., Di Lallo, A., Murphy, R. R., Taylor, R. H., Garibaldi, B. T., and Krieger, A. (2021). "Physical human–robot interaction for clinical care in infectious environments." In: *Nature Machine Intelligence* 3.3, pp. 184–186. DOI: 10.1038/s42256-021-00324-z.

[211] Sugandi, B., Toar, H., and Alfianto, A. (2018). "Hand Tracking-based Motion Control for Robot Arm Using Stereo Camera." In: *2018 International Conference on Applied Engineering (ICAE)*, pp. 1–5. DOI: 10.1109/INCAE.2018.8579360.

[212] Suzuki, R., Karim, A., Xia, T., Hedayati, H., and Marquardt, N. (2022). "Augmented Reality and Robotics: A Survey and Taxonomy for AR-Enhanced Human-Robot Interaction and Robotic Interfaces." In: *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems*. CHI '22. Association for Computing Machinery (ACM). DOI: 10.1145/3491102.3517719.

[213] Tan, H. Z., Gray, R., Young, J. J., and Traylor, R. (2003). "A Haptic Back Display for Attentional and Directional Cueing." In: *Haptics-e* 3.1, pp. 1–20.

[214] Tanaka, K., Mu, S., and Nakashima, S. (2014). "Meal-Assistance Robot Using Ultrasonic Motor with Eye Interface." In: *International Journal of Automation Technology* 8.2, pp. 186–192. DOI: 10.20965/ijat.2014.p0186.

[215] Tang, G., Webb, P., and Thrower, J. (2019). "The development and evaluation of Robot Light Skin: A novel robot signalling system to improve communication in industrial human–robot collaboration." In: *Robotics and Computer-Integrated Manufacturing* 56, pp. 85–94. DOI: 10.1016/j.rcim.2018.08.005.

[216] Thielke, S., Harniss, M., Thompson, H., Patel, S., Demiris, G., and Johnson, K. (2012). "Maslow's Hierarchy of Human Needs and the Adoption of Health-Related Technologies for Older Adults." In: *Ageing International* 37.4, pp. 470–488. DOI: 10.1007/s12126-011-9121-4.

[217] Tidoni, E., Abu-Alqumsan, M., Leonardis, D., Kapeller, C., Fusco, G., Guger, C., Hintermuller, C., Peer, A., Frisoli, A., Tecchia, F., Bergamasco, M., and Aglioti, S. M. (2017). "Local and Remote Cooperation With Virtual and Robotic Agents: A P300 BCI Study in Healthy and People Living With Spinal Cord Injury." In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 25.9, pp. 1622–1632. DOI: 10.1109/tnsre.2016.2626391.

[218] Tsui, K. M., Kim, D.-J., Behal, A., Kontak, D., and Yanco, H. A. (2011). ""I want that": Human-in-the-loop control of a wheelchair-mounted robotic arm." In: *Applied Bionics and Biomechanics* 8.1, pp. 127–147.

[219] Tullis, T. and Albert, B. (2013). Measuring the User Experience. Second Edition. Interactive Technologies. Morgan Kaufmann. DOI: https://doi.org/10.1016/B978-0-12-415781-1.00016-9.

[220] Udupa, S., Kamat, V. R., and Menassa, C. C. (2021). "Shared autonomy in assistive mobile robots: a review." In: *Disability and Rehabilitation: Assistive Technology* 18.6, pp. 827–848. DOI: 10.1080/17483107.2021.1928778.

[221] Urman, A., Wang, C.-K., Dankwa-Mullan, I., Scheinberg, E., and Young, M. J. (2018). "Harnessing AI for health equity in oncology research and practice." In: *Journal of Clinical Oncology* 36.30, pp. 67–67. DOI: 10.1200/jco.2018.36.30_suppl.67.

[222] Van Deurzen, B., Rovelo Ruiz, G. A., and Luyten, K. (2024). “A Visual Design Space for One-Dimensional Intelligible Human-Robot Interaction Visualizations.” In: *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction*. HRI ’24. Association for Computing Machinery, pp. 1067–1071. DOI: 10.1145/3610978.3640557.

[223] Van Krevelen, D. and Poelman, R. (2010). “A Survey of Augmented Reality Technologies, Applications and Limitations.” In: *International Journal of Virtual Reality* 9.2, pp. 1–20. DOI: 10.20870/ijvr.2010.9.2.2767.

[224] Varjo (2024). Varjo XR-3. Published online: https://varjo.com/products/varjo-xr-3/ (last visited on July 19, 2024).

[225] Verschueren, G., Nica, W., Roman, J., and Juwet, M. (2023). “Researching Sensors for Dynamic Forces on Cobots.” In: *Flexible Automation and Intelligent Manufacturing: The Human-Data-Technology Nexus*. Springer International Publishing, pp. 286–296. DOI: 10.1007/978-3-031-17629-6_30.

[226] Verza, R., Carvalho, M. L. L., Battaglia, M. A., and Uccelli, M. M. (2006). “An interdisciplinary approach to evaluating the need for assistive technology reduces equipment abandonment.” In: *Multiple Sclerosis Journal* 12.1, pp. 88–93. DOI: 10.1191/1352458506ms1233oa.

[227] Vines, J., Clarke, R., Wright, P., McCarthy, J., and Olivier, P. (2013). “Configuring Participation: On How We Involve People in Design.” In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, pp. 429–438. DOI: 10.1145/2470654.2470716.

[228] Walker, M., Hedayati, H., Lee, J., and Szafir, D. (2018). “Communicating Robot Motion Intent with Augmented Reality.” In: *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*. HRI ’18. Association for Computing Machinery (ACM), pp. 316–324. DOI: 10.1145/3171221.3171253.

[229] Wang, L., Gao, R., Váncza, J., Krüger, J., Wang, X. V., Makris, S., and Chryssolouris, G. (2019). “Symbiotic human-robot collaborative assembly.” In: *CIRP Annals* 68.2, pp. 701–726. DOI: https://doi.org/10.1016/j.cirp.2019.05.002.

[230] Watanabe, A., Ikeda, T., Morales, Y., Shinozawa, K., Miyashita, T., and Hagita, N. (2015). "Communicating robotic navigational intentions." In: *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. Institute of Electrical and Electronics Engineers (IEEE), pp. 5763–5769. DOI: 10.1109/IROS.2015.7354195.

[231] Weber, B., Schatzle, S., Hulin, T., Preusche, C., and Deml, B. (2011). "Evaluation of a vibrotactile feedback device for spatial guidance." In: *2011 IEEE World Haptics Conference*. Institute of Electrical and Electronics Engineers (IEEE), pp. 349–354. DOI: 10.1109/WHC.2011.5945511.

[232] Wolbrecht, E. T., Chan, V., Reinkensmeyer, D. J., and Bobrow, J. E. (2008). "Optimizing Compliant, Model-Based Robotic Assistance to Promote Neurorehabilitation." In: *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 16.3, pp. 286–297. DOI: 10.1109/TNSRE.2008.918389.

[233] World Health Organization (WHO) (2023a). Coronavirus disease (COVID-19). Published online: https://www.who.int/news-room/fact-sheets/detail/coronavirus-disease-(covid-19) (last visited on July 19, 2024).

[234] World Health Organization (WHO) (2023b). Disability & Health Report. Published online: https://www.who.int/news-room/fact-sheets/detail/disability-and-health (last visited on July 19, 2024).

[235] World Health Organization (WHO) and United Nations Children's Fund (UNICEF) (2022). Global report on assistive technology. World Health Organization.

[236] Wu, Y., Chan, W. L., Li, Y., Tee, K. P., Yan, R., and Limbu, D. K. (2015). "Improving human-robot interactivity for tele-operated industrial and service robot applications." In: *2015 IEEE 7th International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM)*, pp. 153–158. DOI: 10.1109/ICCIS.2015.7274565.

[237] Xu, Y., Ding, C., Shu, X., Gui, K., Bezsudnova, Y., Sheng, X., and Zhang, D. (2019). "Shared control of a robotic arm using non-invasive brain–computer interface and computer vision guidance." In: *Robotics and Autonomous Systems* 115, pp. 121–129. DOI: 10.1016/j.robot.2019.02.014.

[238] Yang, C., Wu, H., Li, Z., He, W., Wang, N., and Su, C.-Y. (2018). "Mind Control of a Robotic Arm With Visual Fusion Technology." In: *IEEE Transactions on Industrial Informatics* 14.9, pp. 3822–3830. DOI: 10.1109/TII.2017.2785415.

[239] Yun, H., Lee, J. W., Yang, H. D., and Yang, J. H. (2018). "Experimental Design for Multi-modal Take-over Request for Automated Driving." In: *Communications in Computer and Information Science*. Springer International Publishing, pp. 418–425. DOI: 10.1007/978-3-319-92285-0_57.

[240] Yun, H. and Yang, J. H. (2020). "Multimodal warning design for take-over request in conditionally automated driving." In: *European transport research review* 12.1, pp. 1–11.

[241] Zein, M. K., Al Aawar, M., Asmar, D., and Elhajj, I. H. (2021). "Deep learning and mixed reality to autocomplete teleoperation." In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. Institute of Electrical and Electronics Engineers (IEEE), pp. 4523–4529.

[242] Złotowski, J., Yogeeswaran, K., and Bartneck, C. (2017). "Can we control it? Autonomous robots threaten human identity, uniqueness, safety, and resources." In: *International Journal of Human-Computer Studies* 100, pp. 48–54.

technologies

Article

Adapt or Perish? Exploring the Effectiveness of Adaptive DoF Control Interaction Methods for Assistive Robot Arms

Kirill Kronhardt [1,*,†], Stephan Rübner [1,*,†], Max Pascher [1], Felix Ferdinand Goldau [2], Udo Frese [2] and Jens Gerken [1,*]

[1] Human-Computer Interaction Group, Department of Media Informatics and Communication, Westphalian University of Applied Sciences, 45897 Gelsenkirchen, Germany; max.pascher@w-hs.de
[2] German Research Center for Artificial Intelligence (DFKI) GmbH, 28359 Bremen, Germany; felix.goldau@dfki.de (F.F.G.); udo.frese@dfki.de (U.F.)
* Correspondence: kirill.kronhardt@studmail.w-hs.de (K.K.); stephan.ruebner@studmail.w-hs.de (S.R.); jens.gerken@w-hs.de (J.G.)
† These authors contributed equally to this work.

Abstract: Robot arms are one of many assistive technologies used by people with motor impairments. Assistive robot arms can allow people to perform activities of daily living (ADL) involving grasping and manipulating objects in their environment without the assistance of caregivers. Suitable input devices (e.g., joysticks) mostly have two Degrees of Freedom (DoF), while most assistive robot arms have six or more. This results in time-consuming and cognitively demanding mode switches to change the mapping of DoFs to control the robot. One option to decrease the difficulty of controlling a high-DoF assistive robot arm using a low-DoF input device is to assign different combinations of movement-DoFs to the device's input DoFs depending on the current situation (adaptive control). To explore this method of control, we designed two adaptive control methods for a realistic virtual 3D environment. We evaluated our methods against a commonly used non-adaptive control method that requires the user to switch controls manually. This was conducted in a simulated remote study that used Virtual Reality and involved 39 non-disabled participants. Our results show that the number of mode switches necessary to complete a simple pick-and-place task decreases significantly when using an adaptive control type. In contrast, the task completion time and workload stay the same. A thematic analysis of qualitative feedback of our participants suggests that a longer period of training could further improve the performance of adaptive control methods.

Keywords: assistive robotics; human–robot interaction (HRI); shared user control; augmented reality; virtual reality; visual cues

Citation: Kronhardt, K.; Rübner, S.; Pascher, M.; Goldau, F.; Frese, U.; Gerken, J. Adapt or Perish? Exploring the Effectiveness of Adaptive DoF Control Interacion Methods for Assistive Robot Arms. *Technologies* **2022**, *10*, 30. https://doi.org/10.3390/technologies10010030

Academic Editor: Fillia Makedon

Received: 4 January 2022
Accepted: 10 February 2022
Published: 14 February 2022

1. Introduction

Robotic solutions are becoming increasingly prevalent in many areas of our professional and personal lives and have started to evolve into collaborators [1,2]. A non-negligible number of people live with motor impairments, ranging from slight limitations to severe paralysis [3]. While a near-complete integration into professional and social life is the final goal, current assistive robotic technologies focus on performing activities of daily living (ADLs). These include tasks ranging from essentials such as eating and drinking to more complex behaviors such as grooming and activities associated with leisure time [4].

A general problem with assistive robotic solutions is finding suitable methods and technologies for controlling such robots. Assistive robotic devices are often characterized as having a large number of Degrees of Freedom (high-DoF). For example, a robotic arm with a simple gripper can freely operate in 3D space and move along Cartesian space as well as yaw, pitch, and rotate. This typically results in five to seven DoFs. Standard input devices, such as joysticks, only cover two DoFs. To control a high-DoF device with a low-DoF input device, mode switching is used. This means that at any point in time, the user has to select

a mode, which then maps the two DoFs of the input device to two of the total available DoFs of the robot while neglecting the others. While high-DoF input devices do exist, they are not often accessible for people with motor impairments.

Using a human–computer interface with a standard button-based mode switching setup, Herlant et al. discovered that more than one-sixth of the total execution time is spent changing the currently selected mode [5]. They showed that automatic mode switching leads to increased user satisfaction within a deterministic simulation environment and with a predefined goal.

Our latest research findings provide a proof-of-concept for a novel method of shared control of an assistive robot. We evaluated the idea within a 2D simulation environment [6]. The novel control method uses a Convolutional Neural Network (CNN) to adaptively generate DoF mappings based on camera data of the current situation. From a user perspective, this system can help the user choose an optimal mapping of available control DoFs for a low-DoF input device, either automatically or upon the user's request. In this paper, we build on this approach, focusing in particular on the user interface. Having an adaptive mapping of control DoFs to the input device can be challenging to understand and learn, which is why there is a need for visual feedback to convey that information to the user. The approach in our previous work included visual cues in the form of arrows. While the results are promising (see Section 2), the limitation of a 2D environment means that it is difficult to predict how this approach transfers to 3D. For example, certain DoF combinations might be more difficult to display with arrows in a 3D environment and lead to visual clutter.

The goal of this paper is to explore the proposed novel control method, as well as possible visual cues for the DoF mappings. In particular, we want to explore how the novel, adaptive control method performs in a 3D environment compared to the standard mode-switch approach with cardinal DoF mappings and whether changes in the visual cues have an impact on the performance of the adaptive control method.

We conducted a remote online study with 39 non-disabled participants, in which we compared three different control types with different DoF mapping behaviors and visual cues. These were *Classic* and *Double Arrow*, which used two arrows attached to the fingers as visual cues, and a visually reduced variant *Single Arrow*. *Single Arrow* only used one arrow through the middle of the gripper (see Section 3 for a detailed description of each control type).

The study was conducted inside a 3D Virtual Reality (VR) environment, utilizing Head-Mounted Displays (HMDs) for an immersive experience (see Section 4.3 for a complete description of the virtual environment). The participants repeatedly performed a simple pick-and-place task, controlling a virtual robot arm using the three control types (see Section 4.5 for a detailed description of the study design).

Due to the ongoing COVID-19 pandemic, we opted to recruit non-specific participants that had access to the required hardware (an *Oculus Quest* VR-HMD) to participate in our study. None of the recruited participants reported living with any motor impairments. We acknowledge this limitation and discuss how our findings can be transferred to the target group of people with motor impairments in Section 7.

As our main contribution, we present findings from our study, which compare our two adaptive control types with the standard mode-switch control type, explicitly focusing on task completion times, number of mode switches and workload. In addition, we contribute an extensive discussion of qualitative results from voice recordings of our participants, providing a deeper understanding of the benefits and challenges of each of the three control types.

2. Related Work

To assist people with physical or cognitive impairments, prior research often suggests possible solutions that use robots that automate specific tasks [7–10]. Assistive robots are found in a variety of designs. There are stationary robots specifically designed for

meal-assistance [11], socially assistive robots for elderly people and people with cognitive impairments [12], navigational robots for blind people [13], and many more examples, both in research and commercially. Besides stationary robots (e.g., fixed to a table) [14], there are also moving robots attached to mobile platforms [15,16] or mounted to the user's wheelchair [9].

To help people with motor impairments, assistive robot arms are widely used, both within the workspace and in performing ADLs [17]. Their flexibility allows for many different applications, such as feeding assistance [18], fetch and pick-up tasks [15], and cataloging of books [7].

Robotic assistance is generally well-received by people with motor impairments. Drolshagen et al. found that people with disabilities quickly accept working with robots, even if the robots are in close proximity [19]. Regarding ADLs specifically, Pascher et al. conducted an ethnographic study with 15 participants with tetraplegia, multiple sclerosis, Locked-In Syndrome, and similar diseases [20]. They found that people with motor impairments would prefer to perform ADLs themselves with the help of a robotic aid as opposed to with the help of another person. People with motor impairments want to "live more independently" and "gain increased autonomy".

However, automating ADLs, as suggested in research, can have unintended consequences. Pollak et al. conducted a study comparing manual and autonomous modes of collaboration with a collaborative robot (cobot) [21]. They found that using the manual mode in which the cobot would perform tasks only upon interaction with the participants decreased stress significantly. The participants felt "more capable of coping with and controlling the situation" than in the autonomous mode.

Similarly, Kim et al. conducted a study with subjects with spinal cord injuries using an assistive robot arm in either a manual or an autonomous mode [22]. They found that overall task completion times for manual and autonomous usage for trained participants were similar, but user satisfaction was higher in manual mode. This is despite the fact that autonomous usage decreased the effort necessary to perform tasks significantly. The authors call for more flexible interfaces to control assistive robot arms.

When interacting with robots that carry out movements, a study by Cleaver et al. showed that users generally prefer to have a visual representation of the robot's future movements. However, having this visualization does not significantly affect the performance when executing tasks using the robot [23]. When using a visual representation of robot motion intent, the most prominent solution is to show the robot's movement using arrows [24–26]. In addition, most of these approaches rely on Augmented Reality to overlay the visual representation on the user's real environment.

Heeding the call for more flexible interfaces, we proposed in our recent work an adaptive control concept for assistive robot arms that promises to allow users to be in control at all times while still providing them with more assistance during ADLs than the standard mode switch control concept [6]. In this proposed concept, a CNN interprets the video feed of a camera attached to the robot arm and adaptively outputs the most likely movement DoFs.

With current control concepts, users with low-DoF input devices, such as simple joysticks, can only move the gripper of an assistive robot arm in cardinal directions (i.e., movement and rotation around Cartesian X-, Y-, and Z-Axes). The user has to switch and choose between the provided mappings of input DoFs to some of the robot's DoFs. This may include the pairings of different DoFs of the robot that are less than ideal for the given situation, resulting in many time-consuming and mentally demanding mode switches. Additionally, in any given mode, an input on an axis of a low-DoF device would move the gripper only in the cardinal direction currently assigned to this input DoF. Combinations of multiple output DoFs (such as orbiting an object, which is the combination of rotation and translation) require more than one input DoF (e.g., both the X- and Y-Axes of a joystick) to be engaged simultaneously in such systems.

To solve this problem, we proposed a representation of these assignments of input DoFs to output DoFs in the form of a matrix similar to the one seen in Figure 1 in our previous work. Each row in that matrix represents a cardinal output DoF, while each column represents the input DoFs of an input device. The values in a column determine which movement the robot's gripper will perform if the input DoF is engaged. For example, an identity matrix would yield a behavior identical to the cardinal mode switch approach, as each input DoF is only mapped to one cardinal output DoF.

$$
\begin{array}{c}
X-Axis \\ Y-Axis \\ Z-Axis \\ Roll \\ Pitch \\ Yaw \\ Gripper
\end{array}
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\quad
\begin{bmatrix}
0.5 & 0 & 0 & 0.5 & 0 & 0 & 0 \\
0.5 & 0 & 0 & 0 & 0.5 & 0 & 0 \\
0 & 0 & 0.5 & 0.5 & 0 & 0 & 0 \\
0 & 0.5 & 0 & 0 & 0.5 & 0 & 0 \\
0 & 0.5 & 0 & 0 & 0 & 0.5 & 0 \\
0 & 0 & 0.5 & 0 & 0 & 0.5 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
$$

Figure 1. Two different DoF mappings as matrices—(**Left**): classic control (one input DoF controls one cardinal output DoF); (**Right**): arbitrarily combined controls (one input DoF controls more than one cardinal output DoF at the same time).

This representation allows for combinations of multiple output DoFs for one input DoF. For example, if the first column contains a value of 0.5 in the first two rows, engaging the first input DoF would result in a diagonal movement along the XY plane of the robot's coordinate system (see the matrix on the right in Figure 1). According to the current situation, the proposed control concept adaptively fills this matrix to create the most useful combination of output DoFs.

We then conducted a small study with a 2D proof-of-concept simulation for our proposed control concept. A total of 23 participants used a "standard" and an "adaptive" control type for a simulated 2D robot that could drive forwards, sideways, rotate around its center, and close its fingers to move blue boxes to target red boxes (see Figure 2). This is the 2D equivalent of a simple pick-and-place task in 3D. Both control types switched modes after five seconds without user input.

The results of our study showed that, subjectively, the "adaptive" control was significantly faster but significantly more difficult than the "standard" control. "Adaptive" control also led to significantly shorter sequence execution times.

While these findings are promising, the concept requires further evaluation in 3D and in a more complex environment with devices that have more DoFs. We set out to do precisely that: evaluate the proposed concept of adaptive control in a more complex environment with a robot arm with seven DoFs.

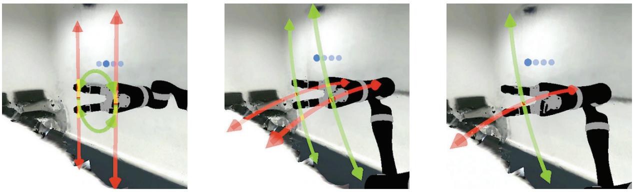
Figure 2. The simulated robot with two out of the four cardinal DoFs (**left**) and two adaptive DoFs (**right**) [6].

3. Control Types for a 3D Environment

To compare the standard control type of switching between cardinal modes to the adaptive approach, we implemented three control types (see Figure 3) in a simulated 3D

environment (see Section 4.3). This simulated environment is meant to act as a proxy for a potential Augmented Reality (AR) implementation. There, users would control an assistive robot arm and see the visual feedback superimposed on the real world and robot via an AR-HMD device. Instead, in our 3D simulation, users wear an *Oculus Quest* VR-HMD, which superimposes the visual feedback directly in the computed 3D scene. An overview of the environment and the control types described in the following sections is provided as a video (see Video S1).

All three control types use arrows as visual cues. Specifically, the arrows show which direction the gripper will move if a user engages the corresponding input DoF. To allow the users to predict the robot's movement when engaging the input DoF with positive values (e.g., pressing the control stick up) and negative values (e.g., pressing the control stick down), the arrows have two heads. Each arrowhead points towards the corresponding movement direction.



(**a**) (**b**) (**c**)

Figure 3. Visualization for the different control types: (**a**) Classic; (**b**) Double Arrow; (**c**) Single Arrow.

Using visual cues in 3D as opposed to 2D often causes visual obstruction, e.g., if the gripper is close to the table, the active DoF would lower the gripper towards the table. In that case, the arrows would clip through the table, making them partially invisible to the user. It would also be common that the robot's gripper itself obstructs parts of the arrows, making them harder to see and interpret. To eliminate these problems, the arrows were made translucent and are always rendered above all other objects yet shown at the correct depth as if looking through whatever is blocking them. This behavior is similar to viewing the scene through Augmented Reality glasses, which would overlay the arrows onto a real scene as opposed to showing the arrows as part of the real world that can be blocked by other real-world objects.

To more easily communicate the currently active mode, all control types show a blue indicator above the robot gripper consisting of four spheres, each representing a mode (see Figure 3). The sphere representing the currently active mode is darker and less translucent than the inactive ones, indicating how many modes are left to switch through before returning to the first.

3.1. Manually Designed DoF-Calculations

The focus of this study was to evaluate how adaptively changing DoF mappings would impact the participant's experience in a more complex 3D environment. While we proposed a CNN to perform these calculations in our previous work [6], there are other ways of calculating these DoF mappings. We developed a manually scripted method of calculating these DoF mappings for the specific task used in the study instead of training a CNN. This method generates a matrix with the same rules described in our previous work (see Figure 1) to represent DoF mapping, thus providing the possibility of equal movements as generated by a CNN trained on camera data. Since our primary focus is the participant's experience with the adaptively changing DoF mappings, we assumed that this approach would significantly decrease the possibility of unpredictable behavior while having little impact on the applicability of our findings to a system using a CNN. A detailed description

of the generated output values is presented in the description of the adaptive control types (see Sections 3.3 and 3.4).

This approach is akin to the widely used "Wizard of Oz" method, in which the output of a proposed system is instead provided by a human to test the user experience of that proposed system before finishing the implementation. In our case, we instead simulated the output of a complex CNN using a simpler system. As with "Wizard of Oz" experiments, our results should therefore be applied to the user experience with the system using a CNN, but the absolute performance measures may vary.

We developed three control types—*Classic*, *Double Arrow*, and *Single Arrow*—to function with different assistive robot arms and different input devices. To conduct the study, we decided to use the widely available stand-alone VR headset *Oculus Quest*. The *Oculus Quest* consists of the headset itself, and two motion controllers, one for each hand, with several buttons and a control stick each. Participants executed a simple pick-and-place task (see Section 4.6) in our VR environment using a virtual model of the *Kinova Jaco* robot arm using each of the control types (see Section 4.3 for a detailed description of the virtual environment and the VR setup).

3.2. Classic Control Type

The *Classic* control type implements the standard mode switch control type most commonly used to control assitive robot arms. This means that an input DoF always corresponds to a cardinal output DoF. Given the seven cardinal DoFs of the Jaco robot arm (X-Translation, Y-Translation, Z-Translation, Roll, Yaw, Pitch, Open/Close fingers) and two input DoFs (the X-Axis and Y-Axis on a motion controller's control stick) four modes are available to the users:

1. X-Translation + Y-Translation;
2. Z-Translation + Roll;
3. Yaw + Pitch;
4. Open/Close fingers + Nothing.

The last mode has no assigned output DoF for the X-Axis on the control stick to allow the users to learn an axis-to-action mapping.

Users can switch modes by pressing the A-Button on the right-hand motion controller. This allows them to perform the tasks at their own pace and assess the usefulness of a mode as long as they need to. Whenever the A-Button is pressed while the fourth mode is active, the first mode is selected again, allowing the users to cycle through modes at will.

Two arrows attached to the fingers of the gripper show the users which motion the gripper would perform, given a user's input on the respective input DoF. Red arrows represent the movement assigned to the Y-Axis of the control stick, and green arrows represent the movement assigned to the X-Axis of the control stick. As the motion controllers are also rendered in the virtual environment, we added a visual representation onto the control stick rendered in-game. A cross with one red axis and one green axis is shown on the motion controller to indicate which direction corresponds to which color. A blue sphere surrounds the A-Button to match it to the blue mode indicator (see Figure 4).

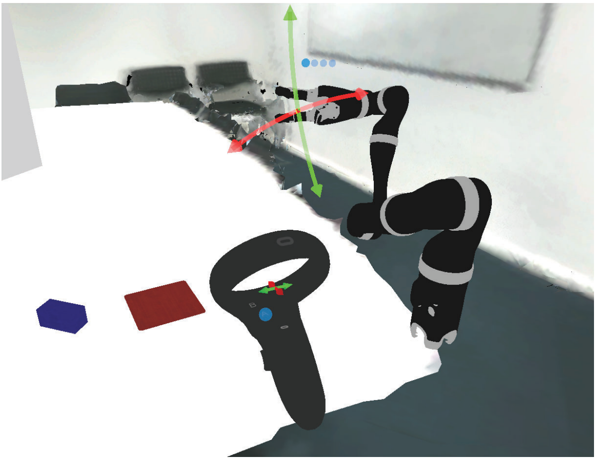


Figure 4. The virtual motion controller with directional indicators and the robot arm with matching arrows visualizing the currently selected mode.

3.3. Double Arrow Control Type

The *Double Arrow* control type implements the proposed adaptive control method using two arrows to show the position of the fingers if a user engages an input DoF. Therefore, each input DoF corresponds to a combination of cardinal DoFs determined based on the current situation. To ensure comparability with the *Classic* control type in regards to the number of mode switches necessary to return to the starting mode, four modes were developed. The modes are ordered by their complexity and usefulness to the users' goal of reaching the next target.

As in the *Classic* control type, two arrows attached to the fingers of the gripper show the users which motion the gripper would perform, given a user's input on the respective input DoF. Red arrows represent the movement assigned to the Y-Axis of the control stick, and green arrows represent the movement assigned to the X-Axis of the control stick.

The first mode assigns the Y-Axis of the control stick to a movement that both rotates and translates the gripper towards the next target simultaneously. More precisely, if the gripper is further than 10 cm away from the target, the movement is oriented towards a point 15 cm above the target. If the gripper is closer than 10 cm to the target, the movement is oriented towards the actual target. This ensures that the gripper tends to grasp and let go of objects from above, as opposed to trying to do so from the sides and thereby possibly crashing into the table. If the gripper is within reach of an object or target point where an object is supposed to be placed by the users, it also allows them to open and close the fingers. The X-Axis of the control stick in the first mode is assigned the same movement as the Y-Axis but rotated by 90° to allow for corrections perpendicular to the Y-Axis movement.

To provide users with more options, the second mode assigns the Y-Axis of the control stick to a linear translational movement towards the object and the X-Axis of the control stick to a rotational movement of the gripper towards the next target. Both of these assignments were chosen since only moving or only rotating are less likely to further the goal of the users. However, the individual movements themselves are still integral movements for coordinating the gripper orientation and some movement towards the goal. In the optimal case, this means that users would not need to use this mode, as both orientation and positioning would be taken care of simultaneously by the first mode.

The third mode assigns the Y-Axis of the control stick to the opening or closing the fingers, depending on whether an object was currently held or not. The X-Axis of the control stick has no assignment in this mode to ensure comparability with the *Classic* control type.

If users stop moving the gripper, they should always be able to move the same way they did before. To ensure this, the fourth mode always assigns the X- and Y-Axis of the control stick the same mappings that were last used to move the gripper. Otherwise, users who would want to assess if they had moved the robot far enough for their personal preference using a given mapping would have no possibility to correct their course.

The system calculates the next movement mappings whenever the users stop moving the robot. However, the system does not instantly assign the first mode to be active, as this would disrupt the users' flow of control (i.e., they might have stopped to asses the situation and then decided to continue with the DoF mapping they were using). Moreover, this would harm comparability to the *Classic* control type (as no automatic mode switches happen in that control type). This means that whenever the users stop moving, the blue mode indicator would show the fourth mode as being active, and a press on the A-Button would lead to the newly calculated first mode.

3.4. Control Type Single Arrow

During the development of *Classic* and *Double Arrow* we discovered that, while two arrows are a perfectly suitable visualization for a 2D environment, these arrows can result in a large amount of visual clutter during complex movement in 3D environments. We decided to develop a visualization that reduces visual clutter in a 3D environment and compare its usage to the *Double Arrow* control type.

Dubbed *Single Arrow*, the input-to-output DoF mappings are calculated in the exact same way as the mappings in *Double Arrow*. Switching between modes is also handled in the same way as in *Double Arrow*. However, the visualization changes from displaying two arrows at the tips of the fingers to displaying one arrow in the middle of the gripper, with a slight offset to allow certain movements to be displayed. This reduces visual clutter for all situations except when the fingers move.

4. Materials and Methods

We present a remote study with 39 participants to compare the proposed concept of adaptive control (in two variations) against the standard mode-switch control concept. In particular, we measured task completion times, the number of mode switches necessary to perform a task, the workload necessary to use the different control concepts via a NASA Raw-TLX (NASA Raw Task Load Index), and the participants' personal ranking of the three presented control types. Participants used their own *Oculus Quest* headset to perform a simple pick-and-place task using a virtual robot inside a realistic 3D environment.

4.1. Hypotheses

We propose the following hypotheses:

- Average Task Completion Time
  - *H1* Double Arrow *leads to lower task completion time than* Classic. The adaptive control of *Double Arrow* should significantly reduce the movements necessary to perform the task by combining different cardinal DoFs into one continuous movement, which otherwise would each have to be adjusted separately.
  - *H2* Single Arrow *leads to lower task completion time than* Double Arrow. Only using one arrow for each DoF mapping should reduce visual clutter. This should lead to a shorter processing time of the suggested movements, reducing the total time to execute a task.
- Average Number of Mode Switches

- *H3* Double Arrow *leads to fewer mode switches than* Classic*.* The adaptive control of *Double Arrow* should reduce the necessity to switch modes significantly. Since different DoFs are combined depending on the current situation, a change in position and rotation brings the robot arm closer to the target and can be performed without mode switches.
- *H4* Single Arrow *and* Double Arrow *need roughly an equal number of mode switches.* The behavior of the two adaptive control types is the same. Thus, while it might take participants longer to understand what movements they can perform with *Double Arrow* as opposed to *Single Arrow*, they should switch modes approximately as often in both control types.

• Workload
- *H5* Double Arrow *leads to lower NASA TLX scores than* Classic*.* The adaptive control of *Double Arrow* calculates sensible movements to reach the next goal position and rotation. Thus, it should alleviate the participants from having to think of a sequence of movements to reach their goal, reducing workload. This is in contrast to the findings of our previous study, in which participants perceived the *Adaptive* control as more complex than the *Standard* control [6]. We expect the benefit of pre-calculated DoF combinations and the workload of developing a sequence of movements in cardinal DoFs to be higher in a 3D environment than in a 2D environment. Therefore, the workload for the adaptive control types should be lower than for *Classic* in 3D.
- *H6* Single Arrow *leads to lower NASA TLX scores than* Double Arrow*.* Since we assume that reduced visual clutter leads to a shorter processing time for the suggested movements, the NASA TLX scores of *Single Arrow* should be lower.

4.2. Participants

In total, 39 people participated in our study (12 female, 26 male, 1 non-binary), which led to a data-set of 936 individual trials (8 per control type, 24 per participant). The age of participants ranged from $\leq$19 to 69, with 20 to 29 being the largest group with 22 participants. Four participants had prior experience with controlling an assistive robot arm, and no participants declared any motor impairments. All participants received EUR 10 as compensation unless they specifically denied the offer.

Due to the ongoing COVID-19 pandemic, we opted to perform a remote study using VR. We did not specifically search for participants with motor impairments because the potential target audience for people with VR setups at home that also have motor impairments appeared too small. There would not be enough time to gather enough participants in a realistic time frame. Instead, we searched for any participants that had access to the necessary equipment (an *Oculus Quest* headset, see Section 4.3) and were able to install our study software on their devices. With these non-specific participants, the performance measures for executing the tasks in our study with the different control types (see Section 4.6) can be compared relative to one another, even though they may not be representative of the intended target audience of such an assistive device. We acknowledge this limitation, which is further discussed in Section 7.

Participants were recruited via announcements in different social media communities relating to VR (e.g., r/OculusQuest: https://www.reddit.com/r/OculusQuest/, accessed on 3 January 2022), social media communities regarding assistive technologies (e.g., r/AssistiveTechnologies: https://www.reddit.com/r/AssistiveTechnology/, accessed on 3 January 2022), and platforms for acquiring participants specifically for XR studies (e.g., XRDRN: https://www.xrdrn.org/, accessed on 3 January 2022) among other more local announcements.

To ensure that VR sickness symptoms did not influence our results, the participants filled out the Virtual Reality Sickness Questionnaire (VRSQ) at the end of the study [27]. The VRSQ measures nine items on a four-point Likert scale and results in a value between 0 and 100, where 0 means no symptoms experienced and 100 means all symptoms were severe.

Reported values were low (Mean: 11.30, Std.-Dev.: 11.38), and none of the participants selected the "Severe" option for any of the items.

4.3. Apparatus

We designed a Virtual Reality environment based on a photogrammetry scan of a real room. The environment included a virtual model of the *Kinova Jaco* (Kinova Jaco robot arm: https://assistive.kinovarobotics.com/product/jaco-robotic-arm, accessed on 3 January 2022) robot arm attached to a table, a red target surface, a blue block, and two virtual screens—one for descriptions and questionnaires and one that would show example photos of the control types (see Figure 5). We decided to use a virtual model of a real robot arm (*Kinova Jaco*) to stay as close to a physical system as possible. Additionally, the *Kinova Jaco* robot arm is specifically designed and often used as an assistive device for people with motor impairments [5].

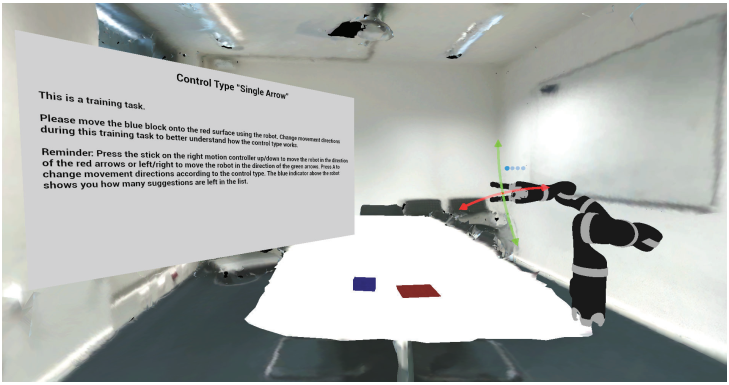


Figure 5. The virtual environment: description screen (**Left**); screen with example photos of the control types (not shown); *Kinova Jaco* with visualisation for control type *Single Arrow* (**Right**); table with blue block and red target (**Bottom**).

The virtual environment was created with the *Unreal Engine 4.26* and was developed to be deployed to the *Oculus Quest* VR headset. Participants had to either own or have access to such a headset and be able to install the study software on that headset using a computer (*Windows*, *macOS*, and *Linux* could be used). Although we tested our software on the original *Oculus Quest* hardware, we did not explicitly exclude the use of the newer and very similar *Oculus Quest 2* headset. The *Oculus Quest* consists of the VR headset and two motion controllers, one for each hand. Each motion controller has several buttons and a control stick. Participants controlled the robot using the right motion controller of the VR headset. In particular, the control stick of the motion controller moved the robot according to the currently active control type. This enabled the participants to control which DoFs were being used and how fast the robot would move. The A-Button was used to switch to the next mode cyclically, returning to the first mode when a mode switch was performed in the last mode.

To simulate the movement of the robot arm, the inputs did not move the joints of the robot as they would with a physical robot arm. Rather, the gripper of the virtual robot arm is moved in 3D space according to the inputs, and the arm of the robot is programmed to adopt a correct pose automatically. This was implemented using the physics system of the *Unreal Engine*.

4.4. Procedure

Participants were directed to a website with a brief introduction to the study, the duration of the study (around 30 to 45 min), the technical and non-technical prerequisites to participate in the study, and a description about what data would be collected during the study. Participants were informed that certain metrics and usage data, such as task completion times, will be recorded and sent to our servers during the study. They were also informed that they would need to fill out a short questionnaire after each condition of the study and that they would be able to record a short audio message after each condition. Lastly, participants were informed that cookies were being used on our website. Each participant gave informed consent by pressing a clearly labeled button to continue and start the study. After giving their consent, participants were instructed on how to install and open the study application and what to do when they were finished with the part of the study inside the VR headset. During the study, neither a video of the participants surroundings through the VR headsets external cameras and sensors nor a screen-recording was captured.

Next, the participants put on their VR headsets and opened our study application. They were greeted with a brief explanation of the study on a large virtual screen. Except for the questionnaires after each control type, any text that was available to read on that screen was also simultaneously read aloud as a prerecorded voice-over. The participants interacted with this screen via a common interaction method that was also used in the menus of the *Oculus Quest* headset: pointing a ray that originated from the motion controller towards the screen and using the trigger to confirm input.

After the study explanation, the participants were presented with a description of the first control type they would be using and the task they would be performing. This explanation was supplemented with an image on a second smaller virtual screen. The descriptions were written in a way that described how the gripper would move in relation to the current situation. We did not explicitly describe the intentions behind the different modes and their order in *Double Arrow* and *Single Arrow* (to provide ideally optimal mappings) to prevent possible biases. Otherwise, the participants might have been inclined to trust the adaptive mappings against their own judgment, thereby changing their behavior.

The explanation of each control type was followed by a series of trials of our pick-and-place task (see Section 4.6) the participants had to execute to progress through the study. For each control type, the task was performed once as a training trial and then eight more times for the same control type. During these eight trials, the task completion time and the number of mode switches performed was recorded.

After executing all trials for a control type, the participants were presented with the NASA Raw-TLX questionnaire to capture the participants' workload. Additionally, the participants could record a short audio message to point out additional things they felt were relevant during the execution of the trials. The recording of the audio message was optional. After filling out the questionnaire and optionally recording an audio message, the participants would continue with the next control type until they had executed all trials for all three control types.

Upon finishing the VR part of the study, participants received a unique code to be entered in a form on our website to complete the VRSQ [27] and our questionnaire. We asked the participants to report their demographic data and rank the control types presented in the VR section of the study. Lastly, participants left their contact information to receive the compensation.

4.5. Study Design

We used a within-subjects design with the control type as an independent variable with three levels: (1) *Classic*, (2) *Double Arrow*, and (3) *Single Arrow*. Each participant performed eight trials of a pick-and-place task for each of the three control types (see Section 4.6). Additionally, they performed one training trial for each control type to

familiarize themselves with the control type. The order of control types shown to the participants was fully balanced.

We measured three dependent variables for each control type: *Average Task Completion Time*, *Average Number of Mode Switches*, and *Workload via a NASA Raw-TLX questionnaire*.

Average Task Completion Time in seconds While participants executed each trial with the robot arm, the time to complete the task was measured for each participant. Then, the average task completion time for each control type was calculated across all participants.
Average Number of Mode Switches While participants executed each trial with the robot arm, each mode switch executed by pressing a button on the input device was counted and stored as the number of mode switches. Then, the average number of mode switches for each control type was calculated across all participants.
Workload via a NASA Raw-TLX questionnaire After completing all trials within each control type, the participants were asked to fill out a NASA Raw-TLX questionnaire to obtain information about the participants' perceived workload. The questionnaire consists of the following six criteria, which participants would rate on a scale of 0 to 100 in steps of 5: mental demand, physical demand, temporal demand, performance, effort, and frustration [28].

In addition, the participants could record a short description of their experiences in the form of a voice message, although this was not mandatory. The recorded voice messages were transcribed and analyzed by multiple researchers to identify underlying themes and common impressions the participants had while using the virtual robot arm (see Section 5.2). Participants also provided a personal ranking of the three control types in a questionnaire at the end of the study.

4.6. Task

Participants were asked to repeatedly place a blue block onto a red target using the assistive robot. Participants performed this task eight times per control type. We did not use two blocks per trial to reduce variability in our results. We decided to use a simple pick-and-place task instead of a specific ADL (e.g., drinking from a glass) since pick-and-place tasks are part of many ADLs. Moreover, a specific ADL might have caused problems with participants' preconceived notions of that task (e.g., they would approach the glass in a particular way, while the adaptive system would approach it differently). This would have possibly distracted them from evaluating the control types as a whole, which we wanted to avoid.

In each of the eight trials per control type, the position of the blue block changed to one of eight predefined positions around the red target surface. The order in which the positions were used in the eight trials was randomized for each participant and control type.

5. Results

We recorded both quantitative and qualitative data from the participants during the trials. This section presents the results of each section from our data analysis.

5.1. Quantitative Results

The recorded quantitative data for each trial included *task completion time* (in seconds) and *the number of mode switches*. For each control type, the quantitative data included the *NASA Raw-TLX* results and the *Rank* given to the control type by the participants (lower rank numbers are better). The used abbreviations and symbols are:

- IQR: Interquartile Range;
- SD: Standard Deviation;
- SE: Standard Error;
- p: p-value as an expression of the level of statistical significance;
- N: Sample Size;
- $\chi^2(2)$: Chi-Squared with two degrees of freedom;

- F: F-Statistic for the Repeated-Measures ANOVA;
- M: Mean;
- df: Degrees of Freedom for the calculation of χ^2 for the Friedman Tests.

5.1.1. Task Completion Time

For each participant, we averaged the task completion times (see Table 1) of the trials for each control type. In an exploratory analysis, we removed outliers that had average task completion times $\geq 2.2 * \text{IQR}$ of the mean task completion time in at least one control type [29] (see Figure 6). Four outliers were excluded this way, leaving 35 participants for analysis of task completion times. An inspection of QQ-plots found the resulting data-set to follow a normal distribution.

Table 1. Statistics for average task completion times (in seconds, N = 35).

	Classic	Double Arrow	Single Arrow
Mean	47.41	42.62	44.04
Median	44.66	37.75	41.23
Std.-Dev.	12.55	19.28	22.24
IQR	14.03	24.33	31.68

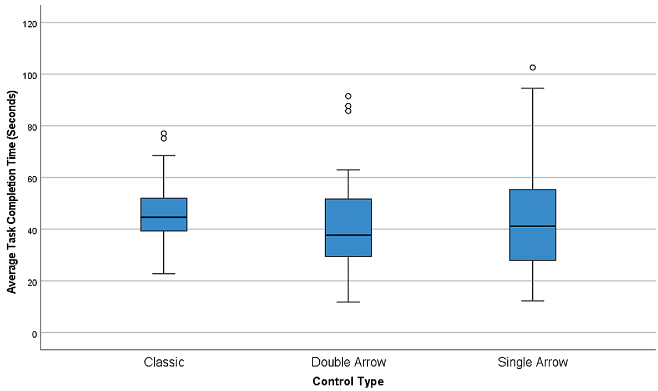


Figure 6. Boxplots for average task completion times.

To determine whether the control types had an effect on average task completion times, we performed a Repeated-Measures ANOVA (RM-ANOVA). However, we found no significant main effect ($F(2, 64) = 1.31$, $p = 0.28$).

In addition to the effect of control types, we examined whether the starting condition of a participant had an impact on task completion times. We included the starting condition as a between-subjects factor for the RM-ANOVA and discovered a significant interaction effect between the starting condition and the task completion times ($F(4, 64) = 8.86$, $p < 0.001$). Analyzing simple main effects, we discovered that the task completion times for *Classic* stayed roughly the same regardless of the starting condition. However, both adaptive control types heavily suffered when they were the starting condition (see Figure 7). A post hoc pairwise comparison (Estimated Marginal Means, Bonferroni adjusted) showed that task completion times for *Single Arrow* (M = 54.66 s, SE = 5.9) were significantly longer than those for *Double Arrow* (M = 33.74 s, SE = 4.6) if *Single Arrow* was the starting condition ($p = 0.001$). Conversely, task completion times for *Double Arrow* (M = 57.89 s, SE = 5) were significantly longer than those for *Single Arrow* (M = 37.82 s, SE = 6.41) if *Double Arrow* was the starting condition instead ($p = 0.002$). Another significant difference was found if *Single*

Arrow was the starting condition: *Classic* task completion times (M = 48.23 s, SE = 3.57) were longer than those of *Double Arrow* (M = 33.74 s, SE = 4.6) in that case ($p = 0.013$). The other comparisons yielded insignificant results.

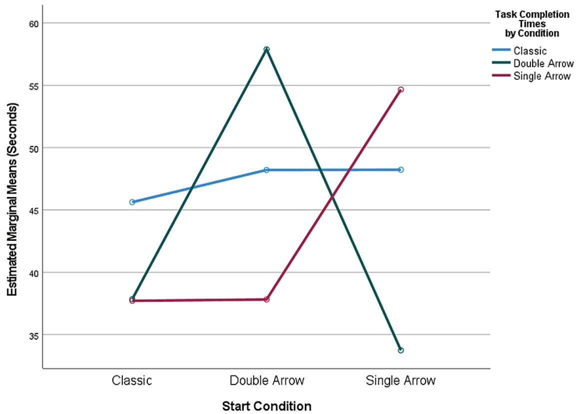


Figure 7. Estimated Marginal Means for average task completion times.

5.1.2. Mode Switches

To determine whether there were differences between the average number of mode switches between control types we used an RM-ANOVA. Due to a software error, mode switch data were only recorded correctly for 20 participants. We found a significant effect of control types on the average number of mode switches ($F(2, 38) = 8.08$, $p = 0.001$). Pairwise comparisons revealed that there were significant differences ($p < 0.05$) between the average number of mode switches for both adaptive control methods (*Double Arrow*: M = 12.93, SD = 3.91; *Single Arrow*: M = 14.23, SD = 5.15) and the *Classic* control method (M = 17.87, SD = 4.8). We found no significant difference between the average number of mode switches for *Single Arrow* compared to *Double Arrow* ($p = 0.11$, see Table 2 and Figure 8).

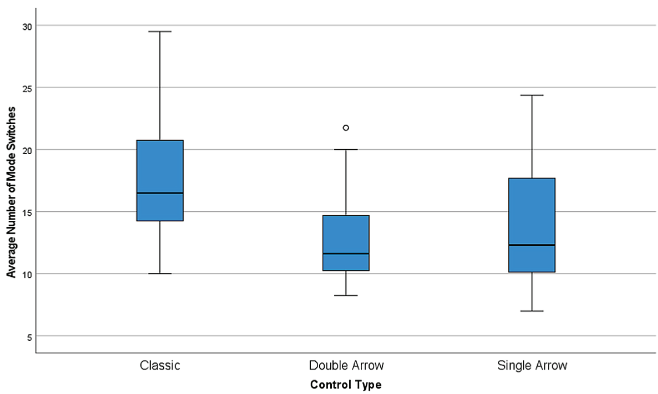


Figure 8. Boxplots for average number of mode switches.

Table 2. Statistics for average number of mode switches (N = 20).

	Classic	Double Arrow	Single Arrow
Mean	17.87	12.93	14.23
Median	16.50	11.63	12.31
Std.-Dev.	4.80	3.91	5.15
IQR	7.00	5.09	7.91

5.1.3. Workload and Rank

Each participant completed a NASA Raw-TLX questionnaire after completing the task with each control type, rating each dimension on a scale from 1 to 100. To evaluate whether there were any differences between the control types regarding workload, Friedman Tests were performed for both the overall NASA TLX value as well as the individual dimensions of the questionnaire. No significant differences were found for either the overall NASA TLX value ($\chi^2(2) = 5.33$, $p = 0.07$) or the individual dimensions (see Table 3).

We also evaluated whether the users preferred one control type over the others. To do so, the participants ranked the control types after completing all tasks. A lower number means the participant ranked that control type higher. No significant differences were found for the ranks ($\chi^2(2) = 0.97$, $p = 0.65$) (see Table 4).

Table 3. Statistics for individual NASA TLX Dimensions on a scale from 1 to 100 (df = 2, N = 39 for all Friedman Tests).

	Mental Demand	Physical Demand	Temporal Demand	Performance	Effort	Frustration
Classic (Mean)	53.33	30.26	36.92	32.05	48.59	41.41
Classic (Std.-Dev.)	24.64	21.67	21.07	20.48	24.84	24.52
Double Arrow (Mean)	56.28	28.21	40.38	38.97	52.82	43.08
Double Arrow (Std.-Dev.)	22.93	16.20	25.06	25.50	24.08	26.40
Single Arrow (Mean)	48.97	27.56	36.03	40.64	51.41	38.33
Single Arrow (Std.-Dev.)	24.69	22.94	20.56	26.61	23.25	26.34
Mean Ranks						
Classic	2.04	1.92	1.96	1.73	1.79	1.92
Double Arrow	2.21	2.17	2.18	2.15	2.18	2.17
Single Arrow	1.76	1.91	1.86	2.12	2.03	1.91
Friedman Tests						
χ^2	4.23	2.07	2.38	4.86	3.15	1.76
Exact Significance	0.12	0.37	0.31	0.09	0.21	0.43

5.2. Qualitative Results

Participants were asked to describe their experience with the control type they used in a voice message. They were asked to elaborate on the ease of controlling the robot, their understanding of movement directions, and the predictability of the next movement directions.

In total, 23 of the 39 participants recorded a message for all three control types. In addition, only four participants recorded voice messages for two of the three control types, and one participant just recorded a single voice message. This resulted in 26 voice messages for each control type.

Table 4. Statistics for NASA TLX on a scale from 1 to 100 and ranking on a scale from 1 to 3 (df = 2, N = 39 for all Friedman Tests).

	NASA TLX	Rank
Classic (Mean)	40.43	1.87
Classic (Std.-Dev.)	17.11	0.77
Double Arrow (Mean)	43.29	2.05
Double Arrow (Std.-Dev.)	15.32	0.86
Single Arrow (Mean)	40.49	2.08
Single Arrow (Std.-Dev.)	17.29	0.84
Mean Ranks		
Classic	1.85	1.87
Double Arrow	2.29	2.05
Single Arrow	1.86	2.08
Friedman Tests		
χ^2	5.33	0.97
Exact Significance	0.07	0.65

5.2.1. Thematic Analysis

The voice recordings were analyzed with the Thematic Analysis method described by Braun and Clarke [30]. This method was chosen because it has the flexibility to identify themes within the unstructured feedback from the recorded voice messages. Throughout the analysis, we identified themes related to our hypotheses, which gave us a better insight into how participants perceived their experience and success in executing the given tasks.

First, we transcribed the voice messages to be able to analyze them. Although most participants recorded their messages in English, a few recorded them in German. Some of the statements in the following chapters were therefore translated into English. Second, two of our researchers performed the Thematic Analysis using the six-phase method described by Braun and Clarke [30]. Each researcher read each transcribed voice message to become familiar with the participant's feedback. They then marked certain paragraphs and phrases to identify underlying topics related to our hypotheses that were relevant within multiple data-sets. Each marked phrase was assigned a short code describing its topic. We used the software *Obsidian* (Obsidian markdown note-taking software: https://obsidian.md, accessed on 3 January 2022) for managing and tagging the transcribed messages in a simple markdown text format with links and tags. Third, codes were organized and grouped into themes, and descriptive titles were assigned to each theme. For a visual representation, we developed visual thematic graphs; one of which is shown in Figure 9. Although some comments were related to several themes, we decided to sort them into the theme with the best fit. Fourth, themes were revised and evaluated by reading the related phrases and codes again to ensure that each theme was internally homogeneous. Fifth, both researchers worked together to refine the themes and compile them into a single thematic map presented in Figure 10. Sixth, a summary of the results was written based on the final thematic map.

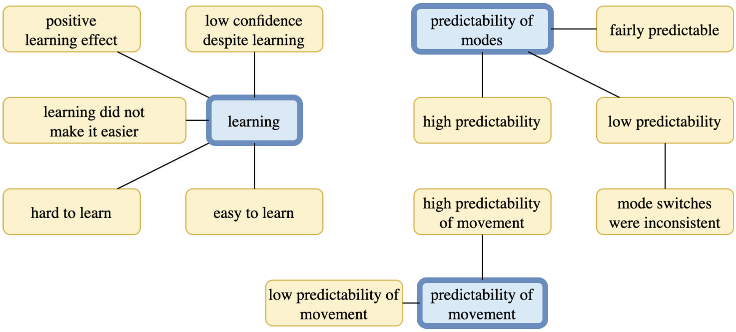


Figure 9. Early thematic map with codes shown in yellow and themes shown in blue.

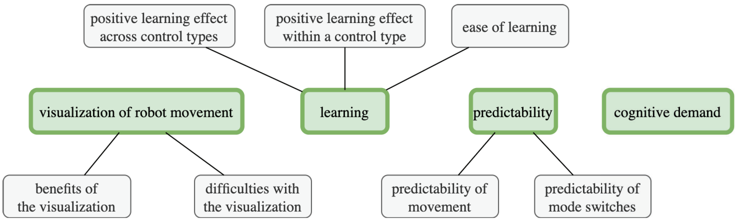


Figure 10. Final thematic map with themes shown in green and sub-themes shown in gray.

5.2.2. Results of the Thematic Analysis

We identified the following themes in the combined thematic map: *visualization of robot movement*, *cognitive demand*, *predictability of mode switching*, *predictability of movement* and *learning*. The excerpts from one participant's audio messages were marked with the participant's unique number (e.g., *P26* for the 26th participant out of the total 39 participants). Since participants often referenced the previous control types they used, we also added which control type they were referring to in brackets when citing them.

Visualization of robot movement: This theme comprises the difficulties and benefits of the visualization of the robot's movement. As expected when transferring over a visualization from a 2D environment to 3D, perspective was one source of errors across all three control types. *P5* stated, "Depending on the orientation of the robot arm, I could not see exactly which way the arrows were going." P4 added, "Sometimes moving the robot was a bit difficult because it just did not feel natural from different perspectives."

Regarding the control types *Double Arrow* and *Single Arrow*, many participants mentioned that the arrows are either hard to interpret or hard to see. Interestingly, the participants did not mention this problem with the *Classic* control type. Participants stated, "[. . .] the arrows that follow the change of the movement direction are a little more difficult and a little bit less intuitive to understand than the previous trial [control type *Classic*]" (*P9*), and "I think it is more difficult than the previous control type because it has more abstract movement [. . .]" (*P25*). Besides the curved arrows, many participants found it difficult to associate the differently colored arrows of the visualization with the different input DoFs across all three control types. *P31* made this clear after using the *Single Arrow* control type. They said, "The hardest part working with this method of motion was determining which direction pushing the analog stick would actually move the robot."

Across both adaptive control types, participants mentioned the helpfulness of the arrows. *P25* commented, "I think it was confusing at first, but those red and green arrows helped a lot to understand how the robot moved." After using the *Double Arrow* control

type, *P8* mentioned, “Controlling the robot was better than before [control type *Single Arrow*], because one could tell more easily where the arm would go, based on the multiple arrows”. This suggests the possible benefits of having multiple arrows in the *Double Arrow* control type.

Cognitive demand: In this theme, we consolidated statements that describe a higher or lower cognitive demand while using a specific control type. Across all three control types, some participants mentioned a high cognitive demand. After using the *Single Arrow* control type, *P17* stated, “This one was more cognitively demanding than the previous one [control type *Classic*], maybe because this one did not have straight movement but a lot of rotational movements”. *P18* found it to be “a bit confusing, but okay.” Participants described the *Classic* control type as “confusing” (*P8*) and “counter intuitive” (*P18*). Using the *Double Arrow* control type, *P21* expressed the need to focus on the task and added, “I do not think you could do anything else while using this control method”.

While mentions of lower cognitive demand were equally frequent in total, many participants found the *Classic* control type to be “easy” or “easy to understand” (*P6*, *P9*, *P25*, *P27*, among others). After using the *Classic* control type, *P39* added, “Here it was best to intuitively remember where each function was”. This suggests a connection with the next two themes regarding predictability.

Predictability of mode switching: This theme describes the ability of the participants to anticipate the next set of movement combinations that the system provides when the participant executes a mode switch. Many of the difficulties participants had with the predictability were with the adaptive control types *Single Arrow* and *Double Arrow*. When using the *Double Arrow* control type, *P17* noted, “In this condition, I was not sure whether cycling through the different types of movements in there always were consistent. That was very confusing.” We also identified this statement as an expression of an increase in cognitive demand. For the same control type, *P21* added, “I did not know which combination would be next when I pressed A”. Using the *Double Arrow* control type, *P23* mentioned, “I could not predict the next movement, because I did not understand in which order the different movements are shown to me next.” We think this participant confused the ever-changing nature of the adaptive suggestions with the different modes. Only a few participants mentioned difficulties with predicting the next mode in the *Classic* control type. *P37* said, “Predictability was uncertain as well, until the later moves where I had enough training to do it effectively.” Additionally, many participants mentioned that they had to switch modes many times to find the proper movement they needed in a given situation, especially with the adaptive control types. Using the *Double Arrow* control type, *P3* stated, “So if I wanted it to go down I would have to switch through multiple modes [. . .]”. Furthermore, using the *Double Arrow* control type, *P5* mentioned, “I had to click through many modes to find the movement that I thought would bring me closer to the block”.

Mentions of good predictability were also spread across all three control types, although these were less common. For the *Classic* control type, *P39* stated, “It was very easy to understand and especially the predictability was the easiest here”. Using the *Single Arrow* control type, *P37* mentioned, “The ease of understanding the movement was a lot easier as well. With some of the movement directions being easier to understand and predict before they show up.” After executing the tasks with the *Double Arrow* control type, *P37* added, “It seemed more predictable and overall, a more optimum way of doing things”.

Predictability of movement: In contrast to the previous theme, this theme is about predicting how and where the robot arm will move when using the currently selected mode. As visualization plays a big part when predicting the robot’s movement, this theme is related to the first theme about visualization. Only a few participants mentioned the predictability of movement directly. After using the *Double Arrow* control type, *P4* said, “So I tried to do one thing and it would do a completely other thing. It felt really unnatural to try and get to the cube and even to pick it up”. For the *Classic* control type, *P10* stated, “Because of the immediate predictability [. . .], it was much easier to control the robot and

to steer it into different vectors to approach the block in the different positions". Using the *Single Arrow* control type, *P10* added, "Therefore I could understand very well how it would move and how it would work out so I could reach the target".

Learning: This theme describes the participants' impression of their learning experience while using the different control types. Across all three control types, participants reported that they grew better at performing the tasks over time. For the *Classic* control type, *P26* stated, "Using this robot arm is pretty easy if you learn how to use them, [. . .]". After using the *Double Arrow* control type, *P25* mentioned, "The predictability of the next movement directions, I think, is easier as you practice with it, [. . .]". For the *Single Arrow* control type, *P39* said, "The more I practiced, the more confidence I got [. . .]".

As participants used the different control types, they noticed a learning effect even across the different control types. After finishing all trials of all control types, ending with the *Double Arrow* control type, *P25* said, "The predictability of the next movement directions, I think, is easier as you practice with it, [. . .]" After using the *Single Arrow* control type, *P33* stated, "Maybe I simply have more experience now, if I performed better in this task in any way".

Even though many participants felt that they needed more practice with the tasks so that they are easier to perform, some described that the process of learning felt relatively easy. When finishing the tasks with the *Classic* control type, *P16* stated, "It was quicker to get familiar with the system." *P33* expressed some difficulties with the *Double Arrow* control type but added, "At least it did not take long to notice a learning effect".

Additionally, we identified many instances where participants reported that they liked the second adaptive control type they used better than the one before, regardless of which control type came first and which came second. This also suggests that a learning effect is taking place. After using the *Double Arrow* and then the *Single Arrow* control types, *P27* stated, "I don't know what is the difference between double arrow and single arrow, but single arrow is much easier to control". For the *Double Arrow* control type, *P31* stated, "This method is a little bit easier to use than the second method [*Single Arrow* control type], but I think that was more a function of having a little bit more experience".

6. Discussion

Initially, our assumptions were that the overall task performance would be best when using the *Single Arrow* control type, followed by *Double Arrow*, and *Classic* would have the worst task performance. In comparison to the results of our previous study [6], the new results are not as pronounced in a realistic virtual 3D setting, at least not without considering the learning effects.

Regarding the task completion times, both Hypothesis 1 and Hypothesis 2 could not be substantiated. However, the interaction effect between the starting condition and task completion times suggests that, with time to learn, the adaptive control types could perform better than the *Classic* type. This is corroborated by participants' reports, as many participants said that their performance and understanding of the adaptive control types improved during the tasks. It is also worth noting that more participants experienced the second adaptive control type as "better" than the first, implying a learning effect not only for one control type but between control types.

Regarding mode switches, Hypothesis 3 and Hypothesis 4 could be substantiated by our results. From *Classic* to *Double Arrow*, we measured a significant reduction in the number of mode switches necessary to perform the task. In contrast, there was no significant difference between *Double Arrow* and *Single Arrow*. Interestingly, this contrasts the participants' opinions that they felt they had to switch many times to get to a mode that performed a movement they expected. However, this reduction in mode switches might be of higher benefit for people with motor impairments than for non-disabled people. Switching modes using a button requires a certain level of dexterity and causes the user to constantly divert their attention away from the original task, so more mode switches can cause more fatigue and time consumption, as explained by Herlant et al. [5]. The impact

of this difference in the number of mode switches on people with motor impairments can thus only be evaluated in a future study with participants with motor impairments.

Regarding workload, Hypothesis 5 and Hypothesis 6 could not be substantiated. This could have multiple reasons. For example, the participants expressed that the predictability of the adaptive control types was low and that they did not necessarily know how the robot would move, even with the arrows. These impressions, combined with the statements regarding positive learning effects and overall high cognitive demand, could mean that with increased exposure to the adaptive control types, users could have a lower workload than with *Classic*.

According to some participants, using visual cues in a 3D environment caused problems with perspective. This made it difficult for them to predict how the robot would move, even with the visual cues provided by the arrows. To mitigate this problem, our concept might be combined with a "digital twin" of the robot arm, which demonstrates the movement virtually before the real robot performs it physically [31].

To improve the overall predictability of the system, both regarding the suggested modes and the movements of the robot, a training mode could be implemented. In this mode, the users would be able to teach the system the way they want specific tasks to be performed [32]. This should increase predictability, as the participants would know the proposed movements will be (partially) based on their own instructions. In addition, Spatial Augmented Reality can help the user's understanding of the robot's perception, e.g., which object the robot assumes the user wants to interact with [33]. In combination with the already implemented visual cues, this can help the users predict the robot's movement more accurately.

After further research and refinement of our proposed control methods, they might allow assistive robot arms to help with ADLs that currently require the help of caregivers or more complex robots, such as dressing [34] or bathing [35]. The fact that the users always stay in control of the robot while the robot performs more fluent, natural movements could also allow people with motor impairments to use the robot in social situations, e.g., at the workplace [36].

7. Limitations

Our study did not specifically involve or focus on people with motor impairments. Thus, we need to discuss how our results can be transferred to this target group. First, the absolute performance measures cannot be generalized to this target group. Individual differences are usually high within people with motor impairments due to varying degrees of physical limitations [37]. However, the study did not aim to provide absolute results in terms of performance but rather an insight into the relative performance of the three different control types. Since they all rely on the same physical interaction concept, we believe that the way motor impairments might affect performance should be comparable for all three control types. Second, Augmented Reality is necessary to provide the user with the type of visual feedback we implemented for our study. We are aware from our prior research that current-generation AR-HMDs are often not accessible to people with motor impairments. AR-HMDs such as the *Microsoft HoloLens* are too heavy and conflict too often with health-supporting systems [38]. We conducted this research with the firm belief that future AR hardware solutions will cope with requirements for people with motor impairments. We acknowledge, however, that this might make the visual feedback designs inapplicable for real-world systems at this point in time or the immediate future.

Additionally, our study involved the use of the *Oculus Quest* system and the *Oculus Quest* Motion Controller as the only input device. In the real world, however, assistive robot arms can be controlled with a wide range of input devices depending on the abilities and preferences of the person using them. We specifically only used the most basic functionality of the Motion Controller (the control stick and one button) to ensure that the results are also applicable when using a different input devices with two input axes. It is still possible

that the use of different input devices might add more complexity to the overall usage of such a system.

Another limitation is the nature of our study being performed as a remote study. The level of control is limited for such a method, which means that the level of engagement of participants can vary. We addressed this limitation by keeping the duration of the study relatively short (30–45 min) and designing the task so that we could easily identify cases in which participants did not follow the study protocol. Our analysis further shows that only a few participants were identified as extreme outliers. In addition, the focus on one set of hardware devices made it possible to harmonize and control the kind of immersive experience that participants engaged with, further reducing potential biasing effects, such as low frame rates or other hardware-performance-related issues. Given the current COVID-19 pandemic, we believe that our study setup is sensible and still able to provide robust results. Still, we aim to replicate at least part of the study in a lab environment and with people with motor impairments in the future.

It is possible that our study does not provide insight into the quality of adaptive control through the means of a CNN. We simulated the adaptive control method to be able to have full control in the study. Otherwise, imperfect DoF mappings would have overshadowed the potential effects of the different visualizations, thus making it difficult to draw conclusions. As discussed, we believe that our approach significantly decreases the possibility of unpredictable behavior while having little impact on the applicability of our findings to a system using a CNN, as long as this CNN is able to perform at a high level of quality regarding the DoF mappings.

8. Conclusions

We conducted a study exploring and evaluating the user experience of an adaptive control concept for assistive robot arms in a realistic virtual 3D environment. Our results suggest a significant benefit of such an adaptive control concept regarding the necessary number of mode switches. However, task completion times and workload do not change when using an adaptive control concept without more intensive training.

By evaluating the interaction between the starting conditions and task completion times and applying a thematic analysis of qualitative data, we conclude that there could be a significant benefit of training that would reveal the potential of an adaptive control concept. Therefore, future work should consider longer training sessions before evaluating task completion times and workload. The targeted user group of assistive robot arms would use such devices not just once but daily and over extended periods and thus have more time to learn how to use the device. Therefore it is important to assess whether the adaptive control concept might have high cognitive demand in the beginning but is better than the *Classic* approach once the users are trained.

Our results seem to suggest that there is little to no difference between *Single Arrow* and *Double Arrow* regarding how well they convey the robots currently active DoF mapping to the users. However, an improved visualization could reduce the overall high cognitive demand users have experienced. Therefore, future work will also focus on different types of visualizations, which will not be restricted to MR-headsets and overlayed arrows but could (additionally) show the robot's future path using spatial Augmented Reality [39].

Future work should (whenever possible) include participants with motor impairments since their experience is vital in designing assistive technology [4]. The impact of a lower number of mode switches enabled by an adaptive control concept should be especially evaluated with people with motor impairments. This could significantly improve their execution of activities of daily living.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/technologies10010030/s1. Video S1: An Overview of the Environment and Control Types.

Author Contributions: Conceptualization, K.K., S.R., M.P., F.F.G., U.F. and J.G.; methodology, K.K., S.R., M.P. and J.G.; software, K.K. and S.R.; validation, M.P. and J.G.; formal analysis, K.K. and S.R.; investigation, K.K., S.R. and M.P.; resources, M.P., F.F.G. and J.G.; data curation, K.K. and S.R.; writing—original draft preparation, K.K. and S.R.; writing—review and editing, M.P., F.F.G., U.F. and J.G.; visualization, K.K. and S.R.; supervision, M.P., U.F. and J.G.; project administration, U.F. and J.G.; funding acquisition, U.F. and J.G.; original study, F.F.G. and U.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the German Federal Ministry of Education and Research (BMBF, FKZ: 16SV8563 and 16SV8565).

Institutional Review Board Statement: Ethical reviews by an Ethics Committee are waived in Germany as they are not compulsory for fields outside medicine. Only the participant's consent was requested.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bauer, A.; Wollherr, D.; Buss, M. Human—Robot Collaboration: A Survey. *Int. J. Hum. Robot.* **2008**, *5*, 47–66. [CrossRef]
2. Fong, T.; Nourbakhsh, I.; Dautenhahn, K. A survey of socially interactive robots. *Robot. Auton. Syst.* **2003**, *42*, 143–166.
3. Statistisches Bundesamt (Destatis). Pflegestatistik—Pflege im Rahmen der Pflegeversicherung—Deutschlandergebnisse—2019. 2020. Available online: https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Pflege/Publikationen/Downloads-Pflege/pflege-deutschlandergebnisse-5224001199004.pdf (accessed on 3 January 2022).
4. Pascher, M.; Baumeister, A.; Schneegass, S.; Klein, B.; Gerken, J. Recommendations for the Development of a Robotic Drinking and Eating Aid—An Ethnographic Study. In Proceedings of the Human-Computer Interaction—INTERACT 2021, Bari, Italy, 30 August–3 September 2021; Ardito, C., Lanzilotti, R., Malizia, A., Petrie, H., Piccinno, A., Desolda, G., Inkpen, K., Eds.; Springer: Cham, Switzerland, 2021. [CrossRef]
5. Herlant, L.V.; Holladay, R.M.; Srinivasa, S.S. Assistive Teleoperation of Robot Arms via Automatic Time-Optimal Mode Switching. In Proceedings of the Eleventh ACM/IEEE International Conference on Human Robot Interaction, Christchurch, New Zealand, 7–10 March 2016; pp. 35–42.
6. Goldau, F.F.; Frese, U. Learning to Map Degrees of Freedom for Assistive User Control: Towards an Adaptive DoF-Mapping Control for Assistive Robots. In Proceedings of the 14th PErvasive Technologies Related to Assistive Environments Conference, Corfu, Greece, 29 June–2 July 2021; Association for Computing Machinery: New York, NY, USA, 2021; PETRA 2021, pp. 132–139. [CrossRef]
7. Gräser, A.; Heyer, T.; Fotoohi, L.; Lange, U.; Kampe, H.; Enjarini, B.; Heyer, S.; Fragkopoulos, C.; Ristic-Durrant, D. A Supportive FRIEND at Work: Robotic Workplace Assistance for the Disabled. *IEEE Robot. Autom. Mag.* **2013**, *20*, 148–159. [CrossRef]
8. Foresi, G.; Freddi, A.; Monteriù, A.; Ortenzi, D.; Pagnotta, D.P. Improving mobility and autonomy of disabled users via cooperation of assistive robots. In Proceedings of the 2018 IEEE International Conference on Consumer Electronics (ICCE), Las Vegas, NV, USA, 2–14 January 2018; pp. 1–2. [CrossRef]
9. Tsui, K.M.; Kim, D.J.; Behal, A.; Kontak, D.; Yanco, H.A. "I want that": Human-in-the-loop control of a wheelchair-mounted robotic arm. *Appl. Bionics Biomech.* **2011**, *8*, 127–147. [CrossRef]
10. Goldau, F.F.; Shastha, T.K.; Kyrarini, M.; Gräser, A. Autonomous multi-sensory robotic assistant for a drinking task. In Proceedings of the 2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR), Toronto, ON, Canada, 24–28 June 2019; pp. 210–216.
11. Tanaka, K.; Mu, S.; Nakashima, S. Meal-assistance robot using ultrasonic motor with eye interface. *Int. J. Autom. Technol.* **2014**, *8*, 186–192. [CrossRef]
12. McColl, D.; Louie, W.Y.G.; Nejat, G. Brian 2.1: A socially assistive robot for the elderly and cognitively impaired. *IEEE Robot. Autom. Mag.* **2013**, *20*, 74–83. [CrossRef]
13. Guerreiro, J.; Sato, D.; Asakawa, S.; Dong, H.; Kitani, K.M.; Asakawa, C. Cabot: Designing and evaluating an autonomous navigation robot for blind people. In Proceedings of the 21st International ACM SIGACCESS Conference on Computers and Accessibility, Pittsburgh, PA, USA, 28–30 October 2019; pp. 68–82.
14. Maheu, V.; Archambault, P.S.; Frappier, J.; Routhier, F. Evaluation of the JACO robotic arm: Clinico-economic study for powered wheelchair users with upper-extremity disabilities. In Proceedings of the 2011 IEEE International Conference on Rehabilitation Robotics, Zurich, Switzerland, 29 June–1 July 2011; pp. 1–5.
15. Hashimoto, K.; Saito, F.; Yamamoto, T.; Ikeda, K. A field study of the human support robot in the home environment. In Proceedings of the 2013 IEEE Workshop on Advanced Robotics and its Social Impacts, Tokyo, Japan, 7–9 November 2013; pp. 143–150.
16. Fattal, C.; Leynaert, V.; Laffont, I.; Baillet, A.; Enjalbert, M.; Leroux, C. Sam, an assistive robotic device dedicated to helping persons with quadriplegia: Usability study. *Int. J. Soc. Robot.* **2019**, *11*, 89–103. [CrossRef]

17. Kyrarini, M.; Lygerakis, F.; Rajavenkatanarayanan, A.; Sevastopoulos, C.; Nambiappan, H.R.; Chaitanya, K.K.; Babu, A.R.; Mathew, J.; Makedon, F. A survey of robots in healthcare. *Technologies* **2021**, *9*, 8. [CrossRef]
18. Canal, G.; Alenyà, G.; Torras, C. Personalization framework for adaptive robotic feeding assistance. In Proceedings of the International Conference on Social Robotics, Kansas City, MO, USA, 1–3 November 2016; Springer: Berlin/Heidelberg, Germany, 2016; pp. 22–31.
19. Drolshagen, S.; Pfingsthorn, M.; Gliesche, P.; Hein, A. Acceptance of industrial collaborative robots by people with disabilities in sheltered workshops. *Front. Robot. AI* **2021**, *7*, 541741. [CrossRef] [PubMed]
20. Pascher, M.; Baumeister, A.; Klein, B.; Schneegass, S.; Gerken, J. Little Helper: A Multi-Robot System in Home Health Care Environments. In Proceedings of the 2019 International Workshop on Human-Drone Interaction (iHDI) as Part of the ACM Conference on Human Factors in Computing Systems, Glasgow, UK, 5 May 2019.
21. Pollak, A.; Paliga, M.; Pulopulos, M.M.; Kozusznik, B.; Kozusznik, M.W. Stress in manual and autonomous modes of collaboration with a cobot. *Comput. Hum. Behav.* **2020**, *112*, 106469. [CrossRef]
22. Kim, D.J.; Hazlett-Knudsen, R.; Culver-Godfrey, H.; Rucks, G.; Cunningham, T.; Portee, D.; Bricout, J.; Wang, Z.; Behal, A. How Autonomy Impacts Performance and Satisfaction: Results From a Study With Spinal Cord Injured Subjects Using an Assistive Robot. *IEEE Trans. Syst. Man Cybern.—Part A Syst. Hum.* **2012**, *42*, 2–14. [CrossRef]
23. Cleaver, A.; Tang, D.V.; Chen, V.; Short, E.S.; Sinapov, J. Dynamic Path Visualization for Human-Robot Collaboration. In Proceedings of the Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (HRI '21 Companion), Boulder, CO, USA, 8–11 March 2021; Association for Computing Machinery: New York, NY, USA, 2021; pp. 339–343. [CrossRef]
24. Walker, M.; Hedayati, H.; Lee, J.; Szafir, D. Communicating Robot Motion Intent with Augmented Reality. In Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction (HRI '18), Chicago, IL, USA, 5–8 March 2018; Association for Computing Machinery: New York, NY, USA, 2018; pp. 316–324. [CrossRef]
25. Shrestha, M.C.; Kobayashi, A.; Onishi, T.; Uno, E.; Yanagawa, H.; Yokoyama, Y.; Kamezaki, M.; Schmitz, A.; Sugano, S. Intent communication in navigation through the use of light and screen indicators. In Proceedings of the 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Christchurch, New Zealand, 7–10 March 2016; pp. 523–524. [CrossRef]
26. Shindev, I.; Sun, Y.; Coovert, M.; Pavlova, J.; Lee, T. Exploration of Intention Expression for Robots. In Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI '12), Boston, MA, USA, 5–8 March 2012; Association for Computing Machinery: New York, NY, USA, 2012; pp. 247–248. [CrossRef]
27. Kim, H.K.; Park, J.; Choi, Y.; Choe, M. Virtual reality sickness questionnaire (VRSQ): Motion sickness measurement index in a virtual reality environment. *Appl. Ergon.* **2018**, *69*, 66–73. [CrossRef] [PubMed]
28. Hart, S.G. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2006**, *50*, 904–908. [CrossRef]
29. Hoaglin, D.C.; Iglewicz, B. Fine-Tuning Some Resistant Rules for Outlier Labeling. *J. Am. Stat. Assoc.* **1987**, *82*, 1147–1149. [CrossRef]
30. Braun, V.; Clarke, V. Using thematic analysis in psychology. *Qual. Res. Psychol.* **2006**, *3*, 77–101. [CrossRef]
31. Park, K.B.; Choi, S.H.; Lee, J.Y.; Ghasemi, Y.; Mohammed, M.; Jeong, H. Hands-free human–robot interaction using multimodal gestures and deep learning in wearable mixed reality. *IEEE Access* **2021**, *9*, 55448–55464. [CrossRef]
32. Kyrarini, M.; Zheng, Q.; Haseeb, M.A.; Gräser, A. Robot learning of assistive manipulation tasks by demonstration via head gesture-based interface. In Proceedings of the 2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR), Toronto, ON, Canada, 24–28 June 2019; pp. 1139–1146.
33. Pascher, M.; Kronhardt, K.; Franzen, T.; Gruenefeld, U.; Schneegass, S.; Gerken, J. My Caregiver the Cobot: Comparing Visualization Techniques to Effectively Communicate Cobot Perception to People with Physical Impairments. *Sensors* **2022**, *22*, 755. [CrossRef]
34. Zhang, F.; Cully, A.; Demiris, Y. Probabilistic real-time user posture tracking for personalized robot-assisted dressing. *IEEE Trans. Robot.* **2019**, *35*, 873–888. [CrossRef]
35. Zlatintsi, A.; Dometios, A.; Kardaris, N.; Rodomagoulakis, I.; Koutras, P.; Papageorgiou, X.; Maragos, P.; Tzafestas, C.S.; Vartholomeos, P.; Hauer, K.; et al. I-Support: A robotic platform of an assistive bathing robot for the elderly population. *Robot. Auton. Syst.* **2020**, *126*, 103451. [CrossRef]
36. Arévalo Arboleda, S.; Pascher, M.; Lakhnati, Y.; Gerken, J. Understanding human-robot collaboration for people with mobility Impairments at the Workplace, a thematic analysis. In Proceedings of the 2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), Naples, Italy, 31 August–4 September 2020; pp. 561–566.
37. Arévalo Arboleda, S.; Becker, M.; Gerken, J. Does One Size Fit All? A Case Study to Discuss Findings of an Augmented Hands-Free Robot Teleoperation Concept for People with and without Motor Disabilities. *Technologies* **2022**, *10*, 4. [CrossRef]
38. Arévalo Arboleda, S.; Pascher, M.; Baumeister, A.; Klein, B.; Gerken, J. Reflecting upon Participatory Design in Human-Robot Collaboration for People with Motor Disabilities: Challenges and Lessons Learned from Three Multiyear Projects. In Proceedings of the 14th PErvasive Technologies Related to Assistive Environments Conference—PETRA 2021, Corfu, Greece, 29 June 2021–2 July 2021. [CrossRef]
39. Watanabe, A.; Ikeda, T.; Morales, Y.; Shinozawa, K.; Miyashita, T.; Hagita, N. Communicating robotic navigational intentions. In Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Hamburg, Germany, 28 September–2 October 2015; pp. 5763–5769. [CrossRef]

Recommendations for the Development of a Robotic Drinking and Eating Aid - An Ethnographic Study

Max Pascher[1,2]( ), Annalies Baumeister[3], Stefan Schneegass[2], Barbara Klein[3], and Jens Gerken[1]

[1] Westphalian University of Applied Sciences, Gelsenkirchen, Germany
max.pascher@w-hs.de
[2] University of Duisburg-Essen, Essen, Germany
[3] Frankfurt University of Applied Sciences, Frankfurt am Main, Germany

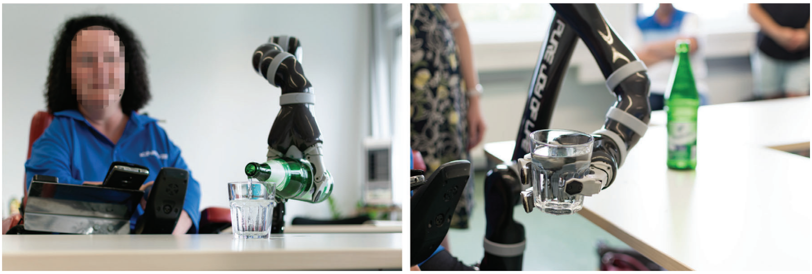
Abstract. Being able to live independently and self-determined in one's own home is a crucial factor or human dignity and preservation of self-worth. For people with severe physical impairments who cannot use their limbs for every day tasks, living in their own home is only possible with assistance from others. The inability to move arms and hands makes it hard to take care of oneself, e.g. drinking and eating independently. In this paper, we investigate how 15 participants with disabilities consume food and drinks. We report on interviews, participatory observations, and analyzed the aids they currently use. Based on our findings, we derive a set of recommendations that supports researchers and practitioners in designing future robotic drinking and eating aids for people with disabilities.

Keywords: Assisted living technologies · Human-centered computing · Meal assistance · Participation design · People with disabilities · Robot assistive drinking · Robot assistive feeding · User acceptance · User-centered design · User participation

1 Introduction

At the end of 2019, 7.9 million people classed as severely disabled were living in Germany [47]. With over 58% of these cases being attributed to physical disabilities, motor impairments affected an total of 4.6 million people; 11.2% of which are suffering from impaired functionality to a complete loss of motor control of their extremities. Additionally a further 10.4% were also affected by impairments in the spinal and torso region.

Published by Springer Nature Switzerland AG 2021
C. Ardito et al. (Eds.): INTERACT 2021, LNCS 12932, pp. 331–351, 2021.
https://doi.org/10.1007/978-3-030-85623-6_21



(a) Filling a glass of water (b) Grasping a glass to drink

Fig. 1. Robotic arms can support users with motor impairments in their everyday drinking and eating task. We explore how such systems should be designed to provide a benefit to the users and support them in living a self-determined life.

Functional loss of the use of extremities can be caused by upper spinal cord trauma and degenerative diseases. Those afflicted are struggling, or are simply unable, to perform every day tasks independently of others. One very prominent area is the one of nutrition. Being self-sufficient in terms of being in control of food and water intake is not only beneficial to ones health but also immensely important for ones self-worth [30].

Assistive technologies are increasingly becoming a vital factor in the field of assisted living; minimising the need for constant care and allowing people with motor impairments to regain some independence [35]. Initial studies by Klein [24] and Merkel and Kurcharski [34] indicated that assistive technology often meets non-acceptance and non-use and propose that devices need to focus more on the needs and preferences of the target group. Using a participatory approach integrating future users in the developing progress is recommended to promote a higher acceptance of the final product [50].

We conducted an ethnographic study in this work to shed light on how users envision future systems supporting them with everyday drinking and eating tasks. We interviewed 15 users with motor impairments, presented a robotic aid as a potential assistive system, and conducted in-situ observations of their drinking and eating behavior and used tools. We gained significant insight into user opinions and derived recommendations regarding structural, social, and collaborative concerns of future assistive systems like a robotic drinking aid (cf., Fig. 1. These recommendations will help designers and engineers in a technology-focused domain to build systems that actually help people.

2 Related Work

Traditionally the focus in the field of developing assistive technologies has been on functionality from an engineering point of view. Recent findings however highlight the need to include future users and their perceived needs in the design process [26]. In this section we first examine previous work done on concepts of user participation and collaborative approaches. In a second step, we present projects that already analysed the use of robotic devices to support people with disabilities and how these aids are valued by their users.

2.1 User Needs

In recent years, there has been growing interest in the concept of user participation in the design of new assistive technologies. Groundwork laid by Thielke et al. [48] and Merkel and Kucharski [34] expressed the need for this collaborated approach to maximize user acceptance. They indicated various methods for integrating the user group as well as family, caregivers and assistants into the innovation process. Focus groups, qualitative interviews, visits of the primary users' homes, and participant observation can provide significant insights into the needs and wants of the user group. The recommendation for this participatory approach that integrates the future users in the developing progress is also noted by Frennert and Östlund [17] and Eftring and Frennert [13], confirming the findings by Klein and Merkel and Kurcharski.

During the development of a robotic therapy support system, Duckworth et al. used three different methods to include the future users preferences into their work [12]. Clinicians and patients were interviewed, given a questionnaire concerning the design of a robotic therapy support system and had the opportunity to use the developed robot during counselling sessions. They came to the conclusion that a participatory design provides essential information for the development of assistive technology and increases the chance of a positive user experience.

Using a similar approach, Mandy et al. conducted a qualitative study with users of the *Neater Eater* to gain an in-depth understanding of their user experience [29]. They report that self-feeding devices increase the life quality of people with disabilities significantly and support a more equal relationship between those who are in need of care and their carers. They stress the need of a positive approach towards assistive technologies for a wide general acceptance.

2.2 Human-Robot Collaboration in the Field of Supporting People with Disabilities

Robotic solutions can make a significant contribution to regaining independence and improving care by supporting and relieving caregivers, thus improving the quality of life of those in need of support [5].

A growing body of literature has examined the impact of assistive robotic systems in supporting people with motor disabilities. Work done by Chen et al. [9] for the *Robots for Humanity* project and Fattal et al. [14] looked into the feasibility and acceptance of robotic systems as assistive technologies. A common finding was that the robotic devices are often designed to assist with several activities of daily living. These devices are usually large; consisting of a robotic arm on a mobile module. They require a barrier-free environment and rooms with sufficient space to fit into and be able to move around safely. In contrast, Pascher et al. noted the potential of smaller, lightweight solutions designed for individual tasks [36], indicating that a specialized aid would be more accessible in terms of size and portability.

Research by Gallenberger et al. used camera and machine learning for an autonomous robotic feeding system to detect types of food items present and to plan the picking-up and transportation to the mouth of the user [18]. An alternative approach is presented by Canal et al. describing a learning-by-demonstration framework to feed the user [8]. Both projects ensure the ability of the robotic arm to fulfill its autonomous tasks without any fine-control of the user focusing on the technical aspects of the development process of assistive technology.

A 2019 study by Beaudoin et al. focused on the long-term use of the robotic arm JACO [4], a recent advance in assistive technologies. They researched improvements of everyday task capabilities, satisfaction with JACO, psychological impact and the implications for users and their caregivers using a similar quantitative approach as employed in this study. Beaudoin et al. reported that almost all participants gained more autonomy in certain life aspects and experienced a number of positive psycho-social impacts. One such success was the increased capability of participants to drink independently of human support using JACO, thus reducing the amount of care and attention needed and increasing well-being and overall health by having a continuous access to beverages.

Interaction technologies such as gaze-based interaction and head movement have been explored to operate, e.g. a PC [11,38,40] and a robot [22,41,45]. Alternatively, brain-computer interfaces were used to control a robotic arm [1]. However, today's ubiquitous technology interaction scenarios are much more tightly integrated in everyday activities and require different interaction interfaces [28].

3 Study

The goal of this work is to understand users' requirements and demands of assistive technology that supports them with drinking and eating. For this, we conducted an ethnographic study consisting of an interview including a VR presentation of a robotic support system and in-situ participatory observations of their drinking and eating habits, with 15 participants afflicted by a varying degree of motor impairments.

3.1 Participants

In preparation for the main study, we opted to evaluate our methods with a pilot participant allowing us to adapt the study design before approaching the remaining participants. Participants were chosen in collaboration with the *Center for Paraplegic Patients Hamburg*, the *Locked-in-Syndrom e.V. Berlin*, and the *State Association of the German Society for Multiple Sclerosis Hessen e.V.* We recruited 15 participants with a permanent and significant degree of compromised mobility of the extremities and the reliance on support for the consumption of food and drinks. Table 1 presents the participants split by gender, age, and diagnosis. 4 female and 11 male participants took part in the main study; the mean age was 42.07 years (SD = 16.68) and all were categorized as severely disabled.

Table 1. Overview of the pilot and main study participants

ID	Gender	Age	Diagnosis
Pilot	Female	60	Multiple sclerosis
P1	Male	18	Spinal cord injury; incomplete at level C3
P2	Male	46	Spinal cord injury; complete at level C4 & some rudimentary mobility until level C5
P3	Male	41	Spinal cord injury; incomplete at level C3 (right body-side has some mobility until level C5)
P4	Male	30	Spinal cord injury; incomplete at level C3
P5	Female	62	Locked-in syndrom
P6	Male	50	Spinal cord injury; incomplete at level C4
P7	Male	38	Spinal cord injury; incomplete at level C3 & complete at level C5
P8	Male	30	Spinal cord injury; complete at level C3
P9	Male	22	Spinal cord injury; complete from level C3 to C7
P10	Male	48	Spinal cord injury; complete at level C4 & C5
P11	Female	60	Multiple sclerosis
P12	Male	50	Inclusion body myositis
P13	Female	51	Locked-in syndrom
P14	Male	34	Spinal cord injury; complete at level C5 & C6
P15	Female	51	Arthrogryposis

3.2 Procedure

Each session took place in the participants' homes which allowed us to conduct the interview, observation of drinking and eating habits as well as analysis of commonly used aids in a natural setting. In most of the cases a caregiver or assistant was present.

Interview. Due to the nature of the physical impairments faced by the participants, obtaining their consent had to be adapted to their particular capabilities. After reading or listening to a researcher reading the consent form, participants signed the form by themselves or had their spoken agreement recorded. In other cases, an authorized caregiver signed the form on behalf of the participant.

The interview part was structured in four sections; each focusing on a different aspect detailing their living situation, attitudes regarding drinking and eating, level of assistance needed as well as wishes towards an ideal robotic aid. In the first part we aimed to understand their current living situation by establishing how many hours they spend in their wheelchair, where they spend most of the time, and where they eat and drink at home.

Next, we were interested in their value propositions and preferences regarding drinking and eating. The participants were asked to describe a typical mealtime routine, what they generally consume, and which preparations are needed.

Further, we wanted to know if drinking and eating is seen as a necessary task or can also convey enjoyment. Participants were also asked if they consume food and drinks if they are not at home (at work, in a restaurant).

The third step focused on the process of drinking and eating in an assistive setting including the communication with their caregivers/assistants and any improvised aids used.

In the final step of the session, we focused on the use of a proposed robotic arm as a drinking and eating aid. To familiarize participants with the concept they were shown images of different eating support systems and wheelchair-extension-type robotic arms that are already on the market, e.g. iEat [3], Obi [10], JACO [23], and iArm [2]. To simulate the situation of sitting in front of an actual robotic arm performing tasks in a close-contact environment we used Google Cardboard [20] and a stereoscopic video of our in-lab robot setting. Conducting the interviews in the participants' homes made this lightweight solution necessary. Figure 2 shows the robotic arm bringing a glass of water to the user's face (in this case simulated by the camera lens). To further the realism of the situation, participants were able to experience the actual sounds of the robotic aid by simultaneously listening to an audio recording.

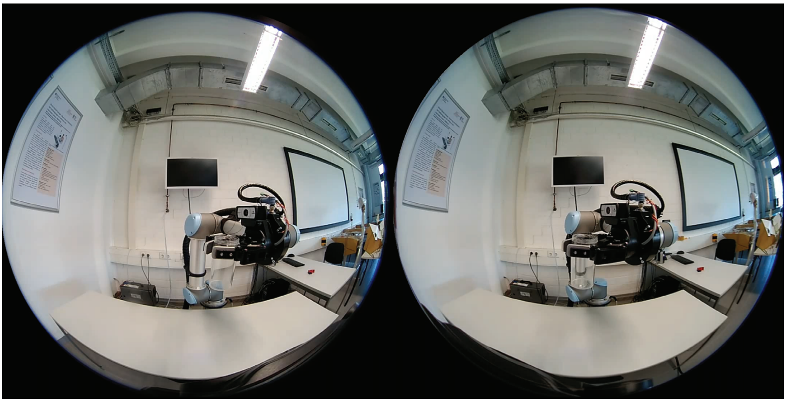


Fig. 2. Stereoscopic video in a first-person perspective of sitting in front of a robotic arm performing tasks in a close-contact environment

Following this experience we inquired about the participants' perception of the robotic aid including their likes and dislikes of the simulation and any changes they would appreciate from an end-user perspective. We encouraged them to express wishful thinking without worrying about current technological capabilities. We were also interested in how the participants would like to interact and collaborate with the robot. Special interest concerned the preferred location of attachment (e.g. table, wheelchair or self-mobile) and which additional functions should ideally be available.

Any additional thoughts, wishes, and suggestions of the participants were recorded for use in future research.

Participatory Observations. This part of the study focused on observing participants consume food and drinks (cf., Fig. 3) with the assistance of their caregiver. Observations of the relative location of the assistant, the methods used, and the communication between both parties were recorded. Depending on the specific type of impairment participants were either laying in their beds or sitting in their wheelchairs. Filming these interactions allowed for easy access during data analysis.

(a) Eating (front view)

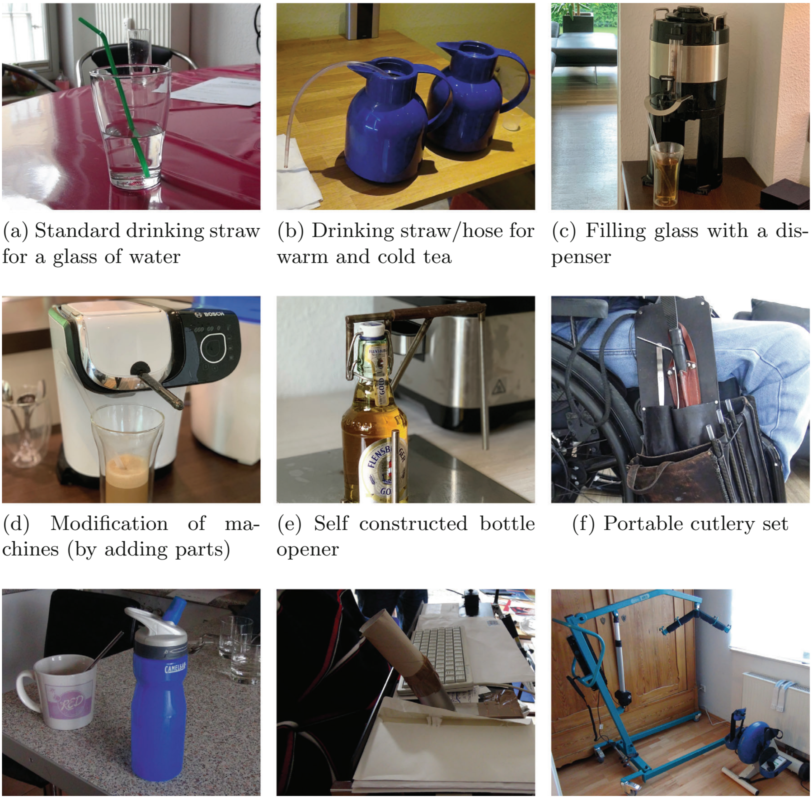
(b) Eating (back view)

(c) Drinking with a straw

(d) Drinking with glass contact

Fig. 3. Observation of eating drinking habits together with their caregiver

Analysis of Commonly Used Aids. During the qualitative interviews we found that every study participant uses some kind of aid to facility food and drink consumption (cf., Fig. 4). In order to consume the necessary amount of fluids, tea and water are provided in teapots (cf., Fig. 4b) or in large dispensers (cf., Fig. 4c). Large dispensers contain enough beverages for all day without the need for re-filling by an assistant. We recorded images of these aids to increase our understanding of the help people need and want when confronted with tasks they cannot independently do anymore.



(a) Standard drinking straw for a glass of water

(b) Drinking straw/hose for warm and cold tea

(c) Filling glass with a dispenser

(d) Modification of machines (by adding parts)

(e) Self constructed bottle opener

(f) Portable cutlery set

(g) Hydro flask

(h) Mouth-stick rest

(i) Lifter

Fig. 4. Analyzed aids in the participants' homes which are currently used

3.3 Limitation of Our Study

The main target group of our study were people with quadriplegia caused by spinal cord injury. The noticeable skewed ratio of more male than female participants reflects statistic by the WHO [51] and German Federal Statistical Office (Destatis) [47] of a 2:1 male-to-female ratio for overall recorded spinal cord injuries worldwide [33]. Additionally, a higher number of women refused to take part in our study as they felt uncomfortable with the study design (e.g. getting filmed while eating and drinking).

Use of a Robotic Arm During the Study. One participant (P15) has rudimentary mobility functions in her lower right arm which allows her to use a joystick-controlled robotic arm for nearly every activity of daily living including consuming food and drinks, manipulating objects and basic hygiene. Figure 5 illustrates how she handles the tasks with her robotic arm.

(a) Drinking (b) Eating

Fig. 5. Observation on the use of a robotic arm (JACO) for the consumption of food and drinks

4 Results and Recommendations

Based on the analysis of interviews, observations, and images we devised a set of user-centered design recommendations for a robotic drinking aid. Recommendations are split into three sections referring to structural, social and collaborative concerns respectively. All materials were transcribed, coded, and categorized independently by two researchers. We focus on processes related to drinking and eating, interactions between participants and caregivers, and additional topics of interest. All participants were interviewed as experts in their own right, as they can accurately describe and explain their situation, their abilities, limitations and needs. Therefore, our analysis method for the interviews and questionnaires followed the qualitative content analysis approach from Mayring [32]. Based on a predetermined interview guide established by the research team, questions belonging into different categories were discussed with study participants. Descriptive and normative statements concerning housing and living situation, individual wishes and needs regarding food and drink intake and attitude towards robotic aids were analysed [7]. The analysis of videos and images was based on the qualitative hermeneutical approach from social sciences by Reichertz and Englert [39] and the photo analysis by Pilarczyk and Mietzner [37]. In a first step, the videos and photos were cataloged according to content (e.g. drinking aid) and subject (e.g. drinking with a straw). Next, the videos were viewed, transcribed and coded. In a last step the photos were viewed again and theme-oriented photo series formed, e.g. photos showing self-made aids. Overall we conclude that photos provide additional information to the videos and interviews or can be used to better describe findings but do not provide much value as standalone objects.

4.1 Structural Concerns

Any design process starts with a structural framework defining size, weight and materials to be used. Whilst the choices might make sense from a purely technical point of view the preferences of the end-user should still be considered. Nobody benefits from the development of an assistive technologies that ends up too big in size to be used in the home of the typical end-user. With the aim being the widespread usage of the new device, taking wishes, where technically possible, regarding size and design into considerations can only be beneficial to future acceptance.

Dimensions of the Robotic Arm. Although all but one of the participants reported living in accessible housing, barriers including narrow hallways remain. During the in-home session we found several of the participants housings to be either too small for current robotic aids or lacking in space due to other large assistive devices present. Care beds and tables, lifters and wheelchairs are essential to support people with disabilities living in their own homes. Adding another large-size device taking up space can be problematic and in some cases impossible.

P10: “There is a second wheelchair somewhere, then maybe there is a bedside table somewhere, and there is a lifter somewhere and the shower chair somewhere. (...) At some point, many run out of space.”

> **Recommendation 1**
>
> A robotic drinking aid should be primarily designed for saving space. The arm has to have the ability to fold itself during waiting/suspended-mode. And include the possibility of space-saving storage when not in use.

Physical Attachment. Types of suitable attachment methods vary depending on individual preferences, type of wheelchair used, and space availability in the participants’ homes. Frequent changes between user location (bedroom/living room and bed/wheelchair) is a further factor to consider. Some participants use a chin-controlled electronic wheelchair. This poses an additional challenge for possible attachment methods and hinders both control and movement of the robot during the drinking and eating process because the joystick is in this case directly in front of the mouth.

P13: "It would be great to be able to fix the arm to the table with a small screw clamp. (...) Adding it to the wheelchair might be good too. But in any case it must be easy to dismantle."

> **Recommendation 2**
>
> Different mounting options for the robotic arm have to be available to allow attachment to different surfaces and care devices including mobility aides, resting chairs or overbed/side tables. Special consideration has to be given to the restrictions imposed by types of wheelchairs used.

4.2 Social Concerns

The advantage of an interdisciplinary approach as outlined in this study is the combination between technical necessities and preferences of the end-user. During our analysis we found that the majority of respondents were much more worried about'social concerns' than technical aspects.

Taking Design Seriously. Stigmatization of people with disabilities is an ongoing problem highlighted in a number of studies and literature e.g. [31]. All participants have reported that they worry about unwanted attention and further stigmatisation by using too-conspicuous aids. Almost all asked for the robotic arm to be unobtrusive and designed with a positive public image in mind.

P14: "I can imagine that design is relatively important, because it is likely the crucial factor whether people accept it and whether they want to integrate it into their environment, right?"

> **Recommendation 3**
>
> The design should range between something plain and unobtrusive to a chic lifestyle product. The arm should be recognizable as a technical tool and not mimic a human arm by using skin-colored coloration or skin-like material.

The Care Situation and Social Aspects. All participants relied on care from in-home relatives for their daily needs. Additionally all but one also employed professional personal care assistants. Due to the limited possibilities of outside interactions the bond with family members and other regular caregivers was observed as particularly strong and important for the mental well-being of the participants.

Interviewer: "Would you describe the exchange between you and your assistant during mealtime as formal or informal?" P4: "Very informal, just about everyday life. Not just about mine. They tell me about themselves. And then you just sit together and talk about everything that is going on. Daily events, personal matters, politics, experiences, about everything really."

> Recommendation 4
>
> Disruptions of conversations and social interactions by the drinking aid have to be minimized. The robotic arm should be placed without obstructing the line of sight between user and assistant. Sound and noises have to be kept to an absolute minimum to avoid distractions.

Safety. Safe use must be guaranteed for primary and secondary users from the onset. Teething problems must be avoided at all costs; therefore strict adherence to safety protocols for direct physical proximity is vital. People within the target group are already faced with numerous health concerns [6,46] and many participants expressed worries about additional injury risks posed by the robot. Worries surrounded their inability/decreased ability to move out of the way if the robotic arm does not stop at a certain distance from their face. A frequently suggested solution would be an adjustment to have the robot bringing a cup with a straw close to the mouth - but not directly touching it; thus enabling the user to cover the last few centimeters on their own accord.

Solutions include the aid of a straw to avoid the drinking cup being delivered directly to the mouth. Allowing users the final approach increases their feeling of autonomy, control and safety.

Interviewer: "What could prevent you from using the robotic arm?"
P2: "Teething troubles, something every device has at the beginning. If problems with the programming come up and the whole weight of the robot would fall on me."

> Recommendation 5
>
> Apply the principle of *safety first* and design for scenarios of use avoiding body-contact.

Privacy. All participants require 24/7 assistance with results in very limited privacy. They all stated that they have to drink a lot during the day for health reasons. Particularly in the case of paraplegia, it is necessary to consume up to three liters of fluid a day to support digestion and temperature regulation [21]. Being able to regulate fluid intake independently and not having to ask for assistance every time they want to drink would allow users to spend several hours at a time without a caregiver. A frequently recorded hope concerns the increase of time being alone gained by integrating the drinking aid into the users' lives.

P4: "And you really sometimes want to be all alone. And even if I send my assistants to go shopping and have an hour alone here or there, that's not comparable to really being alone."

> **Recommendation 6**
>
> Users have to be able to use the device, once set up, independently or with minimal assistance. Once operational, assistants and caregivers should not need to interact with the device at all. Potential components worn by the user need to fit securely to prevent a constant need of re-adjustments (cf., guidelines for wearability [19]).

Data Privacy and Security. Only a small number of participants were concerned about data protection. Some however expressed concerns about the type of data collected, storage options, as well as access to it.

P6: "If there is a camera then I do not know where these images are going. Especially if the robot is connected to the internet."

> **Recommendation 7**
>
> Transparency about collected and stored data have to be maintained to reduce uncertainty and skepticism about modern technology. Storage of personal data, including camera images, should be avoided and frameworks for voice commands should work offline as much as possible. If data have to be stored, it has to be stored securely.

4.3 Collaborative Concerns

Effective assistive technologies only work if they can be used by the target group without major effort. End-users know best what they are capable of and how they feel most comfortable interacting with the robotic device. Therefore it is important to consider the way they want to collaborate with their robotic aid.

Ease of Use. The aim of using assistive technologies is increased independence of the end-user; something that is only possible if the devices are easy and straight-forward to use. Especially in the case of changing caregivers, it is exhausting for users to repeatedly train others in the use of their robotic aid.

P1: "That means that I might not need a nurse anymore, but a technician. Because I already struggle to instruct the carers; and that is just to trigger three commands on my computer."

Recommendation 8

Ease of use, preferably as *ready-to-use* design, should be the aim of the assistive technology. Given the potential of frequently changing assistants, intuitive design and an obvious command structure are required to ensure a short - if at all necessary - familiarization periods. No prior knowledge or training by secondary users can or should be expected and if anything a short introduction guided by the primary user has to be sufficient. The robotic arm should be - once adjusted to a mounting spot - ready to use and easily used.

Interaction Design and Interaction Technology. Participants in the study indicated various desires regarding the interaction with the robotic aid. Due to the frequent changes in position during the day, it is important that the robotic arm is usable in a lying and in a sitting position either from a bed and from a wheelchair.

The majority of the participants already use voice controlled components in their homes, e.g. telephone, door opener, and lighting fixture. However, these components generally cannot be compared with modern smart devices as they do not connect to the internet. Only one participant used an smart speaker for smart home solutions. Other participants refrained from using devices with internet based voice control due to unreliable internet connections or - more often - out of concerns towards data security.

P14: "So I think it would be great if it was using voice control. (...) I think using a joystick or something similar is also very complex. But if I only have to say: "Give me a glass of water", and that would work, that would be great."

Recommendation 9

Whilst voice control is preferred for control and interaction, speech impairments must be taken into account with the extra requirements they pose. Additionally, data security has been identified as a concern when usage of internet based voice control is suggested. Offline solutions are preferable to address these worries. Alternatively, eye-tracking devices and data glasses can be viable options. Participants preferred the former two options compared to head gestures and headsets. Participants preferred eye-tracking control via gaze-dwelling on either real world components or virtual objects in combination with the data glass's user interface.

For users with residual hand and arm functions, a switch among semi-autonomous mode and manual mode via direct joystick control is interesting as it allows greater flexibility and adaptation to daily needs, due to the fact that in a semi-autonomous mode scenarios have to been learnt by the robot. The current mode has to communicated to the users and assistants, e.g. by a ring of two-colored LEDs around the robot's flange like a bracelet.

Robotic Arm as Combined Drinking and Eating Aid. All Participants are excited about the prospect of a functional drinking aid, allowing them to independently regulate their fluid intake. In contrast, few participants can imagine regularly using a robotic arm for food consumption. Those who can still eat independently due to residual functionality in their upper extremities would like to use and maintain this ability. A robotic arm as an eating aid is only interesting for this group if food can be cut into small pieces with the help of the arm. A cutting function would further increase their autonomy and enable participation in meal preparation; in their eyes another step towards social integration. Participants who have their food served to them expressed satisfaction with the assistance they receive from other people. They would like to continue in this way because they value this social interaction and note that people can be more flexible and spontaneous in responding to all eventualities. This includes emergency situations such as choking or spillages, a worry of a number of participants from our focus group voiced.

P5: "I would prefer [the aid] of my husband, because we do communicate a little throughout lunch. I think when the robotic arm feeds you, there is just silence."

Recommendation 10

When prioritising the development of robotic assistance the first focus should be on fluid intake. The scenario of eating with a robotic arm is influenced by various complex aspects, such as social interactions, which need further exploration.

Robotic Arm as a General Aid. Participants frequently expressed a desire for a robotic arm with a distinct grasping function beyond a mere drinking aid. Desired functions include manipulating objects, such as picking something up, taking something out of a cabinet, or being able to lift things. Fine motor tasks such as turning the pages of a book or grasping easily breakable items were also desired. In addition, particularly the younger participants would like to be able to operate a game console. Furthermore, some of the participants would also like to use a robotic arm for aspects of basic care, such as combing hair or brushing teeth.

Participants who still eat independently also showed interest in the topic of cooking. A robotic arm that can cut food, handle cooking utensils, and assist with setting the table would increase autonomy and lead to more participation in the entire process of eating. Participants expressed the wish to handle even fragile objects like raw eggs or eat small but delicate snacks like crisps. Although fears of possible stigmatization due to the use of the robotic arm exist, overall the hope that a robotic arm with various functions could promote independent and self-determined living whilst also giving relieve to caregivers was expressed.

P4: "Having such an arm fulfill different functions such as gripping, I think that makes more sense because it would then be more versatile."

Recommendation 11

Apart from functioning as a drinking and eating aid, a robotic arm should be developed to fulfill other everyday tasks. Since participants fear stigmatization over having too many tools, a robotic aid with various functions would meet greater acceptance.

5 Discussion

Our ethnographic study provides recommendations for future research and development as well as hypotheses that should be tested for further validation. In the context of our target user group, implementing a solution based on our recommendations will still require adaptation to fit individuals with their specific physical abilities, along with further research to verify that a designed assistive system does indeed support the user. Recommendations regarding "Taking Design Seriously" and "Privacy" concern interaction devices that might lead to further stigmatization by drawing unwanted attention and require asking for assistance for wearing or re-calibration - both aspects the target group wants to avoid.

5.1 End-User Involvement in Assistive Technology

Assistive technologies are on the rise, with a number of different robotic aids already on the market or in various stages of development [15,25]. Studies by Scherer[43] and Verza [49] have shown that these devices, albeit useful in an assistive setting, can have a high rate of non-acceptance and non-usage. There is a growing body of literature indicating that this is due to the exclusion of the end-user from the design process [16]. In recent years the field of collaborative work between developers and end-users (or their advocates) has grown but is still in its infancy as discussed by Lee et al. [27] and Simonsen [44]. Our work represents such a collaborative approach investigating the needs and wants of the end-user in regards to a robotic drinking and eating aid. In fact, participants particularly valued the inclusion in the design process of a device developed specifically for them.

5.2 Potential Autonomy

One important finding is that participants want the possibility of spending time without their assistants. Specifically drinking as continuous hydration throughout the day is vital for people with disabilities which results in a near-constant need of care when drinking independently is not possible [21]. Thus, the most important capability of the robotic system should be to support the users with drinking. On the other hand, however, the participants also noted that the number of assistive systems should be limited and, thus, the system should provide multiple tasks. This trade off will be a core challenge for future developers.

5.3 Importance of Structural Concerns

Our results confirm findings by Fattal et al. [14] and others [34,48] as similar structural concerns are also expressed by our participants. Thus, our results also highlight the need for recommendations related to the physical characteristics of the robotic arm and attachment site.

5.4 In-Home Methodology

In this study we recorded the preferences people with severe motor disabilities have towards a robotic drinking and eating aid in terms of functionality and design. We used in-home sessions to interview our participants and record their everyday behaviours in a familiar setting. We opted for this particular approach to increase authenticity of our observation in accordance with Sakowska [42]. Combining all findings from our conversations and observations allowed us to gain significant insight in the actual living situation and challenges faced by the target group. We believe the in-home methods used in this study to represent a much more accurate picture than studies conducted in artificial laboratory or workshop environments. One downside of this approach, however, is the limitation to a small geographical area, potentially limiting the generalizability.

6 Conclusion

People with motor disabilities face a number of obstacles when confronted with everyday tasks such as drinking and eating. Assistive technologies have the potential to greatly improve the quality of life of the target group; however their user acceptance has been challenged by previous work. In this paper we investigated how drinking and eating aids are perceived by conducting interviews and participatory observations. By analyzing the relationship with food and drink intake as well as analyzing the wishes for future assistive technologies we were able to better understand the needs and wants of the target group.

Our research has highlighted the importance of acknowledging structural, social, and collaborative concerns in respect to the design of a robotic arm, defining a set of recommendations for the designs of robotic drinking aids. These recommendations represent an important step in bridging the gap between technological design and the preferences of the target group, thus increasing the likelihood of acceptance of any further assistive technology.

Acknowledgement. We would like to thank all study participants and their assistants for their valuable opinions and time. The authors are also grateful for the support of the *Center for Paraplegic Patients Hamburg*, the *Locked-in-Syndrom e.V. Berlin*, and the *State Association of the German Society for Multiple Sclerosis Hessen e.V.*. This research is supported by the German Federal Ministry of Education and Research (BMBF, FKZ: 16SV7866K and 16SV7868).

References

1. Achanccaray, D., Chau, J.M., Pirca, J., Sepulveda, F., Hayashibe, M.: Assistive robot arm controlled by a p300-based brain machine interface for daily activities. In: 2019 9th International IEEE/EMBS Conference on Neural Engineering (NER), pp. 1171–1174 (2019). https://doi.org/10.1109/NER.2019.8717042
2. Assistive Innovations B.V.: iARM—Robotic arm for humans, mountable on powered wheelchair. https://www.assistive-innovations.com/robotic-arms/iarm. Accessed 2 May 2021
3. Assistive Innovations B.V.: iEAT Robot—Assistive feeding and eating robot for people. https://www.assistive-innovations.com/eatingdevices/ieat-robot. Accessed 2 May 2021
4. Beaudoin, M., Lettre, J., Routhier, F., Archambault, P.S., Lemay, M., Gélinas, I.: Long-term use of the JACO robotic arm: a case series. Disabil. Rehabil. Assist. Technol. **14**(3), 267–275 (2019). https://doi.org/10.1080/17483107.2018.1428692
5. Bemelmans, R., Gelderblom, G.J., Jonker, P., de Witte, L.: Socially assistive robots in elderly care: a systematic review into effects and effectiveness. J. Am. Med. Dir. Assoc. **13**(2), 114–120 (2012). https://doi.org/10.1016/j.jamda.2010.10.002
6. Bickenbach, J.: International Perspectives on Spinal Cord Injury. World Health Organization, Geneva (2013)
7. Bogner, A., Littig, B., Menz, W.: Interviews mit Experten. QS, Springer, Wiesbaden (2014). https://doi.org/10.1007/978-3-531-19416-5
8. Canal, G., Alenyà, G., Torras, C.: Personalization framework for adaptive robotic feeding assistance. In: Agah, A., Cabibihan, J.-J., Howard, A.M., Salichs, M.A., He, H. (eds.) ICSR 2016. LNCS (LNAI), vol. 9979, pp. 22–31. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-47437-3_3
9. Chen, T.L., et al.: Robots for humanity: using assistive robotics to empower people with disabilities. IEEE Rob. Autom. Mag. **20**(1), 30–39 (2013). https://doi.org/10.1109/MRA.2012.2229950
10. DESiN LLC: Obi—The first dining robot of its kind. https://meetobi.com/. Accessed 2 May 2021
11. Duchowski, A.T.: Gaze-based interaction: a 30 year retrospective. Comput. Graph. **73**, 59–69 (2018). https://doi.org/10.1016/j.cag.2018.04.002
12. Duckworth, D., Henkel, Z., Wuisan, S., Cogley, B., Collins, C., Bethel, C.: Therabot: the initial design of a robotic therapy support system. In: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts, pp. 13–14 (2015). https://doi.org/10.1145/2701973.2701993
13. Eftring, H., Frennert, S.: Designing a social and assistive robot for seniors. Zeitschrift für Gerontologie und Geriatrie **49**(4), 274–281 (2016). https://doi.org/10.1007/s00391-016-1064-7

14. Fattal, C., Leynaert, V., Laffont, I., Baillet, A., Enjalbert, M., Leroux, C.: SAM, an assistive robotic device dedicated to helping persons with quadriplegia: usability study. Int. J. Soc. Rob. **11**(1), 89–103 (2018). https://doi.org/10.1007/s12369-018-0482-7
15. Federici, S.: Assistive Technology Assessment Handbook. CRC Press, Taylor & Francis Group, Boca Raton (2018)
16. Federici, S., et al.: Successful assistive technology service delivery outcomes from applying a person-centered systematic assessment process: a case study. Life Span Disabil. **18**(1), 41–74 (2015). http://www.lifespanjournal.it/client/abstract/ENG2902.%20Federici.pdf
17. Frennert, S., Östlund, B.: Review: seven matters of concern of social robots and older people. Int. J. Soc. Rob. **6**(2), 299–310 (2014). https://doi.org/10.1007/s12369-013-0225-8
18. Gallenberger, D., Bhattacharjee, T., Kim, Y., Srinivasa, S.S.: Transfer depends on acquisition: analyzing manipulation strategies for robotic feeding. In: 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 267–276 (2019). https://doi.org/10.1109/HRI.2019.8673309
19. Gemperle, F., Kasabach, C., Stivoric, J., Bauer, M., Martin, R.: Design for wearability. In: Digest of Papers. Second International Symposium on Wearable Computers (Cat. No. 98EX215), pp. 116–122 (1998). https://doi.org/10.1109/ISWC.1998.729537
20. Google VR: Google Cardboard. https://arvr.google.com/cardboard/. Accessed 2 May 2021
21. Haas, U.: Caring for People with Paraplegia: Problems, Needs, Resources and Interventions (nach "Pflege von Menschen mit Querschnittlähmung Probleme, Bedürfnisse, Ressourcen und Interventionen"). Huber, Bern (2012)
22. Jackowski, A., Gebhard, M., Thietje, R.: Head motion and head gesture-based robot control: a usability study. IEEE Trans. Neural Syst. Rehabil. Eng. **26**(1), 161–170 (2018). https://doi.org/10.1109/TNSRE.2017.2765362
23. Kinova Inc.: Kinova Jaco Assistive Robotic Arm. https://www.kinovarobotics.com/en/assistive-technologies/column-a1/kinovaassistive-robotic-arm. Accessed 2 May 2021
24. Klein, B.: Aides, assistive Technologies, and Robotics: Maintaining Independence and Quality of Life in old age (nach "Hilfsmittel, Assistive Technologien und Robotik: Selbstständigkeit und Lebensqualität im Alter erhalten"). Kohlhammer Verlag (2020)
25. Kyrarini, M., et al.: A survey of robots in healthcare. Technologies **9**(1), 8 (2021). https://doi.org/10.3390/technologies9010008
26. Laitano, M.I.: Developing a participatory approach to accessible design. Int. J. Sociotechnology Knowl. Dev. **9**(4), 1–11 (2017). https://doi.org/10.4018/IJSKD.2017100101
27. Lee, H.R., et al.: Steps toward participatory design of social robots. In: HRI 2017, pp. 244–253. IEEE, Piscataway (2017). https://doi.org/10.1145/2909824.3020237
28. Mahmud, S., Lin, X., Kim, J.: Interface for human machine interaction for assistant devices: a review. In: 2020 10th Annual Computing and Communication Workshop and Conference (CCWC), pp. 768–773 (2020). https://doi.org/10.1109/CCWC47524.2020.9031244
29. Mandy, A., Sims, T., Stew, G., Onions, D.: Manual feeding device experiences of people with a neurodisability. Am. J. Occup. Ther. **72**(3), 7203345010p1–7203345010p5 (2018). https://doi.org/10.5014/ajot.2018.025353

30. Martinsen, B., Harder, I., Biering-Sorensen, F.: The meaning of assisted feeding for people living with spinal cord injury: a phenomenological study. J. Adv. Nurs. **62**(5), 533–540 (2008). https://doi.org/10.1111/j.1365-2648.2008.04637.x
31. Matera, C., et al.: Put yourself in my wheelchair: perspective-taking can reduce prejudice toward people with disabilities and other stigmatized groups. J. Appl. Soc. Psychol. **51**(3), 273–285 (2021). https://doi.org/10.1111/jasp.12734
32. Mayring, P.: Qualitative Content Analysis: Basics and Techniques (nach "Qualitative Inhaltsanalyse: Grundlagen und Techniken"). Beltz, Weinheim (2015)
33. McColl, M.A., Charlifue, S., Glass, C., Lawson, N., Savic, G.: Aging, gender, and spinal cord injury. Arch. Phys. Med. Rehabil. **85**(3), 363–367 (2004). https://doi.org/10.1016/j.apmr.2003.06.022
34. Merkel, S., Kucharski, A.: Participatory design in gerontechnology: a systematic literature review. Gerontologist **59**(1), e16–e25 (2018). https://doi.org/10.1093/geront/gny034
35. Park, D., et al.: Active robot-assisted feeding with a general-purpose mobile manipulator: design, evaluation, and lessons learned. Rob. Auton. Syst. **124**, 103344 (2020). https://doi.org/10.1016/j.robot.2019.103344
36. Pascher, M., Schneegass, S., Gerken, J.: SwipeBuddy. In: Lamas, D., Loizides, F., Nacke, L., Petrie, H., Winckler, M., Zaphiris, P. (eds.) INTERACT 2019, Part IV. LNCS, vol. 11749, pp. 568–571. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-29390-1_39
37. Pilarczyk, U., Mietzner, U.: Picture science methods in research on education and social science (nach "Bildwissenschaftliche Methoden in der erziehungs- und sozialwissenschaftlichen Forschung"). Zeitschrift für qualitative Bildungs-, Beratungs- und Sozialforschung **1**(2), 343–364 (2000). https://www.ssoar.info/ssoar/handle/document/28057
38. Plaumann, K., Ehlers, J., Geiselhart, F., Yuras, G., Huckauf, A., Rukzio, E.: Better than you think: head gestures for mid air input. In: Abascal, J., Barbosa, S., Fetter, M., Gross, T., Palanque, P., Winckler, M. (eds.) INTERACT 2015, Part III. LNCS, vol. 9298, pp. 526–533. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-22698-9_36
39. Reichertz, J.: Introduction to Qualitative Video Analysis – A hermeneutic-sociological Analysis Tool (nach "Einführung in die qualitative Videoanalye – Eine hermeneutisch-wissenssoziologische Fallanalyse"). VS Verlag für Sozialwissenschaften, Wiesbaden (2011)
40. Roig-Maimó, M.F., MacKenzie, I.S., Manresa-Yee, C., Varona, J.: Head-tracking interfaces on mobile devices: evaluation using fitts' law and a new multi-directional corner task for small displays. Int. J. Hum. Comput. Stud. **112**, 1–15 (2018). https://doi.org/10.1016/j.ijhcs.2017.12.003
41. Rudigkeit, N., Gebhard, M.: Amicus—a head motion-based interface for control of an assistive robot. Sensors **19**(12), 2836 (2019). https://doi.org/10.3390/s19122836
42. Sałkowska, M.: Carrying out research among persons with disabilities and their relatives – selected ethical issues. Zoon Politikon **8**, 200–217 (2018). https://doi.org/10.4467/2543408XZOP.18.010.10066
43. Scherer, M.J.: Living in the State of Stuck: How Assistive Technology Impacts the Lives of People with Disabilities. Brookline Books, Brookline (2005)
44. Simonsen, J.: Routledge International Handbook of Participatory Design. Routledge, New York (2013)
45. Stalljann, S., Wöhle, L., Schäfer, J., Gebhard, M.: Performance analysis of a head and eye motion-based control interface for assistive robots. Sensors **20**(24), 7162 (2020). https://doi.org/10.3390/s20247162

46. Statistisches Bundesamt (Destatis): Statistical Yearbook 2019 – Chapter 4 Health (nach "Statistisches Jahrbuch 2019 – Kapitel 4 Gesundheit"). https://www.destatis.de/DE/Themen/Querschnitt/Jahrbuch/jb-gesundheit.pdf (2019)
47. Statistisches Bundesamt (Destatis): Disability Facts and Figures - Brief Report 2019 (nach "Statistik der schwerbehinderten Menschen - Kurzbericht 2019") (2020). https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Behinderte-Menschen/Publikationen/Downloads-Behinderte-Menschen/sozial-schwerbehinderte-kb-5227101199004.html
48. Thielke, S., Harniss, M., Thompson, H., Patel, S., Demiris, G., Johnson, K.: Maslow's hierarchy of human needs and the adoption of health-related technologies for older adults. Ageing Int. **37**(4), 470–488 (2012). https://doi.org/10.1007/s12126-011-9121-4
49. Verza, R., Carvalho, M.L.L., Battaglia, M.A., Uccelli, M.M.: An interdisciplinary approach to evaluating the need for assistive technology reduces equipment abandonment. Multiple Scler. J. **12**(1), 88–93 (2006). https://doi.org/10.1191/1352458506ms1233oa
50. Vines, J., Clarke, R., Wright, P., McCarthy, J., Olivier, P.: Configuring participation: on how we involve people in design. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 429–438 (2013). https://doi.org/10.1145/2470654.2470716
51. World Health Organization: Spinal cord injury. https://www.who.int/news-room/fact-sheets/detail/spinal-cord-injury. Accessed 2 May 2021

HaptiX: Vibrotactile Haptic Feedback for Communication of 3D Directional Cues

Max Pascher
max.pascher@w-hs.de
Westphalian University of Applied Sciences
Gelsenkirchen, Germany
University of Duisburg-Essen
Essen, Germany

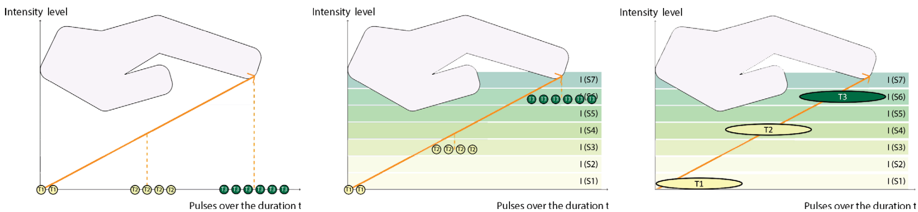
Til Franzen
til.franzen@studmail.w-hs.de
Westphalian University of Applied Sciences
Gelsenkirchen, Germany

Kirill Kronhardt
kirill.kronhardt@.w-hs.de
Westphalian University of Applied Sciences
Gelsenkirchen, Germany

Uwe Gruenefeld
uwe.gruenefeld@uni-due.de
University of Duisburg-Essen
Essen, Germany

Stefan Schneegass
stefan.schneegass@uni-due.de
University of Duisburg-Essen
Essen, Germany

Jens Gerken
jens.gerken@w-hs.de
Westphalian University of Applied Sciences
Gelsenkirchen, Germany



(a) Rabbit Single (b) Rabbit Dual (c) Motion Intensity

Figure 1: Coding of the gradient for a) *Rabbit Single*, (b) *Rabbit Dual*, and (c) *Motion Intensity*. The orange arrow represents the gradient of the directional cue with an intended increase over time. The timing, duration, number of pulses, and intensity (S1 – S7) of the three actuators (T1 – T3) are illustrated for each condition.

ABSTRACT

In Human-Computer-Interaction, vibrotactile haptic feedback offers the advantage of being independent of any visual perception of the environment. Most importantly, the user's field of view is not obscured by user interface elements, and the visual sense is not unnecessarily strained. This is especially advantageous when the visual channel is already busy, or the visual sense is limited. We developed three design variants based on different vibrotactile illusions to communicate 3D directional cues. In particular, we explored two variants based on the vibrotactile illusion of the cutaneous rabbit and one based on apparent vibrotactile motion. To communicate gradient information, we combined these with pulse-based and intensity-based mapping. A subsequent study showed that the pulse-based variants based on the vibrotactile illusion of the cutaneous rabbit are suitable for communicating both directional and gradient characteristics. The results further show that a representation of 3D directions via vibrations can be effective and beneficial.

CCS CONCEPTS

• **Human-centered computing** → *User centered design*; **Haptic devices**; • **Hardware** → **Haptic devices**; • **Computing methodologies** → Virtual reality.

KEYWORDS

directional cues, haptic feedback, vibrotactile feedback

ACM ISBN 978-1-4503-9422-2/23/04.
https://doi.org/10.1145/3544549.3585601

ACM Reference Format:
Max Pascher, Til Franzen, Kirill Kronhardt, Uwe Gruenefeld, Stefan Schneegass, and Jens Gerken. 2023. HaptiX: Vibrotactile Haptic Feedback for Communication of 3D Directional Cues. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 7 pages.
https://doi.org/10.1145/3544549.3585601

1 INTRODUCTION

People perceive objects in their environment primarily through their sense of sight. However, this ability can be reduced or not possible at all in certain situations. Objects might be covered by other things / user interface (UI) elements (visual clutter) or be outside the human field of view. In addition, visual perception may be limited or impossible due to visual impairments. Previous research has shown that the haptic modality can, to some extent, compensate for the lack of visual information and outperform audio-based cues [3]. It can also be applied in combination with other modalities and can offer an additional information channel if, for example, the visual channel is overloaded due to distracting information [5, 14].

Directing attention, guiding, and transmitting patterns via vibrotactile signals have already been researched and found to be useful feedback modalities [11, 12, 26]. Barralon et al. studied pattern recognition using a vibrotactile belt with eight actuators and tasked participants to select the corresponding correct visual representation [2]. Lee and Starner proposed *BuzzWear*, a wearable tactile display with three vibration actuators for notification purposes that function by modulating intensity, pattern, direction, and starting point [16]. After 40 minutes of training, subjects could distinguish between the 24 patterns with up to 99% accuracy. Vibrotactile feedback is also used in the context of guidance. Here, a study by Lehtinen et al., used a vibrotactile glove to support a visual search task on a flat plane on a wall [17].

However, a common challenge is that tactile displays have a limited resolution. Therefore, researchers have simulated smooth movement patterns with the help of tactile illusions [6], such as *Phantom Sensations* [1, 19], *Apparent Tactile Motion* [4, 15, 23], and *Cutaneous Rabbit* [7, 18, 21, 24]. Tan et al. conducted a study using a 3 x 3 tactile display and applied the *Cutaneous Rabbit* sensation to explore the communication of eight 2D directional cues (north, northeast, east, southeast, south, southwest, west, and northwest) and the successful recognition of these cues [25].

While previous work focused on 2D directional cues (e.g., [25]) or allowed users to feel directions upon approach with their hand (e.g., [11]), we are not aware of any work that aims to communicate 3D directional cues. In particular, our work differs from approaches such as [12], who aim to push or pull the hand toward a known target in 3D space but who therefore do not actually need to encode 3D information for the vibration pattern itself. It also differs from work such as [27] which used a Tactile Vision Substitution System (TVSS) to communicate 3D shapes of a static object by directly mapping image features such as contours on a 20 x 20 tactile display.

Our approach builds on the idea of Tan et al. [25] to communicate 2D directions. We combine their base with pulse or intensity mapping to simultaneously communicate the gradient. Furthermore, we explore the influences of different haptic illusions (i.e., *Cutaneous Rabbit* and *Apparent Tactile Motion*) on the comprehension of directional cues. Our work contributes three specific design proposals for communicating 3D directional cues as well as a study on the effectiveness and subjective experience of this non-visual approach to direction mapping.

2 CONCEPT

Within the scope of our experiment, three variants were developed to map vibrotactile 3D directional cues. For the 2D direction, the vibrotactile illusions of the *Cutaneous Rabbit* and *Apparent Tactile Motion* were used. We extended these by a pulse- and intensity-based approach to communicating the gradient of the 3D directional cue (see Figure 1).

2.1 *Rabbit Single*: Cutaneous Rabbit with Pulse-based Approach

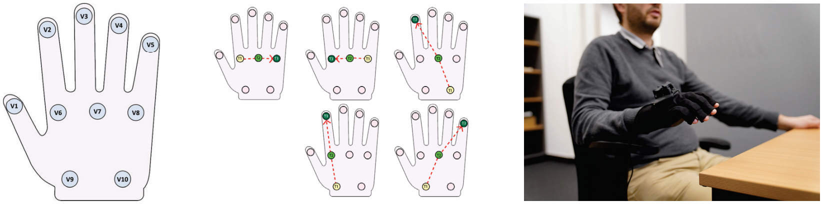
This condition is based on the *Cutaneous Rabbit* for communicating 2D direction, a tactical illusion that can influence the design of vibrotactile patterns. This illusion was discovered in 1972 by Geldard [8]. The sequence of taps on different vibrotactile actuators is perceived as a continuous movement between the different points. Each directional cue is abstracted using three control points for the actuators (illustrated as dashed lines in Figure 1). Depending on the distance resulting from the gradient of the directional cue, the number of pulses triggered at each actuator is determined in a range of 1 – 7 with a Burst Duration (BD) of 125ms, an Inter-Stimulus Interval (ISI) between pulses of 50ms, and an Inter-Burst Interval (IBI) between actuators of 100ms. The closer the control point of the direction cue is to the hand, the higher the number of vibration pulses (see Figure 1a).

2.2 *Rabbit Dual*: Cutaneous Rabbit with a Pulse- and Intensity-based Combined Approach

Rabbit Dual is based on *Rabbit Single* but includes a second additional encoding for the gradient of the 3D directional cue. In addition to the number of pulses, we mapped three different intensity levels on the distance of the directional cue to the palm (see Figure 1b). We based the distinct intensity levels on prior work by Gescheider et al., who measured a just noticeable relative difference threshold – Just-Noticeable Difference (JND) – with values of 0.26 at 4 dB above the perceptual threshold [9]. To communicate and distinguish between up- and downward gradients, three distinct intensity levels were selected - a baseline level in the middle and one low- as well as one high-intensity level. The anticipated benefit of this condition was that gradient comprehension would be improved due to the dual encoding.

2.3 *Motion Intensity*: Apparent Tactile Motion with Intensity-based Approach

This condition applied the same intensity mapping for the gradient as *Rabbit Dual*, but without the pulses. In contrast to the *Cutaneous Rabbit* sensation with distinct pulses as in *Rabbit Single* and *Rabbit Dual*, here we applied the vibrotactile illusion of *Apparent Tactile Motion*. This was first studied in the early 20th century by Burtt [4] and is commonly referred to as the *Phi Phenomenon*. The illusion is created by an overlap in the start times of two actuators – Inter-Stimulus Onset Asynchrony (SOA), calculated as $SOA = 0.32d + 47.3ms$, where d is the vibration period of an actuator – 450ms. Instead of two individual actuators, a single stimulus is perceived as moving from the position of the first triggered actuator T1 to the second actuator T2 – or from actuator T2 to T3 (see Figure 1c). A



(a) Actuator Placement as provided by *SensorialXR* **(b) 2D Directions** **(c) Study Setup**

Figure 2: For the 2D directional cues, we used (a) the placement of all actuators across the hand to (b) communicate five different directions. Here, (c) illustrates the study setup, with the arm resting on the armrest while the hand is in the air.

potential benefit of this illusion is that it may feel more like a natural movement, as it disguises the limited number of actuators. After pilot tests, a starting intensity value of 0.22 and a JND value of 0.3 were chosen, which made the intensity levels easily distinguishable. Thus, a total of seven possible intensity levels were defined.

2.4 Implementation

To develop our approach, we use the 3D game engine *Unreal Engine 4* optimized for usage with a *Meta Quest 2* Virtual Reality (VR) Head-Mounted Display (HMD). This allows for the use of a virtual environment in which the participants can concentrate purely on the haptic feedback without being visually distracted. It also provides a simple way to visually explain the directional cues to the participants and record their responses for rating scales. As a haptic display, we chose the *SensorialXR* glove as a commercially available device with Software Development Kit (SDK) interface to the *Unreal Engine 4*. With ten actuators – Linear Resonant Actuator (LRA) vibration motors, fixed in place – (see Figure 2a), *SensorialXR* gloves are among the models with the most vibration motors per hand. Thus, they offer the potential to map the 3D directional cues with the highest possible vibrotactile resolution [22].

3 STUDY

We conducted a within-subjects experiment with 14 participants to explore and understand the differences and similarities between the three presented designs for vibrotactile feedback (independent variable) regarding their effectiveness in communicating 3D directional cues. As participants were supposed to feel and comprehend directional cues without any additional visual feedback, we conducted the study in person and within a neutral VR environment, which allowed participants to focus entirely on the vibrotactile feedback. The age of participants ranged from 21 to 31 years, with a mean age of 25.71 years ($M = 25.71, SD = 2.972$). Four were female, ten were male, and all were university students of various subjects. None of the participants reported any visual impairment, and all were right-handed.

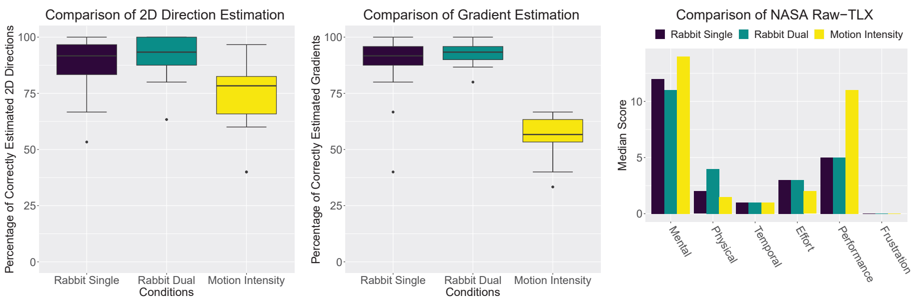
3.1 Procedure

The study was conducted in multiple comparable physical localities. Before commencing, participants were fully informed about the project objective and the various tasks they had to complete. Each participant gave their full and informed consent to partake in the study, have video and audio recordings taken, and have all the relevant data documented. Participants wore a HMD on their head and a vibrotactile glove on the right hand while being asked to keep their right arm rested on an armrest with the palm facing down (see Figure 2c) to avoid any external factors. In the left hand, participants held a controller to control the VR environment.

For each condition, each participant performed six training trials. For each trial, the vibrotactile feedback was repeated three times, and a corresponding visualization was shown to indicate the direction in 3D supporting the participant's mental model. For the actual task, participants were shown a neutral-colored background in VR without any visual representations of the 3D direction. Participants were able to trigger the start of the trial with the VR motion controller. In total, they completed 30 measured trials per condition, resulting in 90 measured trials per participant and 1,260 measured trials in total. The 30 trials consisted of 2 (blocks) x 5 (2D direction) x 3 (gradient). The variable *2D direction* represented a typical set of five possible mappings of straight horizontal, vertical and diagonal directions, which were physically located on the surface of the hand (see Figure 2b). They represented the direction in x-z-coordinates of the overall 3D directional vector. The *gradient* encoded the direction in y-coordinates: either up, down, or neither any gradient. To counter learning and fatigue effects, we applied a *Balanced Latin Square* design for the order of the three conditions. The order of trials was randomized within each block. Between each condition, participants were able to rest their hand for five minutes. The average session lasted for 45 minutes and concluded with a debriefing. Non of the participants mentioned any sensory or muscle fatigue. Participants received 15 EUR in compensation.

3.2 Variables and Research Questions

For dependent variables, we measured the accuracy of the comprehension of the *2D direction* (x-axis, z-axis) and the *gradient* (y-axis). We are measuring the two variables (*2D direction* and *gradient*)



(a) 2D Direction Estimation.

(b) Gradient Estimation.

(c) Task Load.

Figure 3: Measured performance for 2D direction and gradient estimation as well as task load measured with the NASA Raw-TLX (lower score is better). For the task load subscale "frustration" no bars are visible because all three conditions have a median score of 0.

separately, as commonly done within the research community (e.g., estimation of direction and distance for HMDs [10]). The main reasoning here is that orientation in 3D space and especially describing directions in 3D can be challenging for participants and could negatively affect the validity of the measurements. To do so, we presented participants with a UI panel in VR after each trial. The panel showed five pictures with all 2D directions in a top-down view and, subsequently, three pictures of all gradients in a lateral view. Participants used the VR controller to select the fitting representation for each. These two variables were measured with a binary outcome (correct, incorrect) and summarized as percentages of correctly identified outcomes across all trials per person and condition (ratio scale). In addition, we measured mental workload after each condition via the NASA Raw-Task Load Index (NASA RTLX) [13] and additional Likert-scale statements regarding the comprehensibility of the directional cue. We also collected qualitative feedback in a semi-structured interview after each condition as well as at the end of the study, which is when we also asked participants to rank the three conditions.

As the study is exploratory in nature, we were interested in finding out more about the specific features of our three feedback conditions. In particular, we were interested in the following research questions:

RQ1: Do multiple encodings of gradient, as in the condition *Rabbit Dual*, improve the comprehension of gradient information and reduce mental workload?

RQ2: Does apparent movement, as in the condition *Motion Intensity*, improve comprehension of 2D direction? We assume that to be the case as the transition by overlapping of vibration between the actuators may be easier to comprehend and interpret as a path compared to the sensation of isolated pulses as for the *Rabbit* conditions.

RQ3: How would participants experience and rate vibrotactile communication of 3D direction overall and with regard to each individual condition?

4 RESULTS

For our applied inferential statistics, we distinguished between ratio and ordinal data. The estimation percentages for 2D direction and gradient are ratio data, while the Likert items – including task load – are ordinal data. For ratio data only, we first applied a Shapiro-Wilk test to check for normality. We found that none of our ratio data is normally distributed. Thus, we treated all our data in the same way and directly applied non-parametric tests, specifically Friedman tests. Thereafter, we conducted Wilcoxon Signed-rank tests with Bonferroni correction for our post-hoc analysis. The effect sizes of the Wilcoxon tests are reported as r (r: >0.1 small, >0.3 medium, and >0.5 large effect).

4.1 Estimation of 2D Direction

We asked participants to estimate the two-dimensional direction on a ground plane. The median (interquartile range) percentages of correct 2D direction estimations for each condition are (in descending order): *Rabbit Dual* =93.3% (IQR=12.5%), *Rabbit Single* =91.7% (IQR=13.3%), and *Motion Intensity* =78.3% (IQR=16.7%). All percentages are compared in Figure 3a. Since our data is not normally distributed ($p<0.01$), we directly ran a Friedman test that revealed a significant effect of condition on 2D direction estimation ($\chi^2(2)=17.70$, $p<0.001$, N=14). Post-hoc tests showed significant differences between *Rabbit Single* and *Motion Intensity* (W=83, Z=2.62, p=0.018, r=0.50) as well as *Rabbit Dual* and *Motion Intensity* (W=0, Z=-3.30, p<0.001, r=0.62). However, we did not find a significant difference between *Rabbit Single* and *Rabbit Dual* (W=15, Z=-1.88, p=0.182). Here, **we can conclude that both *Rabbit Single* and *Rabbit Dual* result in better estimation performance for 2D direction than *Motion Intensity*.**

4.2 Estimation of Gradient

We asked participants to estimate the gradient behavior of the communicated cue. The median (interquartile range) percentages of

Table 1: Pairwise comparisons for individual statements, Bonferroni-adjusted, p-values: <0.05 (*), <0.01 (**), and <0.001 (***).

	Rabbit Single vs. Dual		Rabbit Single vs. Motion Intensity			Rabbit Dual vs. Motion Intensity		
Statement	test statistic	p-value	test statistic	p-value	effect size	test statistic	p-value	effect size
S1	Z= 0.00	p=1.000	Z=2.89	$p<.001^{***}$	r=0.55	Z=3.04	$p=.004^{**}$	r=0.57
S2	Z=-0.07	p=1.000	Z=2.58	$p=.029^{*}$	r=0.49	Z=2.80	$p=.013^{*}$	r=0.53
S3	Z=-1.17	p=0.838	Z=3.06	$p=.003^{**}$	r=0.58	Z=2.97	$p=.003^{**}$	r=0.56
S4	Z=-0.17	p=1.000	Z=2.84	$p=.006^{**}$	r=0.54	Z=2.69	$p=.018^{*}$	r=0.51

correct gradient estimations for each condition are (in descending order): *Rabbit Dual* =93.3% (IQR=5.8%), *Rabbit Single* =91.7% (IQR=8.3%), and *Motion Intensity* =56.7% (IQR=10.0%). All percentages are compared in Figure 3b. Since our data is not normally distributed (p<0.001), we ran a Friedman test that revealed a significant effect of condition on gradient estimation ($\chi^2(2)$=19.00, p<0.001, N=14). Post-hoc tests showed significant differences between *Rabbit Single* and *Motion Intensity* (W=102, Z=3.11, p=0.002, r=0.59) as well as *Rabbit Dual* and *Motion Intensity* (W=0, Z=-3.30, p<0.001, r=0.62). However, we did not find a significant difference between *Rabbit Single* and *Rabbit Dual* (W=30, Z=-1.42, p=0.501). Here, **we can conclude that both *Rabbit Single* and *Rabbit Dual* result in better gradient estimation performance than *Motion Intensity*.**

4.3 Task Load

The results of task load ratings as measured by the NASA RTLX [13] are shown in Figure 3c. The median (interquartile range) task load scores for each condition are (in ascending order): *Rabbit Single* =22.5 (IQR=12.7), *Rabbit Dual* =24.5 (IQR=7.9), and *Motion Intensity* =28.3 (IQR=20.0). We ran a Friedman test that revealed a significant effect of condition on task load ($\chi^2(2)$=13.50, p=0.001, N=14). Post-hoc tests showed a significant difference between *Rabbit Dual* and *Motion Intensity* (W=105, Z=3.30, p<0.001, r=0.62). However, we did not find any significant differences between *Rabbit Single* and *Rabbit Dual* (W=39, Z=0.00, p=1.000) or between *Rabbit Single* and *Motion Intensity* (W=20, Z=-2.04, p=0.120). Here, **we can conclude that *Rabbit Dual* induces a lower task load than *Motion Intensity*.**

4.4 Individual Statements and Preferences

After each condition, we asked participants to rate four statements, each on a 7-point Likert scale (1=strongly disagree, 7=strongly agree). The results and statements are shown in Figure 4. We found significant main effects for all four statements (N=14; S1: $\chi^2(2)$=14.09, p<0.001; S2: $\chi^2(2)$=11.35, p=0.003; S3: $\chi^2(2)$=17.08, p<0.001; S4:$\chi^2(2)$=12.79, p=0.002). Pairwise comparisons are shown in Table 1. Here, **we can conclude that *Rabbit Single* and *Rabbit Dual* are rated significantly more positively than *Motion Intensity* for all four statements**. No difference was found between *Rabbit Single* and *Rabbit Dual*. Regarding *overall preference*, **eight participants preferred *Rabbit Dual***, while **six voted for *Rabbit Single*** as their favorite. None of the participants preferred *Motion Intensity*.

4.5 Interviews

During the interviews, participants were explicitly asked to comment on the duration of the vibration as well as what may have eased or hindered their comprehension. They also had to explain their overall preference and comment on the overall experience and sensation of interpreting 3D directional cues via vibrotactile feedback. For the analysis, the verbal data was first transcribed by one author and then summarized. The statements were then counted for each question. In addition, across all questions, we applied open coding to identify hidden themes. Data from one interview (P2) was not recorded due to a technical issue. Therefore, only the data from 13 participants was included.

Regarding the duration of the vibration, both *Rabbit* conditions were perceived as having adequate duration (*Rabbit Single*: 10 vs 3 who thought it could have been longer; *Rabbit Dual*: 13:0), while 10 participants would have preferred a longer duration for *Motion Intensity*. For the latter, participants struggled to feel the gradient correctly, as mentioned by five participants (e.g., P7 said that the "[duration was] a little bit short, enough for [2D] direction, but for intensity [gradient] it was really bad.") The varying strength of the vibration was also an issue, as the most distant control point was criticized as having a too weak vibration, which meant that "some vibrations got lost" (P5). This also interfered with the comprehension of 2D direction. The smooth transition of movement in *Motion Intensity* was still found to be a pleasant experience, but the mentioned drawbacks regarding the gradient detection prevailed, according to P4 (RQ2). When comparing pulse with intensity for the mapping of gradient, P12 noted an interesting further advantage of pulse, as "One could decide about the gradient in retrospect even if one wasn't sure before. When the last actuator vibrated many times, then it must have been an upwards gradient." This also implicitly highlights the problem of immediacy, which required attention and did not allow repetition of the feedback. As P10 put it, "in case you did not fully pay attention, there wasn't a repeat to make sure." This sentiment was echoed by P12. Consequently, the dual encoding of a gradient in the *Rabbit Dual* condition was cited by most as the main reason for preferring that condition (RQ1). P11 noted, "I did not just have the number of pulses, but in addition the intensity and that somehow better stuck in my head."

5 DISCUSSION

Overall, responses during the interview and the quantitative data are in agreement. They show significant and substantial advantages of *Rabbit Single* and *Rabbit Dual* compared to *Motion Intensity*, which we did not expect in such clarity. The parity between these two then is again visible from all angles, with preference being nearly balanced (8 vs. 6). Still, the interviews showed that for *Motion Intensity*, participants did like the smooth transition between the individual feedback factors, which however failed to have a measurable advantage (RQ2). A main reason for this may be that

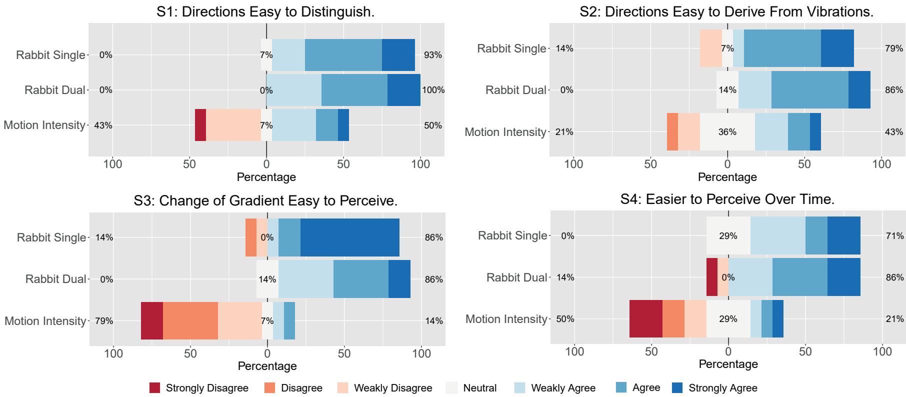


Figure 4: Participant responses to the four rated statements (Likert items ranging from 1: strongly disagree to 7: strongly agree).

we found that the mechanisms to communicate 2D direction and gradient can interfere with one another. In particular, the intensity gradient mapping had a negative effect on the 2D direction mapping in *Motion Intensity*, as the minimum vibration sometimes "got lost" (P5), when users did not pay close attention. While pre-tests suggested otherwise, individual differences among the perception of our participants as well as potential fitting issues with the glove (see *limitations* below) may have resulted in this issue. From the comments of participants, we have to assume that *Rabbit Dual* was affected by this problem as well, although to a lesser degree; the simultaneous pulse mapping implicitly included repeated vibrations of the same actuator at least twice.

Our analysis also showed that the type of feedback may require more training for participants to get accustomed to. P13 summarized this point nicely: "Maybe if you market that [...] and someone develops a game for it [...], then I might like it, and in a year, no one can imagine a world without it." Others noted the effort involved, with P10 saying they "found that it was really exhausting since you are not used to it." Still, the overall experience of using vibrotactile feedback to interpret 3D directional cues was described as "surprisingly good" (P7) and prompted many ideas for use cases, such as medical scenarios (operating table with limited visuals), people with visual impairments in daily life as well as when driving a bike or motorcycle.

Limitations: As an exploratory study, our results should be perceived as preliminary and require further testing and confirmation. In particular, our results may be limited due to the number of participants (14), which also led to the design not being fully balanced. In addition, all participants in our study were right-handed, which could affect our results. We also found that more training may be required to compensate for initial learning effects, as the type of feedback is so unusual and novel for participants. In addition, the *SensorialXR* glove only provides a fixed setting of the actuators and offers a "one-size-fits-all" size, which showed to be problematic for some users with smaller hands, where the actuators were not always in tight contact with the skin. For future research prototypes adding additional Velcro around the actuators might help.

6 CONCLUSION

This work aimed to explore different design approaches to communicate three-dimensional directional cues with vibrotactile feedback. We developed two conditions based on the *Cutaneous Rabbit* illusion and one based on *Apparent Tactile Motion* to communicate 2D direction. The gradient of the overall 3D direction was then encoded by the number of discrete vibration pulses, the vibration intensity, or a combination of both. Our study showed that three-dimensional directional cues can be communicated by *Rabbit Single* and *Rabbit Dual* with a high success rate for both the 2D direction and gradient (median for *Rabbit Single*: 91.7%, *Rabbit Dual*: 93.3%) – significantly better compared to *Motion Intensity*. With respect to our research questions, we found partial evidence for RQ1, as multiple participants specifically mentioned the dual mapping for gradient as a benefit. Still, for the quantitative data, both *Rabbit* conditions performed more or less identical. RQ2 has to be dismissed at this point. However, as revealed by our qualitative analysis, we believe that the *Apparent Tactile Motion* illusion can also be a viable option for future designs, as the smooth transition between actuators was appreciated by participants. The challenge will lie in overcoming the inferences we found between 2D directional and gradient intensity mapping.

Future Research: In our work, we aim to apply this approach to communicate the intended movements [20] of a semi-autonomous robot in collaborative scenarios, where vision alone may not be sufficient to successfully predict robot motion. In Figure 5 an assistive robot arm is illustrated, which is manually controlled by the

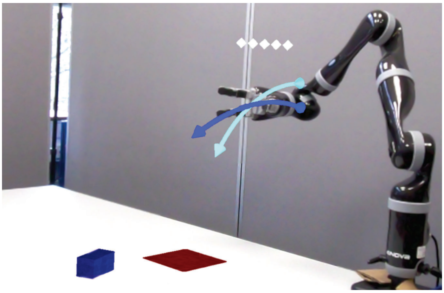


Figure 5: An assistive robotic arm with AI-created directional movement recommendations. The cyan arrow indicate the current movement direction of the arm, while the blue arrow shows the recommendation, which would be mapped as 3D directional cues on the glove. Note: The cyan and blue arrows are only for presentation purposes.

user but is supported through an Artificial Intelligence (AI) which provides real time directional movement recommendations. Here, our approach could be used to map these directional movement recommendations as vibration input on the hand. Changes in the intensity of the actuators indicate the amount of directional change, thus enabling the user to better imagine the generated trajectory. We also encourage researchers to both replicate our design and study and apply it to different use cases. Future research should also investigate variables such as the effect of higher-resolution tactile displays, different setting of actuators, or other approaches to encode gradient (e.g. through different vibration frequencies, varying linear and non-linear intensity levels), which were not possible with the *SensorialXR* technology. Furthermore, results of our study should also be evaluated with participants with a dominant left hand or their non-dominant hand.

REFERENCES

[1] David S. Alles. 1970. Information Transmission by Phantom Sensations. *IEEE Transactions on Man-Machine Systems* 11, 1 (1970), 85–91. https://doi.org/10.1109/TMMS.1970.299967

[2] Pierre Barralon, Ginna Ng, Guy A. Dumont, Stephan K. W. Schwarz, and J. Mark Ansermino. 2009. Design of Rhythm-Based Vibrotactile Stimuli Around the Waist: Evaluation of Two Encoding Parameters. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 39, 5 (2009), 1062–1073. https://doi.org/10.1109/TSMCA.2009.2025026

[3] Jennifer L. Burke, Matthew S. Prewett, Ashley A. Gray, Liuquin Yang, Frederick R.B. Stilson, Michael D. Coovert, Linda R. Elliot, and Elizabeth Redden. 2006. Comparing the effects of visual-auditory and visual-tactile feedback on user performance: A meta-analysis. In *ICMI'06 (ICMI'06: 8th International Conference on Multimodal Interfaces, Conference Proceeding)*. 108–117. https://doi.org/10.1145/1180995.1181017 null ; Conference date: 02-11-2006 Through 04-11-2006.

[4] Harold Ernest Burtt. 1917. Tactual illusions of movement. *Journal of Experimental Psychology* 2 (10 1917), 371–385.

[5] Taizhou Chen, Yi-Shiun Wu, and Kening Zhu. 2018. Investigating Different Modalities of Directional Cues for Multi-Task Visual-Searching Scenario in Virtual Reality. In *Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology* (Tokyo, Japan) *(VRST '18)*. Association for Computing Machinery, New York, NY, USA, Article 41, 5 pages. https://doi.org/10.1145/3281505.3281516

[6] R. W. Cholewiak and A. A. Collins. 2000. The generation of vibrotactile patterns on a linear array: influences of body site, time, and presentation mode. *Perception & Psychophysics* 62, 6 (2000), 1220–1235. https://doi.org/10.3758/bf03212124

[7] Frank A. Geldard. 1975. *Sensory saltation : metastability in the perceptual world / Frank A. Geldard.* Lawrence Erlbaum Associates ; distributed by the Halsted Press Division of Wiley Hillsdale, N.J. : New York. 133 p. : pages.

[8] Frank A. Geldard and Carl E. Sherrick. 1972. The Cutaneous "Rabbit": A Perceptual Illusion. *Science* 178, 4057 (1972), 178–179. https://doi.org/10.1126/science.178.4057.178 arXiv:https://www.science.org/doi/pdf/10.1126/science.178.4057.178

[9] George A. Gescheider, Stanley J. Bolanowski, Ronald T. Verrillo, D J Arpajian, and T F Ryan. 1990. Vibrotactile intensity discrimination measured by three methods. *The Journal of the Acoustical Society of America* 87 1 (1990), 330–8.

[10] Uwe Gruenefeld, Dag Ennenga, Abdallah El Ali, Wilko Heuten, and Susanne Boll. 2017. EyeSee360. In *Proceedings of the 5th Symposium on Spatial User Interaction*. ACM. https://doi.org/10.1145/3131277.3132175

[11] Stefan Grushko, Aleš Vysocký, Dominik Heczko, and Zdenko Bobovský. 2021. Intuitive Spatial Tactile Feedback for Better Awareness about Robot Trajectory during Human-Robot Collaboration. *Sensors (Basel, Switzerland)* 21, 17 (2021). https://doi.org/10.3390/s21175748

[12] Sebastian Günther, Florian Müller, Markus Funk, Jan Kirchner, Niloofar Dezfuli, and Max Mühlhäuser. 2018. TactileGlove: TactileGlove: Assistive Spatial Guidance in 3D Space through Vibrotactile Navigation. In *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference*. ACM, New York, NY, USA, 273–280. https://doi.org/10.1145/3197768.3197785

[13] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology*. Elsevier, Amsterdam, The Netherlands, 139–183. https://doi.org/10.1016/s0166-4115(08)62386-9

[14] Oliver Beren Kaul and Michael Rohs. 2016. HapticHead: 3D Guidance and Target Acquisition through a Vibrotactile Grid. *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (2016).

[15] Jacob H. Kirman. 1974. Tactile apparent movement: The effects of interstimulus onset interval and stimulus duration. *Perception & Psychophysics* 15, 1 (1974), 1–6. https://doi.org/10.3758/BF03205819

[16] Seungyon Claire Lee and Thad Starner. 2010. BuzzWear: Alert Perception in Wearable Tactile Displays on the Wrist. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 433–442. https://doi.org/10.1145/1753326.1753392

[17] Ville Lehtinen, Antti Oulasvirta, Antti Salovaara, and Petteri Nurmi. 2012. *Dynamic Tactile Guidance for Visual Search Tasks.* Association for Computing Machinery, New York, NY, USA, 445–452. https://doi.org/10.1145/2380116.2380173

[18] Troy McDaniel, Morris Goldberg, Daniel Villanueva, Lakshmie Narayan Viswanathan, and Sethuraman Panchanathan. 2011. Motor learning using a kinematic-vibrotactile mapping targeting fundamental movements. In *Proceedings of the 19th ACM international conference on Multimedia*. ACM.

[19] Gunhyuk Park and Seungmoon Choi. 2018. Tactile Information Transmission by 2D Stationary Phantom Sensations. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3173574.3173832

[20] Max Pascher, Uwe Gruenefeld, Stefan Schneegass, and Jens Gerken. 2023. How to Communicate Robot Motion Intent: A Scoping Review. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems - CHI '23*. https://doi.org/10.1145/3544548.3580857

[21] Jukka Raisamo, Roope Raisamo, and Veikko Surakka. 2009. Evaluating the effect of temporal parameters for vibrotactile saltatory patterns. In *Proceedings of the 2009 international conference on Multimodal interfaces*. ACM. https://doi.org/10.1145/1647314.1647381

[22] Ismo Rakkolainen, Ahmed Farooq, Jari Kangas, Jaakko Hakulinen, Jussi Rantala, Markku Turunen, and Roope Raisamo. 2021. Technologies for Multimodal Interaction in Extended Reality - A Scoping Review. *Multimodal Technologies and Interaction* 5, 12 (2021). https://doi.org/10.3390/mti5120081

[23] Carl E. Sherrick and Ronald Rogers. 1966. Apparent haptic movement. *Perception & Psychophysics* 1 (1966), 175–180.

[24] Daniel Spelmezan, Mareike Jacobs, Anke Hilgers, and Jan Borchers. 2009. Tactile Motion Instructions for Physical Activities. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 2243–2252. https://doi.org/10.1145/1518701.1519044

[25] Hong Z. Tan, Robert Gray, J. Jay Young, and Ryan Traylor. 2003. A Haptic Back Display for Attentional and Directional Cueing. Vol.3.No.1.,Haptics-e, 11-June-2003 (2003).

[26] B. Weber, S. Schatzle, T. Hulin, C. Preusche, and B. Deml. 2011. Evaluation of a vibrotactile feedback device for spatial guidance. In *2011 IEEE World Haptics Conference*. IEEE, 349–354. https://doi.org/10.1109/WHC.2011.5945511

[27] Juan Wu, Yu Ding, Dejing Ni, Guangming Song, and Wei Liu. 2013. Vibrotactile Representation of Three-Dimensional Shape and Implementation on a Vibrotactile Pad. *Sensors and Materials* (2013), 79–97. https://doi.org/10.18494/sam.2013.855

AdaptiX – A Transitional XR Framework for Development and Evaluation of Shared Control Applications in Assistive Robotics

MAX PASCHER, TU Dortmund University, Germany and University of Duisburg-Essen, Germany
FELIX FERDINAND GOLDAU, German Research Center for Artificial Intelligence (DFKI), Germany
KIRILL KRONHARDT, TU Dortmund University, Germany
UDO FRESE, German Research Center for Artificial Intelligence (DFKI), Germany
JENS GERKEN, TU Dortmund University, Germany

(a) (b) (c)

Fig. 1. Setup with (**a**) a user's view in the Virtual Reality (VR) simulation environment, (**b**) setup of interaction with a physical robot, and (**c**) a combined view of physical robot and visual cues in Mixed Reality (MR).

With the ongoing efforts to empower people with mobility impairments and the increase in technological acceptance by the general public, assistive technologies, such as collaborative robotic arms, are gaining popularity. Yet, their widespread success is limited by usability issues, specifically the disparity between user input and software control along the autonomy continuum. To address this, shared control concepts provide opportunities to combine the targeted increase of user autonomy with a certain level of computer assistance. This paper presents the free and open-source *AdaptiX* XR framework for developing and evaluating shared control applications in a high-resolution simulation environment. The initial framework consists of a simulated robotic arm with an example scenario in Virtual Reality (VR), multiple standard control interfaces, and a specialized recording/replay system. *AdaptiX* can easily be extended for specific research needs, allowing Human-Robot Interaction (HRI) researchers to rapidly design and test novel interaction methods, intervention strategies, and multi-modal feedback techniques, without requiring an actual physical robotic arm during the early phases of ideation, prototyping, and evaluation. Also, a Robot Operating System (ROS) integration enables the controlling of a real robotic arm in a *PhysicalTwin* approach without any simulation-reality gap. Here, we review the capabilities and limitations of *AdaptiX* in detail and present three bodies of research based on the framework. *AdaptiX* can be accessed at https://adaptix.robot-research.de.

Authors' addresses: Max Pascher, TU Dortmund University, Emil-Figge-Str. 50. 44227 Dortmund, Germany, max.pascher@udo.edu; Felix Ferdinand Goldau, Enrique-Schmidt-Str. 5, 28359 Bremen, Germany, felix.goldau@dfki.de; Kirill Kronhardt, TU Dortmund University, Emil-Figge-Str. 50. 44227 Dortmund, Germany, kirill.kronhardt@udo.edu; Udo Frese, Enrique-Schmidt-Str. 5, 28359 Bremen, Germany, udo.frese@dfki.de; Jens Gerken, TU Dortmund University, Emil-Figge-Str. 50. 44227 Dortmund, Germany, jens.gerken@udo.edu.

ACM 2573-0142/2024/6-ART244
https://doi.org/10.1145/3660243

CCS Concepts: • **Computer systems organization** → **Robotic control**; • **Human-centered computing** → *Visualization techniques*; *Virtual reality*.

Additional Key Words and Phrases: assistive robotics, human–robot interaction, shared user control, augmented reality, virtual reality, mixed reality, visual cues

ACM Reference Format:
Max Pascher, Felix Ferdinand Goldau, Kirill Kronhardt, Udo Frese, and Jens Gerken. 2024. AdaptiX – A Transitional XR Framework for Development and Evaluation of Shared Control Applications in Assistive Robotics. *Proc. ACM Hum.-Comput. Interact.* 8, EICS, Article 244 (June 2024), 28 pages. https://doi.org/10.1145/3660243

1 INTRODUCTION

Robotic arms as assistive technologies are a powerful tool to increase self-sufficiency in people with limited mobility [33, 44], as they facilitate the performance of Activities of Daily Living (ADLs) – usually involving grasping and manipulating objects in their environment – without human assistance [50]. However, a frequent point of contention is the assistive robot's autonomy level. The reduction of user interaction to just oversight with purely autonomous systems elicits stress [51] and feelings of distrust in their users [67]. On the other side of the autonomy spectrum, manual controls can be challenging - or even impossible - to operate, depending on the significance and type of impairment. Shared control – a combination of manual user control through standard input devices plus algorithmic support through computer software adjusting the resulting motion – may have the potential to mitigate both concerns [1]. Here, both the user and the robot share a task on the operational level, enabling people with motor impairments to get involved in their assistance. As a result, such approaches can increase the feeling of independence while improving ease of use compared to manual controls [17].

A characteristic real-world scenario, motivated by our research, has an assistive robotic arm (e.g., a Kinova Jaco 2) attached to a wheelchair to support the user in ADLs. Here, the user is challenged with operating six or more Degrees-of-Freedom (DoFs), which requires complex input devices or time-consuming and confusing mode switches. This potentially results in increased task completion time and user frustration [21]. Addressing this, shared control systems can facilitate more straightforward and accessible robot operation. However, they may require well-designed communication of robot (motion) intent, so that the user retains awareness and understands the level of support they get from the system [45]. Also, different users might need distinct input devices or require multi-modal input to account for varying abilities.

Based on our experiences, we identified several challenges that currently influence and potentially impede the effective development of shared control approaches:

- Shared control systems for assistive technologies still pose open questions requiring considerable experimentation, tweaking and balancing between user and robot interaction [34].
- While much research explored robot motion intent, there is little insight into what works best in which situation and for which type of user. In assistive robotics, the visualization and feedback modality must be carefully adapted to the user's needs and abilities as there is no "one size fits all" solution [23].
- Similarly, suitable input devices may vary between users. Depending on individual preferences and capabilities, multi-modal input or the choice between different input modalities may be required [2].
- Bringing robots and humans physically together during research studies is difficult due to the laborious and costly transportation, safety concerns with robots and general availability of the user group [6].

Contribution. To allow researchers, designers and developers to address these challenges holistically and flexibly, we present *AdaptiX* – a free, open-source XR framework [1]. Aimed at Design and Development (D&D), *AdaptiX* combines a physical robot implementation with a 3D simulation environment. The simulation approach (analogous to simulations in industrial settings [37, 42, 59]) mitigates the assistive robotic arm's bulky, expensive, and complex nature. It also makes the integration of visualization feedback or different input modalities easier to explore and test, while a Robot Operating System (ROS) interface allows the direct transfer to the real robot. Testing new interaction and control options becomes much less time-consuming while simultaneously excluding potentially dangerous close-contact situations with users before glitches are managed [44]. In total, the framework facilitates the development and evaluation of assistive robot control applications *in-silico* and creates a practical and effective step between ideation, development, and evaluation, allowing HRI researchers more flexibility and facilitating efficient resource usage.

To summarize, the *AdaptiX* framework contributes the following:

- *AdaptiX* allows researchers to rapidly design and test novel visualization and interaction methods.
- The framework integrates an initial concept and implementation of a shared control approach.
- The integrated ROS interface facilitates connection to a non-simulated – physical – robotic arm to perform bidirectional interactions and data.
- The framework's concept enables a code-less trajectory programming by hand-guiding the simulated or physical assistive robotic arm to the specific location and saving the position and orientation of the Tool Center Point (TCP).
- Recording TCP data enables replaying user-controlled robot movements and results in a fully customizable system. Options include changing specific details during replaying, such as repositioning cameras or re-rendering background scenes.
- Finally, the entire continuum of Mixed Reality (MR) can be exploited in the *AdaptiX* environment. This allows applications in Virtual Reality (VR), pure screen space, Augmented Reality (AR), simultaneous simulation and reality, and pure reality (cf. the *virtuality continuum* of Milgram and Kishino [41]).

2 RELATED WORK

While robotic arms are a particularly useful and versatile subset of assistive technologies, their widespread success is limited by a number of design challenges concerning the interaction with their human user. In recent years, a growing body of research addressed these concerns and associated optimization options to increase their usability, e.g., [12, 20, 34]. During the *AdaptiX* development process, we aimed to include functionality to address the challenges of shared control optimization [19], intent communication [45], and attention guidance [48].

2.1 Shared Control for Assistive Robots

Current shared control systems operate along an autonomy continuum, respectively balancing user input and system adjustments. At one extreme, the systems tend to be heavily manual, with only minor adjustments to the user's input [56]. On the other end are systems where users primarily provide high-level commands for the robot to execute [60]. A number of different approaches – including time-optimal [21] and blended mode switching [16], shared-control-templates [52] and body-machine-interfaces [29] – are currently employed in various settings.

[1] *AdaptiX* framework. https://adaptix.robot-research.de, last retrieved May 20, 2024.

A fundamentally different approach is the shared control system proposed by Goldau and Frese [19]. Their concept combines a robotic arm's cardinal DoFs according to the current situation and maps them to a low-DoF input device. The mapping is accomplished by attaching a camera to the robotic arm's gripper and training a Convolutional Neural Network (CNN) by having people without motor impairments perform ADLs [19] – similar to the learning-by-demonstration approach for autonomous robots by Canal et al. [7]. The CNN returns a set of newly mapped DoFs, ranked by their assumed likeliness based on the CNN for the given situation, allowing users to access a variety of movements for each situation. In addition, the CNN-based approach allows the system to be easily extendable as the same system can be trained to discriminate between many different situations – making it a viable concept for day-to-day use. Goldau and Frese [19] conducted a proof-of-concept study comparing the control of a simulated 2D robot with manual or CNN-based controls. Task execution was faster with their proposed concept; however, users experienced it as more complex than manual controls [19].

Our framework *AdaptiX* is influenced by Goldau and Frese's approach, but extends it from 2D to 3D space. This increases the number of possible DoFs, which allows for an accurate representation of ADLs in the framework. By adding functionality, visualizations, and a ROS integration, *AdaptiX* can be used to develop and evaluate novel interaction control methods based on this approach for shared control, which we refer to as *Adaptive DoF Mapping Control (ADMC).*

2.2 Robot Motion Intent

Regardless of the specific interaction details, it is necessary to effectively communicate the intended assistance provided by the (semi-)autonomous system [4]. Clear communication between robots and humans enhances the shared control system's predictability, avoids accidents, and increases user acceptance.

A crucial element of the D&D process of robotic devices is, therefore, the testing of intent communication methods. *Choreobot* – an interactive, online, and visual dashboard – proposed by van Deurzen et al. [61] supports researchers and developers to identify where and when adding intelligibility to the interface design of a robotic system improves the predictability, trust, safety, usability, and acceptance. Moreover, Pascher et al. [45] provide an extensive overview of the various types of visualization and modalities frequently used in communicating robot motion intent. These range from auditory [10] and haptic [9] modalities to anthropomorphizing the robot and using its gaze [38] or gestures [18]. Their findings are substantiated by Holthaus et al. [24], who used an ethnographic approach to derive a comprehensive communication typology.

While all these intent communication modalities are viable, visual representations of future movements are often quoted as less workload-intense for the end-user [13]. AR is, therefore, unsurprisingly a frequently used tool to convey detailed motion intent [8, 22, 53, 63, 65], allowing interactions to become more intuitive and natural to humans [36]. Suzuki et al. emphasize the benefits of AR-based visualizations for communicating movement trajectories or the internal state of the robot [58].

The visual feedback employed by *AdaptiX* mimics AR in a VR environment with directional cues registered in 3D space. This approach allows the user to understand different movement directions for the actual control and the suggested DoF combinations. To streamline understanding the control methods, one of our primary approaches is the usage of arrows – a straightforward and common visualization technique to communicate motion intent [54, 55, 63].

2.3 Feedback Modalities for User Attention Guidance

When creating systems using shared control, it is crucial to guide the user's focus to the assistance the robot is offering [49]. This guidance is particularly important if either party is moving the robot in a way that could lead to collisions or worsen the situation. To enhance the predictability of shared control systems, various feedback modalities have been proposed to guide user attention as a secondary feedback mechanism to AR. The goal is to provide a feedback solution that results in short reaction times, enabling users to quickly direct their focus to the information provided by the robot.

In the related discipline of autonomous driving systems, if the vehicle encounters a situation it was not programmed or trained to handle, it will issue a Take-Over-Request (TOR). This TOR prompts the driver to take manual control of the vehicle to prevent a collision or to drive in areas the vehicle cannot handle autonomously.

Auditory, visual, and tactile/haptic modalities are commonly used for TORs [64] – either as a single sensory input [49] or a combination of multiple variants [48]. Simulation studies, along with research on reaction times to different sensory stimuli, indicate that multi-modal feedback results in the lowest possible reaction times in shared control systems [5, 14, 31].

Implementing these feedback methods into existing assistive robot systems would be straightforward as the necessary output devices – like screens, speakers, or vibration motors – are commonly already present. To allow researchers to evaluate the benefits of the different modalities, *AdaptiX* includes three modes for attention guiding: visual, auditory, and tactile/haptic. Developers can either choose one modality or follow a multi-modal approach.

3 FRAMEWORK CONCEPT

The *AdaptiX* XR framework facilitates the development and evaluation of HRI shared control applications in an easy-to-use, high-resolution transitional MR environment. Equipped with a VR simulation environment containing a virtual *Kinova Jaco 2* and ample customization options, researchers can streamline their D&D process while simultaneously reducing overhead and boosting efficiency. Figure 2 provides an overview of the framework's architecture.

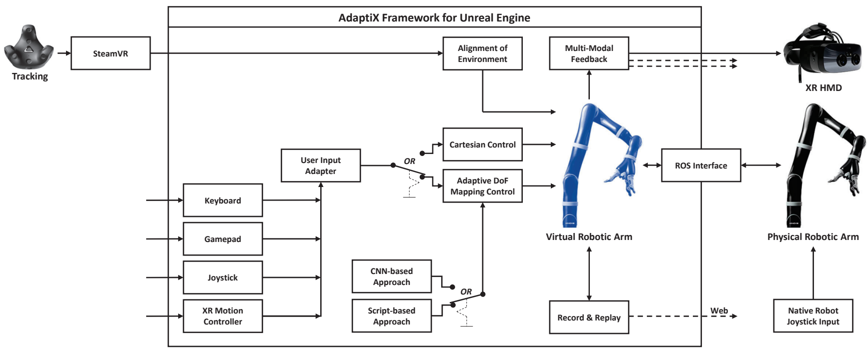


Fig. 2. Overview of *AdaptiX*' architecture, illustrating each component, their directional communication, and the crossover from and to the framework. The user input is either used for *Cartesian Control* or Adaptive DoF Mapping Control (ADMC). For ADMC, either a CNN-based or script-based rule engine can be selected.

In addition to an Cartesian robot control, we propose ADMC as an initial shared control approach, using suggestions by a rule engine (e.g., a CNN or script-based approach) to be controlled by the user. ADMC is implemented directly into the *Unreal Engine* to enable researchers and developers to fully customize the control methods, systems behavior, and feedback techniques by coding in *C++* or *Blueprints*.

AdaptiX supports several pre-implemented input devices and provides an adapter class for an easy development and implementation of further input devices. This supports researchers and developers to easy implement their ideas and concepts. The integrated ROS interface facilitates connection to a non-simulated – physical – robotic arm to perform bidirectional interactions and data exchange in a *DigitalTwin* and *PhysicalTwin* approach.

AdaptiX enables effortless trajectory programming by manually guiding the TCP of a simulated or physical robotic arm to a desired location and recording its position and orientation. Recorded data of user-controlled robot movements can be replayed. Offering the adjustment of specific details, such as camera positions and background scenes, results in a highly customizable system.

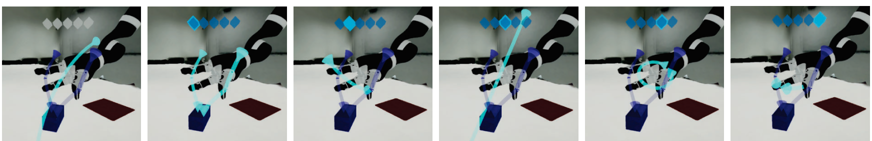
The aim is to provide a modular and extensible framework so that research teams do not need to start from scratch when implementing their shared control applications.

3.1 Adaptive DoF Mapping Control (ADMC)

For the adaptive DoF mapping – referred to as *ADMC* – of the robotic arm, the goal is to present a set of DoF mappings ordered based on their effectiveness in accomplishing the pick-and-place task used in the experiment. The concept of "usefulness" assumes that maximizing the cardinal DoFs of the robot assigned to an input-DoF while progressing towards the next goal is the most advantageous option.

This DoF mapping, referred to as the *optimal suggestion*, is assumed to be the best choice due to a significant reduction in the need for mode switches when multiple DoFs are combined into a single movement. The more DoFs are combined (assuming it is sensible for the given situation), the fewer mode switches are required. As a result, the DoF mappings are ordered based on the number of DoFs they combine.

In addition to the optimal suggestion, the second suggestion is a selection of an orthogonal variation of the first suggestion, which has the highest probability and most variation in spatial direction and keeps the number of combined DoFs unchanged. This secondary suggestion is likely useful to users as they can utilize it to adjust their position while maintaining a sensible orientation toward the next goal. The following DoF mappings were used (see Figure 3):



(a) (b) (c) (d) (e) (f)

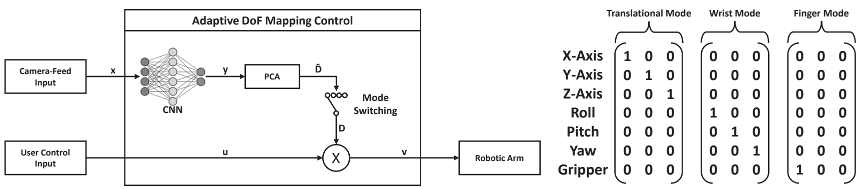
Fig. 3. Suggestions as Visualized in the ADMC, (**a**) Continue previous movement, (**b**) Optimal Suggestion, (**c**) Adjustment Suggestion, (**d**) Translation Suggestion, (**e**) Rotation Suggestion, (**f**) Gripper Suggestion. Colors: Bright cyan arrow: Currently active DoF mapping. Dark blue arrow: Next most likely DoF mapping.

(1) *Optimal Suggestion:* Combining translation, rotation, and finger movement [opening and closing] into one suggestion, causing the gripper to move towards the target, pick it up, or release it on the intended surface.

(2) *Adjustment Suggestion:* An orthogonal suggestion based on (1) but excluding the finger movement. Allows the users to adjust the gripper's position while still being correctly orientated.
(3) *Translation Suggestion:* A pure translation towards the next target, disregarding any rotation.
(4) *Rotation Suggestion:* A pure rotation towards the next target disregarding any translation.
(5) *Gripper Suggestion:* Opening or closing of the gripper's fingers.

3.1.1 CNN-based Approach. For the CNN approach, a color-and-depth camera is attached to the gripper of an assistive robotic arm. The live video feed is transmitted to a CNN, which is trained using data collected from non-impaired individuals performing ADLs using the robotic arm along with a high-DoF input device. The CNN does not need a model of the environment to provide these mappings. Principal Component Analysis (PCA) is employed to transform the CNN's output into a matrix $\hat{D}$, where each column represents a combination of cardinal DoFs along which the robotic arm can move.

Next, a subset of $\hat{D}$ is selected, containing as many columns as the number of DoFs provided by the input device. This selected subset is referred to as D, and it serves to map input-DoFs to output-DoFs. When an input-DoF is engaged, the robot's movements are determined by the values in the corresponding vector of D, which proportionally activate the robot's cardinal DoFs. A mode switch is defined as the exchange of D with a different subset of $\hat{D}$. This enables the system to switch between various mappings of input-DoFs to output-DoFs, adapting the robot's control according to the user's needs and preferences. A visual representation of this control pipeline is depicted in Figure 4a.



(a) (b)

Fig. 4. Concept of adaptive DoF mapping control. **(a)** Control pipeline for proposed adaptive shared control and **(b)** matrix representation of DoF mappings: Columns represent input-DoFs. Rows represent output-DoFs. Subsets represent modes. Two empty columns were added to represent zero movement mappings in *Finger Mode*.

$\hat{D}$ is a square matrix with dimensions based on the number of cardinal DoFs available on the robot to be controlled. In the case of the *Kinova Jaco 2* [30], this results in a 7×7 matrix. This matrix represents a mapping of input-DoFs to output-DoFs when the number of DoFs in both cases is equal. The values in each column, ranging from -1 to 1, indicate the proportion with which the specific cardinal DoF is utilized when engaging the corresponding input-DoF.

By defining $\hat{D}$ as an identity matrix, each input-DoF is mapped to a single output-DoF. Selecting an equal number of columns from $\hat{D}$ to form matrix D allows for manual control with mode switching along cardinal DoFs. Moreover, this representation enables the combination of multiple cardinal movements into arbitrary output DoF mappings. For example, a (transposed) column of *(0.5, 0.5, 0, 0, 0, 0, 0)* would result in diagonal movement along the X- and Y-Axes of the robot. Such combinations enable the offering of complex movements with different proportions depending on the situation, enhancing the control options available to users. The identity matrix for a *Kinova Jaco 2* with a 3-DoFs joystick is illustrated in Figure 4b.

3.1.2 Script-based Approach. As an alternative rule engine for our ADMC concept, we implemented a task-specific script. This approach eliminates potential biases that a more generic, but currently limited method like a CNN-based control might introduce. It is essential to note that our task-specific script is effective only in a controlled experimental environment.

The task-specific script assesses the end effector's current position, rotation, and finger position relative to a target, allowing it to adaptively calculate the matrix $\hat{D}$. This script recommends optimal movements to pick up an object and place it onto a target drop area, maximizing the combination of as many DoFs as possible. Additionally, it provides other DoF combinations that may be less beneficial to mimic the idea that each subsequent column in $\hat{D}$ has a decreasing likelihood of being useful. These additional DoF mappings are ordered by the number of combined DoFs in a decreasing manner.

To validate the effectiveness of this approach, we conducted pilot tests, comparing it to a *Wizard-of-Oz* method. In this scenario, a human "simulated a CNN" to explore user interaction with such a system.

3.1.3 Point of Time to Communicate the Suggestion. Our ADMC concept uses an adaptive DoF mapping system to recommend DoF mappings to the users depending on the current situation. The system visualizes the currently active DoF mapping as a bright cyan and the suggestion as a dark blue arrow (see Figure 3). This suggestion can be communicated – based on the the configuration – either continuously or only if the next most likely movement direction differs from the currently active DoF mapping by a certain threshold.

To calculate this threshold – the difference between the currently active and new most likely DoF mapping –, *cosine similarity* [57] is used, ranging from exact alignment [0%] to total opposite direction [100%]. The formula for cosine similarity of two n-dimensional vectors is defined as:

$$\text{cosine similarity} = \cos\left(\vec{a}, \vec{b}\right) = \frac{\vec{a}\vec{b}}{\|\vec{a}\|\|\vec{b}\|} = \frac{\sum_{i=1}^{n} a_i b_i}{\sqrt{\sum_{i=1}^{n} (a_i)^2}\sqrt{\sum_{i=1}^{n} (b_i)^2}} \tag{1}$$

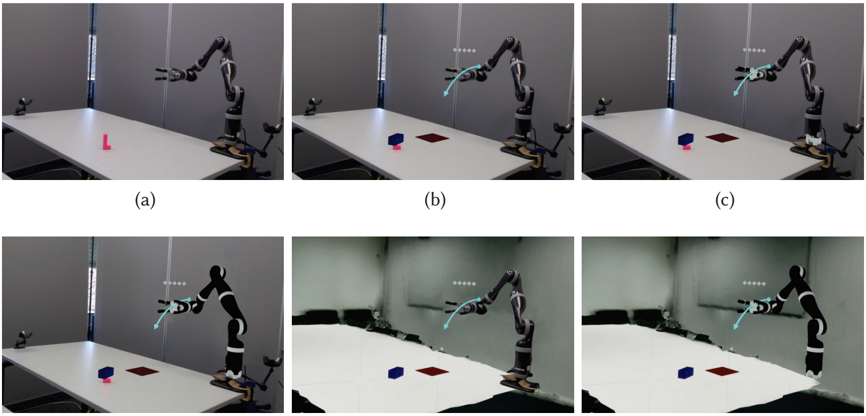
To implement a difference value, the cosine similarity needs to be transformed. As a cosine similarity of -1 indicates completely opposed vectors, the difference value needs to return 1 – i.e. the maximum possible difference – for a cosine similarity value of -1. A cosine similarity of 1, indicating exact similarity, should return a difference value of 0 – i.e. no difference. Perpendicular vectors with cosine similarity 0 should return a difference value of 0.5 – i.e. a 50% difference. To calculate the difference value **d**, the following formula is used:

$$\text{difference d} = 1 - \frac{\cos\left(\vec{a}, \vec{b}\right) + 1}{2} \tag{2}$$

This difference value represents the difference between two vectors. While the user moves the robot with an active DoF mapping, the adaptive DoF mapping system reevaluates the situation and calculates new suggested DoF mappings. The default difference value is set to 0.2 (20% difference between currently active and new most likely DoF mapping).

3.2 Full Mixed Reality Continuum

In our framework, we created an environment in which the entire continuum of MR is exploitable. This extends the use of *AdaptiX* to new scenarios and environments – including the real world. The



(a) (b) (c)

(d) (e) (f)

Fig. 5. MR continuum with (**a**) only the real robotic arm in real environment, (**b**) augmenting of directional cues in the real environment with the real robotic arm, (**c**) additional visualizing the gripper and base of the virtual robotic arm in the real environment, (**d**) visualizing the simulated robotic arm in the real environment, (**e**) visualizing the real robotic arm in the virtual environment, and (**f**) visualizing the simulated robotic arm in the virtual environment.

virtual and real environments of the robotic arm are aligned, allowing researchers to seamlessly switch between the user controlling the real and virtual robot. The level of MR can be adjusted in various steps (cf. the *virtuality continuum* of Milgram and Kishino [41]).

The MR environment setups include:

(1) the completely real environment with the real robotic arm,
(2) the real environment extended with visual cues,
(3) the real environment into which the virtual robot is transferred and displayed (with and without visual cues),
(4) the virtual environment into which the real robot is transferred and displayed (with and without visual cues),
(5) the completely virtual environment with the virtual robotic arm.

A comparison of the user's view in reality and simulation can be seen in Figure 5. MR continuum level (1) is suitable for study baseline-condition, without any multi-modal feedback to the user. In level (2) an AR visualization technique is mimicked, showing the whole physical setup augmented by basic cues. Especially level (3) and (4) enable customizing either the robot itself or the environment to extent/exchange the physical setup but still not loosing the context. In (3) users can interact with a totally new or customized robot while being in a familiar environment. World's distractions can be excluded in (4) while the the original robot is presented. Finally, level (5) provides a VR environment that can be fully customized.

3.3 Interfaces

We designed *AdaptiX* to facilitate the comparison of different interaction designs, intervention strategies, and feedback techniques for shared robot control. The initial version of the framework includes interface types for extending user input, ROS integration, and multi-modal feedback. However, this baseline can easily be customized and extended by future development.

3.3.1 User Input. We provide a standard control approach where pressing a keyboard button moves the end effector along cardinal DoFs (x, y, z, roll, pitch, yaw, opening and closing the gripper). Using build-in functionalities, the designated keyboard input can easily be adjusted to other input devices like gamepads, joysticks, or customized assistive input appliances.

In contrast to tele-operating the robotic arm, a *follow-me* approach for any trackable object in 3D space – e.g., the user's handheld VR motion controller – was implemented. The robot's end effector directly follows the movement of the trackable object, which corresponds functionally to direct control. This can be used to generate high-dimensional input and record intended behavior quickly, providing an easy way of interacting and controlling the robot, especially for inexperienced users.

3.3.2 ROS Integration. The ROS integration allows for a bidirectional exchange of information between the simulation and a real robot, mirroring the robot's state *in-silico* and vice versa. Figure 6 shows the involved components: a ROS bridge facilitates the multi-device connection between the framework and the real robot while exchanging robot data. On the ROS side, the messages for the arm position and orientation control and the values for the angle-accurate control of the gripper fingers are read in via the ROS subscriber node. They are then processed, and the robot arm and gripper are controlled through our action client. In addition, the joint angles, the TCP, and the position of all three gripper fingers are published via ROS, which are then input by our *Unreal Engine* framework. The virtual and real robots are synchronized via ROS every 0.1 seconds.

Based on this, our framework provides – depending on the specific context – both a *DigitalTwin* and *PhysicalTwin* approach, allowing the control of either with the other.

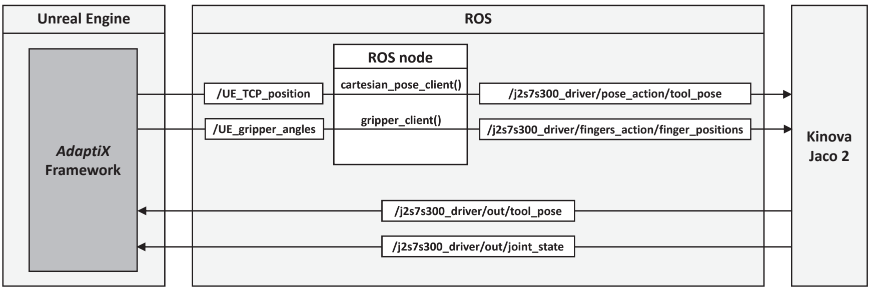


Fig. 6. Component connections of the ROS interface for mixed reality.

3.3.3 Multi-Modal Feedback. To communicate any combination of DoFs, our framework supports several visual cues to illustrate the intended movement trajectory and provides multi-modal feedback extensions via audio and haptic-tactile feedback. Visual feedback can be either provided dynamically attached to the virtual/physical robot's end effector, stationary in the world, or attached to the user's view.

AdaptiX aims to support the development of novel multi-modal interaction and feedback designs either in the pure VR simulation testbed environment or by interacting with a real robot in MR, which mimics an AR setting due to the stereoscopic video-feed. Moreover, it is also possible to show the real robot in our VR simulation environment instead of the simulated one.

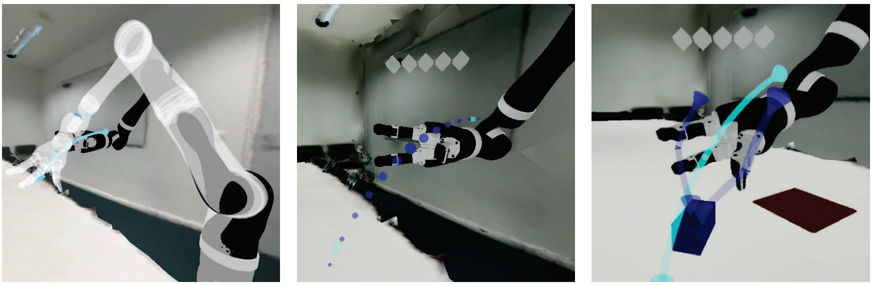
Figure 7 shows three exemplary AR-style visualizations provided by the framework, including (a) a robotic ghost overlay, (b) discrete waypoints in 3D, and (c) a variety of multidimensional arrows. Though varying in design, these visualizations can effectively communicate the robot's motion intent to the user.

Ghost: A visualization of robot motion intent by showing an additional version of the robot (or specific components) registered in 3D space, in another color and/or opacity. These visualizations communicate the exact position and orientation a robot at a given time, behaving precisely as though the real robot had been moved this way.

Waypoints: This visualization technique augments the position of a robot (or in our case, the gripper of the robotic arm) in 3D space at a certain point in the future. Usually, the robot navigates linearly between these *Waypoints*, which increases predictability.

Arrow: Among visualizations arguably the most basic but certainly also the most familiar (as seen in traffic navigation systems, road signs, and on keyboards). *Arrows* are found both in straight and curved varieties, where curved arrows indicate a rotation. Given the abundance of *Arrows* in daily life, it makes sense that many robot motion intent visualizations use them.

Classic: This visualization also uses *Arrows*, but in our prototype they are used as a baseline condition to evaluate adaptive and non-adaptive controls. Here, as with the standard input device *Kinova Jaco 2*, two axes can be controlled simultaneously and the user has to choose between different translations and rotations by mode-switching.



(a) Ghost (b) Waypoints (c) Arrows

Fig. 7. Visualization examples pre-implemented in the framework.

All interfaces are modular, enabling quick adaptations and switching between variations. This flexibility allows for studies with clean methodologies and easy comparisons without additional overhead. The community is invited to extend the implementations with any interfaces or control methods desired for their research.

3.4 Recording and Replay

AdaptiX contains an easy-to-use general-purpose system to record, store and replay simulation data, including detailed information about robot states, execution times, or the states of various objects in the environment. The recording system generates Comma-separated values (CSV) text files, which can be accessed with any data manipulation software (e.g., Python or MATLAB). The added output functionality differs significantly from the replaying system provided by the underlying *Unreal Engine*, which is mainly designed for visual replays and – among other things – does not support a CSV file format.

In addition, *AdaptiX*'s recording and replaying system is entirely customizable. Camera repositioning or re-rendering background scene options are included in the initial version. By default, the recording system tracks the user's view, the robotic arm, and all moveable actors in the virtual environment. All other objects are assumed to be stationary, thus part of the level, and ignored as such. This approach allows for the randomization of background scenes by re-rendering.

The system stores the assigned virtual meshes, scales, possible segmentation tags for each tracked object, and the complete pose data per frame. During the replay process, all objects that were initially recorded in a specific level are swapped with the corresponding data stored in the loaded recording. However, if a different scene is being loaded, the objects from that scene are used instead. In every subsequent frame, all objects are positioned at their respective position until the loaded recording has finished. The system permits custom code to be run at the end of each loaded frame, thus enabling de-bugging and data rendering during replays.

Overall, *AdaptiX* facilitates the lightweight storage of recordings as CSV files with the option to render and store complex and large-scale data (e.g., images or videos) for subsequent evaluation. This lightweight approach is particularly useful when deploying experiments on external devices or recording extensive datasets.

4 FRAMEWORK IMPLEMENTATION

The *AdaptiX* simulation environment is based on the game engine *Unreal Engine 4.27* [15]. The advanced real-time 3D photoreal visuals and immersive experiences provide a suitable foundation for our framework, and assets for future extensions are readily available. *Unreal Engine 4.27* includes integrated options for various hardware setups, thus enabling the framework to be deployed on different operating systems while utilizing most currently available VR/MR/AR headsets, gamepads, and joysticks. At the time of writing, *Unreal Engine 4.27* is free to use, has a considerable user space, and allows unrestricted publications of non-revenue generating research products like the *AdaptiX* framework. Detailed implementation descriptions can be accessed in the *README* provided in the repository at https://adaptix.robot-research.de.

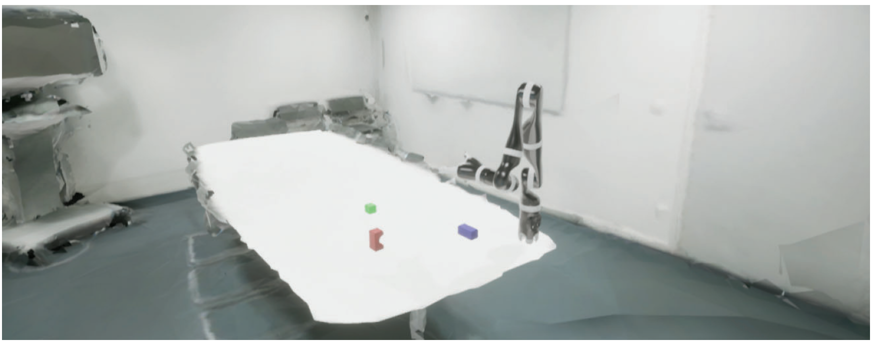


Fig. 8. Example scenario provided in *AdaptiX* including a table, a virtual *Kinova Jaco 2* robotic arm and colored blocks on the tabletop.

4.1 Simulation Environment

The *AdaptiX* default scenario centers on the photogrammetry scan of an actual room that contains a table with an attached virtual robotic arm (see Figure 8). A simulated camera is mounted on the arm's gripper. We added a toggle-off option to hide the camera from the user's view.

The framework includes a straightforward testbed scenario for pick-and-place operations, mimicking the basic principles of most ADLs. The simulation centers around a red surface as a drop target and a blue block as the to-be-manipulated object. Once the object has been successfully placed, the setup randomly re-positions the blue block on the table surface, and the task can be repeated.

We optimized the robotic arm simulation for operation via a VR motion controller with an analog stick, several playable buttons, and motion capture capabilities (e.g., *Meta Quest 2* [39]). These options provide a workable foundation to implement and test diverse interaction concepts, including adaptive concepts which can be configured to match the individual physical abilities of the intended user.

By incorporating the *Varjo XR-3* [62] – a particularly high-resolution XR-Head-Mounted Display (HMD) – we implemented a transitional MR environment. Using two *HTC VIVE* trackers [26], the virtual and real worlds are synchronized so that the robots' working areas are identical. By including the *HTC VIVE* motion controller [25], it is then possible to control the physical robot directly via the *PhysicalTwin* approach of *AdaptiX* (see Figure 1).

The virtual robotic arm is designed as a modular entity, allowing easy integration to new levels following the *Unreal Engine*'s *ActorBlueprint* class structure.

4.1.1 Simulated Robotic Arm. The commercially available *Kinova Jaco 2* assistive robotic arm [30] is specifically designed as an assistive device for people with motor impairments. It is frequently used by a) the target audience and b) researchers – e.g., [3, 21] – during HRI studies, hence the suitability for inclusion in *AdaptiX*.

We designed the simulated *Kinova Jaco 2* as close as possible to the actual product, using virtual meshes generated directly from computer-aided design (CAD) files provided by the manufacturer. Much like in reality, the virtual arm consists of a series of individual links connected by angular joints as shown in the annotated rendering of the assembled model Figure 9.

As *AdaptiX* – including the operation of its simulated robotic arm – is optimized for HRI studies, it focuses on user interaction rather than low-level robot control, whilst also able to incorporate those. Hence, rather than following the standard base-up control, the simulated arm moves in reverse: the user's input directly controls the end effector's motion; the connected joints are positioned to connect the end effector with the base.Each intermediate joint is modeled as a dampened spring with the links unaffected by gravity. This also resolves the redundancy, i.e., joint angle ambiguity a 7-jointed robot has.

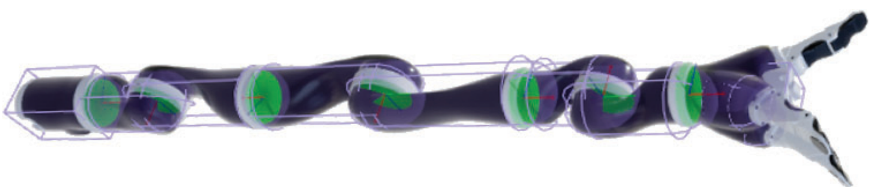


Fig. 9. Virtual Robotic Arm with Physics Constraints: purple capsules represent links, green discs represent angular constraints.

This approach allows for nearly arbitrary motion of the end effector and a semi-realistic interaction of the arm with the environment. As a beneficial side effect, developers can disconnect the end effector from the rest of the arm and allow the user to control a free-floating robot hand without any constraints. However, the internal physics engine to realistically handle collisions and interactions between the end effector and the environment is still active.

Likewise, we based the grasp concept on a custom interaction design for robotic grasping rather than physics. Physics-based grasping in a virtual environment is a challenging task [27] and would require substantial preparation and asset fine-tuning from future developers who use the framework. Instead, we defined a logic-based approach that we consider sufficiently realistic for shared control applications: an object is regarded as grasped once it has contact with two opposite fingers while closing the gripper until the fingers open again. The grasped object is rigidly attached to the end effector, keeping its relative position stable and moving alongside the end effector until released.

4.1.2 Simulated Camera System. Computer-aided robot control usually requires a camera system – or a comparable sensor – to measure context information about the current environment for the underlying software function. To provide a realistic equivalent in simulation, *AdaptiX* contains a virtual version of the commercially available *Intel Realsense D435* [28]. This camera system is commonly used in research applications [11, 66] and can deliver aligned color and depth images. The built-in color sensor generates depth data by applying a stereo-vision algorithm using grayscale image data of two built-in infrared (IR) imagers. To improve the texture information captured by the IR imagers, the camera also includes an IR projector, which projects a static pattern on the scene.

As with the simulated robotic arm, the virtual camera system is a modular actor that can be arbitrarily placed within the simulation environment. Its mesh and texture are derived directly from the manufacturer's CAD files to optimize authenticity. The virtual camera system includes all image sensors of the original, plus an optional virtual sensor generating a segmented image of the scene. We designed the virtual sensor parameters to be as close as possible to those of the actual sensors. They include – but are not limited to – sensor dimensions, lens structure, focal length, and aperture.

Because the framework can provide depth information directly from the 3D simulation, the virtual depth camera does not need to calculate its data using stereo-vision but instead yields perfect per-pixel depth information. If stereo-vision-generated depth data with realistic noise, errors, and other algorithm-specific effects is needed, the virtual system also delivers the IR images for a manual calculation.

Additionally, the simulated camera system supports the usage of the image data in-simulation and storing the data on disk for applications such as dataset generation or logging.

4.2 Adaptive DoF Mapping Control (ADMC)

The adaptive DoF mapping is implemented in the object *Axis Wizard*, which provides functions to calculate the optimal suggestion, as well as the other possible optimizations. The calculation relies solely on the virtual objects in the simulation environment instead of object recognition or camera data to enable development and evaluation without a physical robot setup. However, the camera feed for object recognition can be activated by developers to read positions and orientations. In addition to the positions and orientations of the *Gripper Mover* and the *Current Target* (which can be an object to pick up or a target surface to place the object on, depending on the context), two other parameters of *Axis Wizard* are important to ensure the correct calculations for the pick-and-place task – *Minimal Hover Distance* and *Hover Height*.

Disregarding the handling of edge cases, the calculation of the optimal suggestion is taken care of in three steps: 1) calculating *Translation*, 2) calculating *Rotation*, and 3) calculating the finger movement variable *Gripper*. The Blueprints for implementation details are provided in Appendix A.

4.2.1 Calculation of the Optimal Suggestion. *Minimal Hover Distance* represents the distance – projected on the XY-plane – between the *Gripper Mover* and the *Current Target*. When this distance is smaller than the *Minimal Hover Distance* (see Figure 12 in the appendix), the *Axis Wizard* uses a point above the *Current Target* for its calculations – referred to as the *Target Point*, instead of the *Current Target*'s position to prevent the robot from getting too close to the table, allowing for proper gripper rotation. Then, a vector from the *Gripper Mover*'s position towards the *Target Point* is calculated, normalized, and inversely rotated by the *Gripper Mover*'s rotation. This calculation returns a unit vector pointing from the *Gripper Mover* toward the *target point* in the *Gripper Mover*'s reference frame. This vector is then scaled by the *Vel Trans* value of the *Kinova Jaco 2* to get a translation of the size of the movement performed by the *Kinova Jaco 2* during one frame.

Hover Height determines the height of the aforementioned point above the *Current Target*. If the XY-projected distance between the *Gripper Mover* and the *Current Target* is smaller than the *Minimal Hover Distance*, the *Axis Wizard* directly uses the *Current Target*'s position for its calculations instead of the point above it.

To calculate the optimal suggestion's *Rotation*, the *Translation* – calculated in the first step – is used as input for the *Make Rot from X* node. This node returns a *rotator* representing the rotation required to make an object point toward the direction indicated by the input vector – *target point*. To mitigate an additional *roll* of *Gripper Mover*, the inverse value is added, keeping the *Gripper Mover's* orientation largely steady. Additionally, since only a small part of the rotation is performed during one frame, the *rotator* is scaled down. The calculation for the *Rotation*, excluding edge cases, is depicted in Figure 13 in the appendix.

4.2.2 Calculation of Gripper values. The *Gripper* value only depends on whether the target point is within reach of the robotic fingers, either with or without additional movement (i.e. if the fingers are almost close enough, there will be a movement towards the target point, otherwise the fingers will engage without moving the gripper) and whether or not an object is currently being grasped (i.e. if an object is grasped and the gripper is close to the target point, it suggests to open the fingers, otherwise close them).

4.2.3 Calculation of the Adjustment Suggestion. The adjustment suggestion is calculated by rotating the optimal suggestion's *Translation* by 90° around the Y-Axis, keeping the same *Rotation* and setting the *Gripper* value to 0. This results in a DoF mapping which moves roughly along the *Gripper Mover's* Z-Axis, or colloquially "up and down" between the fingers if the optimal suggestion is seen as "forward and backward". As *Rotation* is kept the same between the optimal and adjustment suggestions, the resulting movement keeps the fingers roughly facing the direction of the *Current Target*.

The translation, rotation, and gripper suggestions use much simpler calculations. The translation suggestion calculates a vector from the *Gripper Mover* towards the *Current Target*, inversely rotates it by the *Gripper Mover's* rotation to put it into the *Gripper Mover's* reference frame and uses that as the *Translation* value for the suggested *Adaptive Axis*. This vector is also what the rotation suggestion uses to calculate a *Rotator* representing a rotation towards the *Current Target*. The gripper suggestion checks whether an object is currently being grasped. If so, the suggestion is to open the fingers (*Gripper* = -1). Otherwise, the suggestion is to close the fingers (*Gripper* = 1).

4.2.4 Attention Guidance in Threshold*.* Both the *Continuous* and *Threshold* approaches share the same core calculation for DoF mappings. However, the *Threshold* approach has an additional task:

determining whether the optimal suggestion significantly differs from the currently active DoF mapping. This task is more related to visualization than the DoF mapping calculation itself and is managed by the *Gizmo* object.

The *Gizmo* object contains a *Realtime Threshold* variable, which represents the threshold as a value between 0 and 1. It also includes a function called *Adaptive Axes Nearly Equal*, which determines whether two *Adaptive Axes* are nearly equal by checking if their difference is below the *Realtime Threshold*. The threshold value is chosen to be between 0 and 1 to align with a percentage of difference (see Section 3.1.3), providing a more intuitive understanding of the amount of difference compared to the cosine similarity value used as the basis for the difference calculation.

As the *Unreal Engine* does not provide an arbitrarily sized vector structure, the calculations required needed to be programmed manually rather than with built-in vector operations. Therefore, two math expression nodes were defined, one calculating the dot product of two 7D vectors and the other calculating the magnitude of a 7D vector. Using these, the cosine similarity between two *Adaptive Axes* could be calculated in *Unreal Blueprints* (see Figure 14 in the appendix). To forego the transformation of the cosine similarity into a percentage difference, the *Unreal Engine's Nearly Equal* node was used to determine whether the cosine similarity was nearly equal to *1* – meaning the vectors align – with a threshold of *2 * Realtime Threshold*. The threshold needed to be multiplied by *2* as the range of the cosine similarity has a magnitude of 2. The result of this calculation is a boolean value that is true if the difference between the *Adaptive Axes* is below the threshold and false otherwise.

The resulting value is then used by the *Gizmo* to show the arrow corresponding to the optimal suggestion. It is also used to notify the *Game Mode* – an object representing the game, keeping track of study variables, etc. – that the threshold was broken. This triggers an event that causes a 1kHz sound to play and a haptic effect to occur on the motion controller. A reset variable is used to prevent the sound from constantly triggering. However, there appears to be a specific point during movement at which it is possible for users to stop their input and the software to get caught in a loop of firing the event and resetting it, causing a constant sound and vibration. If users continued their movement, the software stopped firing the event, seizing the sound and vibration. Unfortunately, this was only noticed during the experiment, which is why the problem persists in the current software version. Assuming *Threshold* is to be used in future research, a better solution for a single fire execution of the notification needs to be developed.

5 LIMITATIONS

In HRI research, the leading factor impacting user experience is usually the chosen method of (shared) control and the respective interfaces. Using frameworks like *AdaptiX* allows researchers to tweak these variables toward high user satisfaction through methodological studies and experiments.

However, like any simulation, *AdaptiX* only approximates reality and contains ingrained limitations when working with the system and evaluating generated results.

5.1 Scenario Selection

In the initial version, *AdaptiX* provides only a single level, as seen in all screenshots of this work. This scenario functions mainly as a model for simple tasks. As such, it lacks environment interactions or varying backgrounds and is not designed for a specific assistive task.

This single level might need to be revised to represent the complete application range of assistive shared control, which is why extensions are required. As such, *AdaptiX*'s modular design allows the community to generate custom levels for their specific research interests effortlessly.

5.2 Simulation Sickness caused by Head Mounted Display

HMDs are a popular tool to create immersive virtual environments, frequently used in research and industrial settings. However, using a HMD in HRI can create a significant displacement between the virtual object and the physical world through effects related to the resulting limited field of view, reduced depth perception, and distorted spatial cues.

For applications within the *AdaptiX* framework, these issues could result in users experiencing motion sickness, disorientation, discomfort, and potentially decreased performance when interacting with the simulated robotic arm or virtual objects. Researchers must consider these artifacts when designing experiments, especially when developing studies including qualitative questionnaires or when comparing different levels of MR continuum.

5.3 Simulation Environment

The simulation environment centers on the photogrammetry scan of an actual room that contains a table with an attached virtual robotic arm. Compared to a 3D modeling of a room, the photogrammetry does not provide a high resolution, leading to a partial blurred appearance.

AdaptiX does not provide a photo realistic virtual environment (yet). However, in our studies, the slightly blurred appearance never seemed to have had a negative effect. On the contrary, it has helped participants focus on the scene's relevant parts (i.e. the robot and objects). Researchers and developers are invited to create and evaluate a 3D modeled environment.

5.4 Simulated Robotic Arm

If controlled entirely in simulation, the robotic arm (as described in Section 4.1.1) does not move identically to an actual *Kinova Jaco 2* because of implementation decisions favoring physical interactions over accurate per-joint robot actions. In most other cases, the individual joints are in relatively realistic positions, even though they might not be identical to the underlying solution provided by an inverse kinematic of the real robot.

Especially in the *follow-me* approach (see Section 3.3.1), it is possible to reach outside of the mechanical range of the robotic arm. Due to the entirely physics-based connection, this results in partially disconnected joints. However, this is only an issue of visualizing the robotic arm in the simulation environment and does not affect the control or the TCP data recording.

Likewise, grasping simulated objects is based on a custom implementation, and grabbed objects are firmly attached to the end effector. Care must be taken for objects that are – in reality – too heavy for the gripper, have slippery surfaces, or have mechanical dimensions that make the object unstable when held. Theoretically, this "ideal kind of grasping" allows the virtual robot to move any arbitrarily large and heavy object. To address this, we added the object tag *Graspable* that allows developers to define permitted – and by omission – unpermitted objects.

5.5 Simulated Camera System

Although the simulated camera is based on manufacturer CAD files, comparison tests failed to deliver completely identical data to the actual recording system. These variances stem from environmental differences between simulation and reality, as light or dust/other particles in the air will cause effects in the produced image. However, these effects can be added in post-production or – if required – activated in the framework. By default, the respective settings are disabled as they would primarily introduce noise that not every developer might want.

On a technical level, the images generated by the virtual system differ slightly in terms of data types. The virtual grayscale IR images consist of three identical color channels instead of a single channel in reality. Also, the virtual IR and color images include an additional fourth alpha channel, which is not used in our framework. The generated depth data format also differs, as the actual camera system generates images as 16-bit unsigned integer, and the simulation provides them as 16-bit signed floats. The depth data generated by the framework is pixel-perfect, which ignores various camera system effects that occur in reality by the calculation of depth using stereo-vision.

All these technical differences are addressed within the framework through data transformation and should not noticeably affect the output of *AdaptiX*. However, researchers and developers should be aware of these adjustments for future developments and extension.

5.6 ROS Interface

The ROS interface connects the virtual with a real robot, each with its own environmentally-determined set of limitations. This results in some logical inconsistencies while using the interface. The obvious velocity limitations of the real system result in delayed execution if reality is to follow the simulation. Therefore, the maximum velocity of the virtual robotic arm is set automatically to the physical characteristics after enabling ROS. Also, as the virtual joints are not controlled by an inverse kinematics (IK) but instead based on physics, the interface sends only end effector poses to the real robot, omitting individual joint poses. This may result in differing robot configurations, with only the end effector point being aligned in some instances.

When sending pose data from the real robot to the virtual twin in simulation, most of these restrictions do not apply. The simulated robot can move arbitrarily fast, and its configuration aligns automatically with the real system. The only restriction is that, by default, no further information about the natural environment is available, resulting in a relatively empty virtual environment if relying purely on the ROS interface.

When designing expansions, developers also must be aware that ROS and *Unreal Engine* differ in handedness. ROS is based on a right-handed coordinate system, while the *Unreal Engine* uses a left-handed approach. *AdaptiX* internally does the necessary transformation for the robotic arm but will not automatically calculate this for other position and orientation data, e.g., obstacles. However, researchers can mitigate this by applying the provided coordinate transformation methods of the robotic arm to any further object.

6 FRAMEWORK EXAMPLE ADAPTIONS

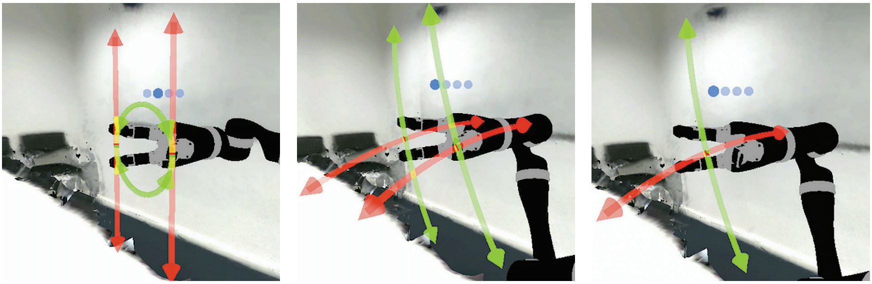
The *AdaptiX* framework has been successfully used and adapted in three case studies evaluating interaction concepts and multi-modal feedback with remote and laboratory-based focus groups.

6.1 Example Adaption 1: Adaptive Control of an Assistive Robot

In an initial study [32], the *AdaptiX* framework was used to explore the proposed ADMC control method with associated visual cues for various DoF mappings.

In particular, we analyzed how the novel adaptive control method – proposed by Goldau and Frese [19] – performs in a 3D environment compared to the standard mode-switch approach with cardinal DoF mappings. They also investigated whether changes in the visual cues' appearance impact the performance of the adaptive control method. Three different types of control with varying visual cues and methods of mapping DoFs were compared in a remote online study. These included the *Classic* visualization, one based on *Double Arrow* using two arrows attached to the gripper's fingers, and a visually reduced variant *Single Arrow*, using only one arrow through the middle of the gripper. See Figure 10 for a graphical comparison.

Due to the ongoing COVID-19 pandemic, the study was conducted entirely in a VR environment created by *AdaptiX*. Non-specific participants were recruited that had access to the required hardware (an *Oculus Quest* VR-HMD) for an immersive experience.



(a) Classic (b) Double Arrow (c) Single Arrow

Fig. 10. Evaluated interaction design and visualizations [32].

The participants repeatedly performed a simple pick-and-place task by controlling the virtual *Kinova Jaco 2* using one of the three control types. Comparative results established that adaptive controls require significantly fewer mode switches than the classic control methods. However, task completion time and workload did not improve. Study participants also mentioned concerns about the dynamically changing mapping of combined DoFs and the 2-DoF input device.

Framework contribution: *AdaptiX* demonstrated its effectiveness in this remote study to evaluate new interaction designs and feedback techniques. The innovative advantage is that the physical robotic device does not need to be present during these preliminary studies when testing and evaluating essential design elements. The *Record & Replay* functionality of *AdaptiX* allowed a remote analysis of participants data. This VR approach significantly increases the potential to include end-users in the research and design process while at the same time decreasing cost, time involvement, and accessibility concerns.

6.2 Example Adaption 2: Communicating Adaptive Control Recommendations

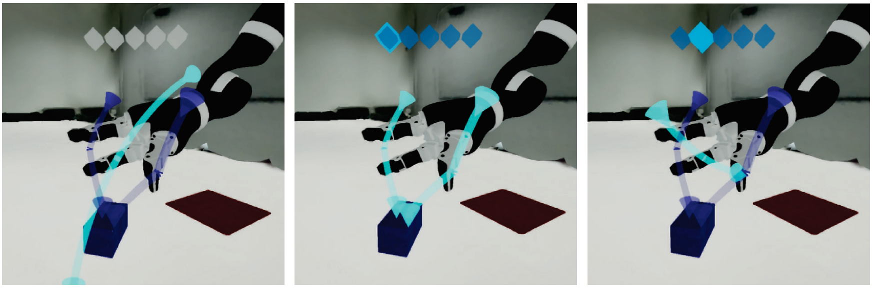
A follow-up study [46] evaluated two new adaptive control methods for an assistive robotic arm, one of which involves a multi-modal approach for attention guiding of the user.

We used *AdaptiX* in a laboratory study to cross-validate the initial study's findings on how participants interact with the environment. The adaptive system re-calculated the best combination of DoFs to complete the task during movement. These calculations were presented to the user as alternative control options for the current task. Users cycled through these suggestions – by pressing a button on the input device – to make a suitable selection or continue moving with the previous active DoFs (see Figure 11).

They contrasted the variants *Continuous* and *Threshold*, differing in the time at which suggestions are communicated to the user, against a non-adaptive *Classic* control method. Possible effects on task completion time, the number of necessary mode switches, perceived workload, and user opinions on each control method were compared. Further, we establish that *Continuous* and *Threshold* performed equally well in quantitative and qualitative insights. Consequently, both are promising approaches to communicating proposed directional cues effectively.

Framework contribution: The integrated multi-modal feedback is an integral feature of *AdaptiX*, capable of supporting the system's real-time suggestions by user attention guiding. Although some participants experienced the combined visual-auditory-haptic multi-modal feedback as "irritating" [46], it effectively communicated updated suggestions. One application of virtual frameworks like *AdaptiX* might be the differentiation between different modality types and corresponding user preferences in an easy-to-set-up study. Highlighting the advantage of our framework, we could evaluate our different visualizations and multi-modal feedback without implementing a VR environment [46].

Based on the successful implementation of *AdaptiX* in this laboratory study, we are confident that the framework performs well in remote and in-person studies.



(a) (b) (c)

Fig. 11. Suggested control alternatives in light blue, visualized as in case study 2: (**a**) Moving forward and downward towards the object, (**b**) Closing the fingers to grasp the object, and (**c**) Moving towards the target area.

6.3 Example Adaption 3: Comparing Input Devices for Controlling a Physical Robot in Mixed Reality

A third study [47] highlights the MR capability of the framework and the integration options with different input devices. This study used the *Varjo XR-3* XR-HMD to explore a similar interaction design and feedback technique to our *Threshold* approach [46]. By incorporating this XR-HMD, the prototype mimics an AR environment (see Section 3.2) to the user, seeing the physical setup augmented by visual cues. Instead of a virtual pick-and-place task as before, this study combined a physical object, a physical drop area, and a physical robotic arm with AR cues delivered via the headset.

Participants compared three assistive input techniques: 1) a head-based control by using the deflection of the head on the *pitch* axis for continuous input and on the *roll* axis for mode-switching, 2) a gamepad input by using the *Xbox Adaptive Controller* [40] extended with *Logitech Adaptive Gaming Kit* [35] buttons for a discrete input, and 3) the control-stick of a *Nintendo Joy-Con* [43] motion controller – as a baseline to our previous study [46].

Framework contribution: With its real-world setting augmented by virtual cues, the research moved closer to reality on the MR-continuum than the previous two case studies. *AdaptiX* successfully performed as an easy-to-use interface between the usage of a physical robot and virtual communication via a XR-HMD.

It also allowed the research team to quickly evaluate the efficiency of different input devices with the potential to control the robotic arm along the adaptive DoF mapping. The standardized *User Input Adapter* enables researchers to easily chose between different technologies – supporting continuous, discrete, and absolute user input – and further extend it to their needs by its modular nature.

7 CONCLUSION

Integrating *AdaptiX* into HRI research can streamline the development and evaluation of new interaction designs and feedback techniques for controlling assistive robotic arms. The framework is advantageous in remote and in-person studies as its usage negates the need for a physical robotic device during the initial ideation and prototyping stages, thus increasing flexibility, accessibility, and efficiency.

An initial shared control concept by adaptive DoF mapping is provided and implemented to support researchers and developers to either change, extend, or exchange methods with their ideas. In studies using a physical robot, the integration of ROS bridges the gap to reality, by enabling a bidirectional connection between virtual and physical robotic arm. ROS allows developers and users to choose between a *DigitalTwin* and *PhysicalTwin* approach while interacting with *AdaptiX*. Using *AdaptiX*, researchers benefit from the entire continuum of MR. As the simulated and real-world environments of the robotic arm are perfectly aligned, nearly seamless switching between controlling the real and virtual robot is possible. This functionality allows applications in pure screen space, VR, AR, simultaneous simulation/reality, and pure reality. *AdaptiX*'s 3D teach-in interface facilitates a code-less trajectory programming of an assistive robot by hand-guiding the simulated or real robot to the specific location and saving the position and orientation of the tool center point. These waypoints are interpolated to a combined movement trajectory. The framework's recording/replaying system is entirely customizable. It includes options to change details during replay, such as repositioning cameras or re-rendering background scenes. A fully integrated recording of participants interacting with the robot is possible, which can be analyzed afterward to evaluate the specific research variables.

Taken together, *AdaptiX* is a free and open-source tool that enables HRI researchers to test and evaluate their shared control concepts for assistive robotic devices in a high-resolution virtual environment. The cited case studies clearly demonstrate the benefits researchers and developers can draw from using the framework. The near-endless customization options allow users to tweak the initial version to their specific research needs, resulting in practically tailor-made environments.

7.1 Framework Extensions

We invite the community to extend the *AdaptiX* framework based on their requirements needs by creating custom levels/scenarios and integrating new interfaces. *AdaptiX* can be accessed free-of-charge at https://adaptix.robot-research.de. Refer to the *README* provided in the repository for a detailed description of how to implement experiments in *AdaptiX*.

ACKNOWLEDGMENTS

This research is supported by the *German Federal Ministry of Education and Research* (BMBF, FKZ: 16SV8563, 16SV8565).

REFERENCES

[1] David A. Abbink, Tom Carlson, Mark Mulder, Joost C. F. de Winter, Farzad Aminravan, Tricia L. Gibo, and Erwin R. Boer. 2018. A Topology of Shared Control Systems—Finding Common Ground in Diversity. *IEEE Transactions on Human-Machine Systems* 48, 5 (Oct. 2018), 509–525. https://doi.org/10.1109/thms.2018.2791570

[2] Stephanie Arévalo-Arboleda, Max Pascher, Annalies Baumeister, Barbara Klein, and Jens Gerken. 2021. Reflecting upon Participatory Design in Human-Robot Collaboration for People with Motor Disabilities: Challenges and Lessons Learned from Three Multiyear Projects. In *The 14th PErvasive Technologies Related to Assistive Environments Conference - PETRA 2021* (2021-06-29). ACM. https://doi.org/10.1145/3453892.3458044

[3] Maude Beaudoin, Josiane Lettre, François Routhier, Philippe S. Archambault, Martin Lemay, and Isabelle Gélinas. 2018. Impacts of robotic arm use on individuals with upper extremity disabilities: A scoping review. *Canadian Journal of Occupational Therapy* 85, 5 (2018), 397–407. https://doi.org/10.1177/0008417418820878 PMID: 30866682.

[4] Connor Brooks and Daniel Szafir. 2020. Visualization of Intended Assistance for Acceptance of Shared Control. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE Press, Piscataway, NJ, USA, 11425–11430. https://doi.org/10.1109/IROS45743.2020.9340964

[5] Jennifer L Burke, Matthew S Prewett, Ashley A Gray, Liuquin Yang, Frederick RB Stilson, Michael D Coovert, Linda R Elliot, and Elizabeth Redden. 2006. Comparing the effects of visual-auditory and visual-tactile feedback on user performance: a meta-analysis. In *Proceedings of the 8th international conference on Multimodal interfaces*. Association for Computing Machinery, New York, NY, USA, 108–117.

[6] Samuel Bustamante, Jan Peters, Bernhard Schölkopf, Moritz Grosse-Wentrup, and Vinay Jayaram. 2021. ArmSym: A Virtual Human–Robot Interaction Laboratory for Assistive Robotics. *IEEE Transactions on Human-Machine Systems* 51, 6 (2021), 568–577. https://doi.org/10.1109/THMS.2021.3106865

[7] Gerard Canal, Guillem Alenyà, and Carme Torras. 2016. Personalization Framework for Adaptive Robotic Feeding Assistance. In *Social Robotics*. Vol. 9979. Springer International Publishing, Cham, 22–31. https://doi.org/10.1007/978-3-319-47437-3_3

[8] Ravi Teja Chadalavada, Henrik Andreasson, Maike Schindler, Rainer Palm, and Achim J Lilienthal. 2020. Bi-directional navigation intent communication using spatial augmented reality and eye-tracking glasses for improved safety in human–robot interaction. *Robotics and Computer-Integrated Manufacturing* 61 (2020), 101830.

[9] Yuhang Che, Allison M. Okamura, and Dorsa Sadigh. 2020. Efficient and Trustworthy Social Navigation via Explicit and Implicit Robot–Human Communication. *IEEE Transactions on Robotics* 36, 3 (2020), 692–707. https://doi.org/10.1109/TRO.2020.2964824

[10] Tiffany L. Chen, Chih-Hung King, Andrea L. Thomaz, and Charles C. Kemp. 2011. Touched by a robot. In *Proceedings of the 6th international conference on Human-robot interaction - HRI '11*, Aude Billard, Peter Kahn, Julie A. Adams, and Greg Trafton (Eds.). ACM Press, New York, New York, USA, 457. https://doi.org/10.1145/1957656.1957818

[11] Marianna Ciccarelli, Alessandra Papetti, Cecilia Scoccia, Giacomo Menchi, Leonardo Mostarda, Giacomo Palmieri, and Michele Germani. 2022. A system to improve the physical ergonomics in Human-Robot Collaboration. *Procedia Computer Science* 200 (2022), 689–698. https://doi.org/10.1016/j.procs.2022.01.267

[12] Yann-Seing Law-Kam Cio, Maxime Raison, Cédric Leblond Ménard, and Sofiane Achiche. 2019. Proof of Concept of an Assistive Robotic Arm Control Using Artificial Stereovision and Eye-Tracking. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 27, 12 (2019), 2344–2352. https://doi.org/10.1109/TNSRE.2019.2950619

[13] Andre Cleaver, Darren Vincent Tang, Victoria Chen, Elaine Schaertl Short, and Jivko Sinapov. 2021. Dynamic Path Visualization for Human-Robot Collaboration. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction* (Boulder, CO, USA) *(HRI '21 Companion)*. Association for Computing Machinery, New York, NY, USA, 339–343. https://doi.org/10.1145/3434074.3447188

[14] Adele Diederich and Hans Colonius. 2004. Bimodal and trimodal multisensory enhancement: effects of stimulus onset and intensity on reaction time. *Perception & psychophysics* 66, 8 (2004), 1388–1404.

[15] Epic Games. 2024. Unreal Engine. Published online (last visited on May 20, 2024). https://www.unrealengine.com Version 4.27.2.

[16] Chinemelu Ezeh, Pete Trautman, Louise Devigne, Valentin Bureau, Marie Babel, and Tom Carlson. 2017. Probabilistic vs linear blending approaches to shared control for wheelchair driving. In *2017 International Conference on Rehabilitation Robotics (ICORR)* (London, United Kingdom). IEEE Press, Piscataway, NJ, USA, 835–840. https://doi.org/10.1109/ICORR.2017.8009352

[17] F. Flemisch, D. A. Abbink, M. Itoh, M.-P. Pacaux-Lemoine, and G. Weßel. 2019. Joining the blunt and the pointy end of the spear: towards a common framework of joint action, human–machine cooperation, cooperative guidance and control, shared, traded and supervisory control. *Cognition, Technology & Work* 21, 4 (Aug. 2019), 555–568. https://doi.org/10.1007/s10111-019-00576-1

[18] Michael J Gielniak and Andrea L Thomaz. 2011. Generating anticipation in robot motion. In *2011 RO-MAN*. IEEE Press, Piscataway, NJ, USA, 449–454.

[19] Felix Ferdinand Goldau and Udo Frese. 2021. Learning to Map Degrees of Freedom for Assistive User Control: Towards an Adaptive DoF-Mapping Control for Assistive Robots. In *The 14th PErvasive Technologies Related to Assistive Environments Conference* (Corfu, Greece) *(PETRA 2021)*. Association for Computing Machinery, New York, NY, USA, 132–139. https://doi.org/10.1145/3453892.3453895

[20] Muhammad Abdul Haseeb, Maria Kyrarini, Shuo Jiang, Danijela Ristic-Durrant, and Axel Gräser. 2018. Head Gesture-Based Control for Assistive Robots. In *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference* (Corfu, Greece) *(PETRA '18)*. Association for Computing Machinery, New York, NY, USA, 379–383. https://doi.org/10.1145/3197768.3201574

[21] Laura V. Herlant, Rachel M. Holladay, and Siddhartha S. Srinivasa. 2016. Assistive Teleoperation of Robot Arms via Automatic Time-Optimal Mode Switching. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction* (Christchurch, New Zealand) *(HRI '16)*. IEEE Press, Piscataway, NJ, USA, 35–42.

[22] Nicholas J. Hetherington, Elizabeth A. Croft, and H.F. Machiel Van der Loos. 2021. Hey Robot, Which Way Are You Going? Nonverbal Motion Legibility Cues for Human-Robot Spatial Interaction. *IEEE Robotics and Automation Letters* 6, 3 (jul 2021), 5010–5015. https://doi.org/10.1109/lra.2021.3068708

[23] Catherine Holloway. 2019. Disability Interaction (DIX): A Manifesto. *Interactions* 26, 2 (feb 2019), 44–49. https://doi.org/10.1145/3310322

[24] Patrick Holthaus, Trenton Schulz, Gabriella Lakatos, and Rebekka Soma. 2023. Communicative Robot Signals: Presenting a New Typology for Human-Robot Interaction. In *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction* (Stockholm, Sweden) *(HRI '23)*. Association for Computing Machinery, New York, NY, USA, 132–141. https://doi.org/10.1145/3568162.3578631

[25] HTC. 2024. VIVE Controller 2.0. Published online (last visited on May 20, 2024). https://www.vive.com/de/accessory/controller2018/

[26] HTC. 2024. VIVE Tracker 3.0. Published online (last visited on May 20, 2024). https://www.vive.com/de/accessory/tracker3/

[27] Markus Höll, Markus Oberweger, Clemens Arth, and Vincent Lepetit. 2018. Efficient Physics-Based Implementation for Realistic Hand-Object Interaction in Virtual Reality. In *2018 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE Press, Piscataway, NJ, USA, 175–182. https://doi.org/10.1109/VR.2018.8448284

[28] Intel Corporation. 2022. Intel Realsense - Product Family D400 Series. Published online (last visited on May 20, 2024). https://www.intelrealsense.com/wp-content/uploads/2022/11/Intel-RealSense-D400-Series-Datasheet-November-2022.pdf Revision 014.

[29] Siddarth Jain, Ali Farshchiansadegh, Alexander Broad, Farnaz Abdollahi, Ferdinando Mussa-Ivaldi, and Brenna Argall. 2015. Assistive robotic manipulation through shared autonomy and a Body-Machine Interface. In *2015 IEEE International Conference on Rehabilitation Robotics (ICORR)*. IEEE Press, Piscataway, NJ, USA, 526–531. https://doi.org/10.1109/ICORR.2015.7281253

[30] Kinova inc. 2021. Jaco assistive robot - User guide. Published online (last visited on May 20, 2024). https://assistive.kinovarobotics.com/uploads/EN-UG-007-Jaco-user-guide-R05.pdf EN-UG-007-r05-202111.

[31] Robert J Kosinski. 2008. A literature review on reaction time. *Clemson University* 10, 1 (2008), 337–344.

[32] Kirill Kronhardt, Stephan Rübner, Max Pascher, Felix Goldau, Udo Frese, and Jens Gerken. 2022. Adapt or Perish? Exploring the Effectiveness of Adaptive DoF Control Interaction Methods for Assistive Robot Arms. *Technologies* 10, 1 (2022), 24 pages. https://doi.org/10.3390/technologies10010030

[33] Maria Kyrarini, Fotios Lygerakis, Akilesh Rajavenkatanarayanan, Christos Sevastopoulos, Harish Ram Nambiappan, Kodur Krishna Chaitanya, Ashwin Ramesh Babu, Joanne Mathew, and Fillia Makedon. 2021. A survey of robots in healthcare. *Technologies* 9, 1 (2021), 8.

[34] Audrey Lebrasseur, Josiane Lettre, François Routhier, Philippe S. Archambault, and Alexandre Campeau-Lecours. 2019. Assistive robotic arm: Evaluation of the performance of intelligent algorithms. *Assistive Technology* 33, 2 (May 2019), 95–104. https://doi.org/10.1080/10400435.2019.1601649

[35] Logitech. 2024. Adaptive Gaming Kit. Published online (last visited on May 20, 2024). https://www.logitechg.com/en-us/products/gamepads/adaptive-gaming-kit-accessories.943-000318.html

[36] Zhanat Makhataeva and Huseyin Atakan Varol. 2020. Augmented Reality for Robotics: A Review. *Robotics* 9, 2 (2020). https://doi.org/10.3390/robotics9020021

[37] Elias Matsas and George-Christopher Vosniakos. 2015. Design of a virtual reality training system for human–robot collaboration in manufacturing tasks. *International Journal on Interactive Design and Manufacturing (IJIDeM)* 11, 2 (Feb. 2015), 139–153. https://doi.org/10.1007/s12008-015-0259-2

[38] Alyxander David May, Christian Dondrup, and Marc Hanheide. 2015. Show me your moves! Conveying navigation intention of a mobile robot to humans. In *2015 European Conference on Mobile Robots (ECMR)*. 1–6. https://doi.org/10.1109/ECMR.2015.7324049

[39] Meta. 2024. Meta Quest 2. Published online (last visited on May 20, 2024). https://www.meta.com/de/quest/products/quest-2/

[40] Microsoft. 2024. Xbox Adaptive Controller. Published online (last visited on May 20, 2024). https://www.microsoft.com/en-us/d/xbox-adaptive-controller/8nsdbhz1n3d8

[41] Paul Milgram and Fumio Kishino. 1994. A taxonomy of mixed reality visual displays. *IEICE TRANSACTIONS on Information and Systems* 77, 12 (1994), 1321–1329.

[42] Sarah Luisa Müller, Stefan Schröder, Sabina Jeschke, and Anja Richert. 2017. Design of a Robotic Workmate. In *Lecture Notes in Computer Science*. Springer International Publishing, 447–456. https://doi.org/10.1007/978-3-319-58463-8_37

[43] Nintendo. 2024. Nintendo Joy-Con. Published online (last visited on May 20, 2024). https://www.nintendo.de/Hardware/Zubehor/Zubehor-1183144.html

[44] Max Pascher, Annalies Baumeister, Stefan Schneegass, Barbara Klein, and Jens Gerken. 2021. Recommendations for the Development of a Robotic Drinking and Eating Aid - An Ethnographic Study. In *Human-Computer Interaction*

– *INTERACT 2021* (2021-09-01), Carmelo Ardito, Rosa Lanzilotti, Alessio Malizia, Helen Petrie, Antonio Piccinno, Giuseppe Desolda, and Kori Inkpen (Eds.). Springer, Cham. https://doi.org/10.1007/978-3-030-85623-6_21

[45] Max Pascher, Uwe Gruenefeld, Stefan Schneegass, and Jens Gerken. 2023. How to Communicate Robot Motion Intent: A Scoping Review. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems - CHI '23*. https://doi.org/10.1145/3544548.3580857

[46] Max Pascher, Kirill Kronhardt, Felix Ferdinand Goldau, Udo Frese, and Jens Gerken. 2023. In Time and Space: Towards Usable Adaptive Control for Assistive Robotic Arms. In *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2300–2307. https://doi.org/10.1109/RO-MAN57019.2023.10309381

[47] Max Pascher, Kevin Zinta, and Jens Gerken. 2024. Exploring of Discrete and Continuous Input Control for AI-enhanced Assistive Robotic Arms. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction* (Boulder, CO, USA) *(HRI '24 Companion)*. Association for Computing Machinery, New York, NY, USA, 5 pages. https://doi.org/10.1145/3610978.3640626

[48] Sebastiaan Petermeijer, Pavlo Bazilinskyy, Klaus Bengler, and Joost de Winter. 2017. Take-over again: Investigating multimodal and directional TORs to get the driver back into the loop. *Applied Ergonomics* 62 (2017), 204–215. https://doi.org/10.1016/j.apergo.2017.02.023

[49] Sebastiaan Petermeijer, Fabian Doubek, and Joost de Winter. 2017. Driver response times to auditory, visual, and tactile take-over requests: A simulator study with 101 participants. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. 1505–1510. https://doi.org/10.1109/SMC.2017.8122827

[50] Laura Petrich, Jun Jin, Masood Dehghan, and Martin Jagersand. 2022. A Quantitative Analysis of Activities of Daily Living: Insights into Improving Functional Independence with Assistive Robotics. In *2022 International Conference on Robotics and Automation (ICRA)*. 6999–7006. https://doi.org/10.1109/ICRA46639.2022.9811960

[51] Anita Pollak, Mateusz Paliga, Matias M. Pulopulos, Barbara Kozusznik, and Malgorzata W. Kozusznik. 2020. Stress in manual and autonomous modes of collaboration with a cobot. *Computers in Human Behavior* 112 (2020), 106469. https://doi.org/10.1016/j.chb.2020.106469

[52] Gabriel Quere, Annette Hagengruber, Maged Iskandar, Samuel Bustamante, Daniel Leidner, Freek Stulp, and Jörn Vogel. 2020. Shared Control Templates for Assistive Robotics. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE Press, Piscataway, NJ, USA, 1956–1962. https://doi.org/10.1109/ICRA40945.2020.9197041

[53] Emanuele Ruffaldi, Filippo Brizzi, Franco Tecchia, and Sandro Bacinelli. 2016. Third point of view augmented reality for robot intentions visualization. In *International Conference on Augmented Reality, Virtual Reality and Computer Graphics*. Springer, 471–478.

[54] Ivan Shindev, Yu Sun, Michael Coovert, Jenny Pavlova, and Tiffany Lee. 2012. Exploration of Intention Expression for Robots. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction* (Boston, Massachusetts, USA) *(HRI '12)*. Association for Computing Machinery, New York, NY, USA, 247–248. https://doi.org/10.1145/2157689.2157778

[55] Moondeep C. Shrestha, Ayano Kobayashi, Tomoya Onishi, Erika Uno, Hayato Yanagawa, Yuta Yokoyama, Mitsuhiro Kamezaki, Alexander Schmitz, and Shigeki Sugano. 2016. Intent communication in navigation through the use of light and screen indicators. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. 523–524. https://doi.org/10.1109/HRI.2016.7451837

[56] Joris Sijs, Freek Liefhebber, and Gert Willem R.B.E. Romer. 2007. Combined Position & Force Control for a robotic manipulator. In *2007 IEEE 10th International Conference on Rehabilitation Robotics*. IEEE. https://doi.org/10.1109/icorr.2007.4428414

[57] Amit Singhal. 2001. Modern information retrieval: A brief overview. *IEEE Data Eng. Bull.* 24, 4 (2001), 35–43.

[58] Ryo Suzuki, Adnan Karim, Tian Xia, Hooman Hedayati, and Nicolai Marquardt. 2022. Augmented Reality and Robotics: A Survey and Taxonomy for AR-Enhanced Human-Robot Interaction and Robotic Interfaces. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 553, 33 pages. https://doi.org/10.1145/3491102.3517719

[59] Emmanuele Tidoni, Mohammad Abu-Alqumsan, Daniele Leonardis, Christoph Kapeller, Gabriele Fusco, Cristoph Guger, Cristoph Hintermuller, Angelika Peer, Antonio Frisoli, Franco Tecchia, Massimo Bergamasco, and Salvatore Maria Aglioti. 2017. Local and Remote Cooperation With Virtual and Robotic Agents: A P300 BCI Study in Healthy and People Living With Spinal Cord Injury. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 25, 9 (Sept. 2017), 1622–1632. https://doi.org/10.1109/tnsre.2016.2626391

[60] Katherine M Tsui, Dae-Jin Kim, Aman Behal, David Kontak, and Holly A Yanco. 2011. "I want that": Human-in-the-loop control of a wheelchair-mounted robotic arm. *Applied Bionics and Biomechanics* 8, 1 (2011), 127–147.

[61] Bram van Deurzen, Herman Bruyninckx, and Kris Luyten. 2022. Choreobot: A Reference Framework and Online Visual Dashboard for Supporting the Design of Intelligible Robotic Systems. *Proc. ACM Hum.-Comput. Interact.* 6, EICS, Article 151 (jun 2022), 24 pages. https://doi.org/10.1145/3532201

[62] Varjo. 2024. Varjo XR-3. Published online (last visited on May 20, 2024). https://varjo.com/products/xr-3/

[63] Michael Walker, Hooman Hedayati, Jennifer Lee, and Daniel Szafir. 2018. Communicating Robot Motion Intent with Augmented Reality. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction* (Chicago, IL, USA) *(HRI '18)*. Association for Computing Machinery, New York, NY, USA, 316–324. https://doi.org/10.1145/3171221.3171253

[64] Hanna Yun and Ji Hyun Yang. 2020. Multimodal warning design for take-over request in conditionally automated driving. *European transport research review* 12, 1 (2020), 1–11.

[65] Mohammad Kassem Zein, Majd Al Aawar, Daniel Asmar, and Imad H Elhajj. 2021. Deep learning and mixed reality to autocomplete teleoperation. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 4523–4529.

[66] He Zhang and Cang Ye. 2019. Human-Robot Interaction for Assisted Wayfinding of a Robotic Navigation Aid for the Blind. In *2019 12th International Conference on Human System Interaction (HSI)*. IEEE. https://doi.org/10.1109/hsi47298.2019.8942612

[67] Jakub Złotowski, Kumar Yogeeswaran, and Christoph Bartneck. 2017. Can we control it? Autonomous robots threaten human identity, uniqueness, safety, and resources. *International Journal of Human-Computer Studies* 100 (2017), 48–54.

A BLUEPRINTS OF ADMC IMPLEMENTATION

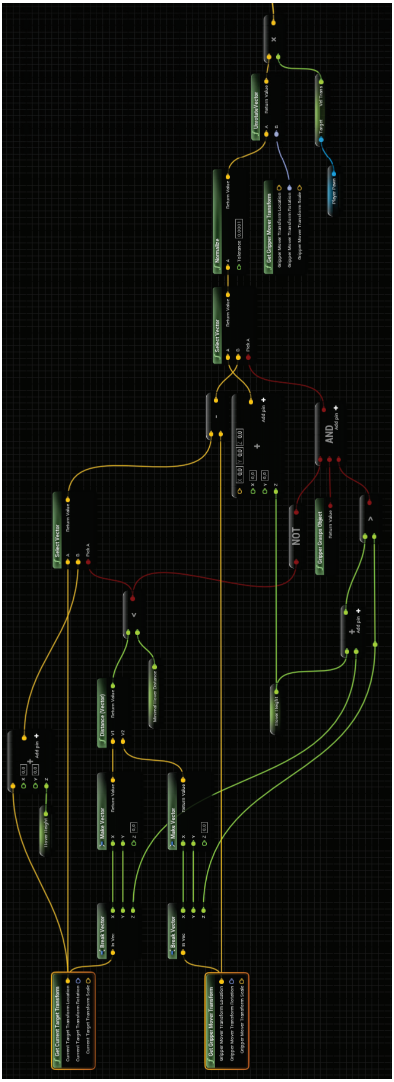


Fig. 12. Calculation of the translation for the *Optimal Suggestion*: Excerpt of *Blueprint* code calculating the *Translation* value of the *Adaptive Axis* for the *Optimal Suggestion*. Not pictured: Edge case handling for gripping an object.

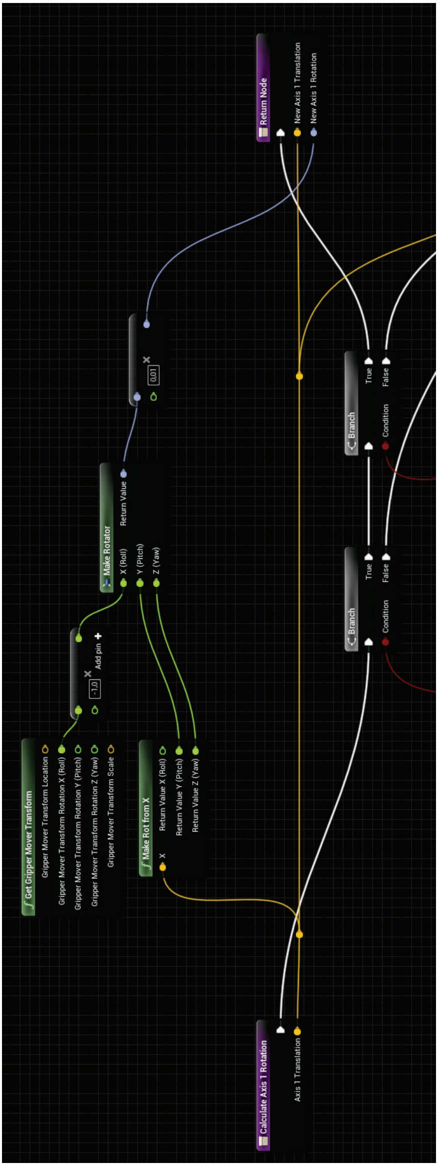


Fig. 13. Calculation of the Rotation for the *Optimal Suggestion*: Excerpt of *Blueprint* code calculating the *Rotation* value of the *Adaptive Axis* for the *Optimal Suggestion*. Not pictured: Edge case handling.

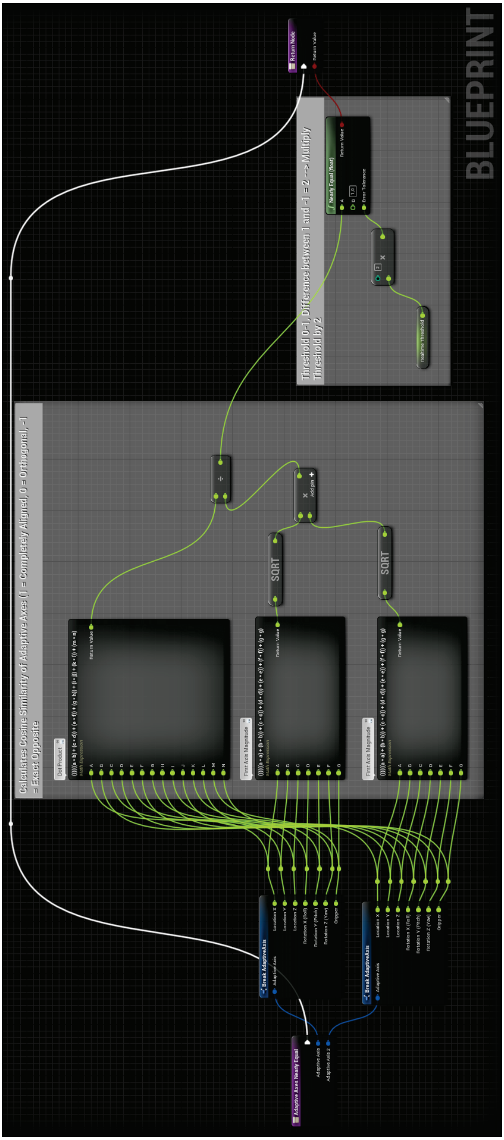


Fig. 14. *Adaptive Axes Nearly Equal* function to prepare the multi-modal attention guiding of the user.

First Submitted February 2023; Revised July 2023; Accepted October 2023

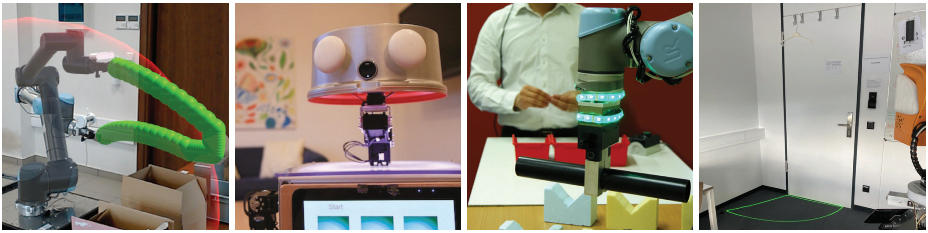
How to Communicate Robot Motion Intent: A Scoping Review

Max Pascher
max.pascher@w-hs.de
Westphalian University of Applied Sciences
Gelsenkirchen, Germany
University of Duisburg-Essen
Essen, Germany

Uwe Gruenefeld
uwe.gruenefeld@uni-due.de
University of Duisburg-Essen
Essen, Germany

Stefan Schneegass
stefan.schneegass@uni-due.de
University of Duisburg-Essen
Essen, Germany

Jens Gerken
jens.gerken@w-hs.de
Westphalian University of Applied Sciences
Gelsenkirchen, Germany



(a) *Motion* [120] (b) *Attention* [67] (c) *State* [117] (d) *Instruction* [128]

Figure 1: (a) *Robot motion intent*: The robot communicates its intended motion (e.g., a trajectory of the robot's intended movement path is visualized in Augmented Reality [120]). Furthermore, our analysis revealed three additional types of intent that complement robot motion intents. (b) *Attention*: A robot aims to catch the user's attention for subsequent movement activity (e.g., by moving its whole body [67]). (c) *State*: A robot communicates its state so that a human can predict future motions and identify potential conflicts before they occur (e.g., the robot communicates its movement activity with the help of a colored LED stripe [117]). (d) *Instruction*: The robot aims to provide specific instructions so that the human can assist further movement (e.g., by requesting to open a door [128]).

ABSTRACT

Robots are becoming increasingly omnipresent in our daily lives, supporting us and carrying out autonomous tasks. In Human-Robot Interaction, human actors benefit from understanding the robot's motion intent to avoid task failures and foster collaboration. Finding effective ways to communicate this intent to users has recently received increased research interest. However, no common language has been established to systematize *robot motion intent*. This work presents a scoping review aimed at unifying existing knowledge. Based on our analysis, we present an intent communication model that depicts the relationship between robot and human through different intent dimensions (*intent type*, *intent information*, *intent location*). We discuss these different intent dimensions and their interrelationships with different kinds of robots and human roles. Throughout our analysis, we classify the existing research literature along our intent communication model, allowing us to identify key patterns and possible directions for future research.

CHI '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9421-5/23/04.
https://doi.org/10.1145/3544548.3580857

CCS CONCEPTS

• **General and reference** → **Surveys and overviews**; • **Human-centered computing**; • **Computer systems organization** → *Robotics*;

KEYWORDS

intent, motion, robot, cobot, drone, survey

ACM Reference Format:
Max Pascher, Uwe Gruenefeld, Stefan Schneegass, and Jens Gerken. 2023. How to Communicate Robot Motion Intent: A Scoping Review. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 17 pages. https://doi.org/10.1145/3544548.3580857

1 INTRODUCTION

The field of Human-Computer Interaction (HCI) has moved beyond traditional user interfaces and interaction technologies. The omnipresence of Artificial Intelligence (AI) research and development requires our field to scrutinize the applicability of established design practices [2, 106]. Human interaction with AI is evolving away from being like operating a tool to being more like interacting with a partner, which is particularly interesting concerning Human-Robot Interaction (HRI) [53]. The area of HRI has been studied for a long time in HCI and, in particular, the CHI community [4, 65, 75, 90, 122]. For example, Arevalo Arboleda et al. [4] and Villanueva et al. [122] investigated combining robots and Augmented Reality (AR) technology to enable intuitive teleoperation, while others have explored on-site control of robot swarms [65] and home robots [75] as well as communication of emotions and intentions to the human [90].

Robots are versatile, they can assist us in our workplaces, support us at home, and accompany us in public spaces [1, 9, 76]. The applications of robots are manifold, significantly increasing human capabilities and efficiency [46]. While robots come in many forms, robotic arms in particular have been shown to be suitable for and adaptable to different use cases, such as production lines [15] and domestic care [96]. Here, they are known as cobots who support their users in Activities of Daily Living (ADLs), such as eating and drinking, grooming, or activities associated with leisure time.

As robots have a physical form, they tend to move and operate in the same space as humans. With advances in the degree of autonomy allowing for effective close-contact interaction, there is a need for a shared understanding between humans and robots. While robotic research tackles this from a sensory and path planning perspective (e.g., human-aware navigation [69]), the field of HCI (and HRI in particular) has been concerned with how humans may better understand robot behavior [12, 99, 124]. The subtleties of human communication are usually lost in this context, and robotic behavior needs to be understood from its own frame of reference. Robots are not a monolithic entity; with the many different types come just as many unique ways of conveying information, which could lead to erroneous interpretations by their human counterpart. An added complication is the increasing number of close-contact situations that allow little time to recognize and correct errors. This has led to numerous research efforts in recent years to find ways for robots to effectively communicate their intentions to their users [68]. This includes the direct communication of planned movements in space [54], but also less obvious means, such as drawing a user's attention to the robot [67], communicating the robot's movement activity state (e.g., active or inactive due to failure) [110], and facilitating human oversight by communicating their external perception of the world [57].

While all of these examples are concerned with communicating *robot motion intent*, they differ tremendously in their methods and goals. Other researchers, such as Suzuki et al., have subsequently identified *robot motion intent* as an essential research area [113]. But beyond further solution approaches, the field needs a common understanding of the concept of *robot motion intent* (i.e., what do we actually mean by intent, what are relevant intent dimensions, and how does the communication of *robot motion intent* influence the relationship between robot and human).

To this end, we conducted a scoping review of current approaches to communicate *robot motion intent* in the literature. Based on our findings, we introduce an intent communication model of *motion intent*, which depicts the relationship between robot and human through the means of different intent dimensions (*intent type*, *intent information*, and *intent location*; see Figure 1). We further discuss these different intent dimensions and their interrelationships with different kinds of robots and human roles. Throughout our analysis, we classify the existing research literature along our intent communication model to form a design space for communicating *robot motion intent*. Practitioners and researchers alike may further benefit from this work for the design and selection of specific mechanisms to communicate *motion intent*. We identify future research directions and current gaps, which are further highlighted in an interactive website that lists the papers and allows comparisons based on user-selected categories.[1]

Our contribution is two-fold: 1) a survey contribution that includes our analysis and classification of previous literature as well as future research (cf. contribution from Wobbrock and Kientz [129]), and 2) a theoretical contribution that introduces an intent communication model and describes the relationship of its entities.

2 BACKGROUND

In this section, we will illustrate the need for communicating *robot motion intent* and discuss the current understanding of the term, which provides the foundation for our scoping review.

Robot is an umbrella term that describes a miscellaneous collection of (semi-)automated devices with various capabilities, technologies, and appearances[52]. These cyber-physical systems are often differentiated by their Degrees-of-Freedom (DoF) or ability to move and manipulate their environment. In industrial assembly lines, robotic arms manipulate and weld heavy parts [126], often in restricted areas [59]. Enabled by lightweight materials and safety sensors, robots have started to adapt to their users – today, they shut down when humans get too close or when resistance to the robot's movement is detected. This has led to the development of *cell-less* HRI [10], which has also paved the way for further scenarios, such as supporting people with disabilities in their daily lives [97]. Ajoudani et al. trace in their review paper several approaches of HRI, how it evolved, and how it increased over the last two decades [1]. They conclude that the success of HRI comes from combining human cognitive skills (i.e., intelligence, flexibility, and ability to act in case of unexpected events) with the robot's high precision and ability to perform repetitive tasks.

Matheson et al. proposed different types of such *cell-less* HRI, defined by their closeness of interaction [78]. They include *coexistence* (separation in space but not in time), *synchronized* (no separation in space but in time), *cooperation* (no separation in space or in time, but still not working on the same task), and *collaboration* (human and robot work on a task together, where the action of one has immediate consequences for the other). These works indicate that communication and interaction between robots and humans are critical to successful HRI. While research in human-aware navigation aims to make the robot smart enough to understand human

[1] Interactive Data Visualization of the Paper Corpus. https://rmi.robot-research.de, last retrieved February 16, 2023.

behavior and react to it [69], supporting humans in understanding robot behavior is equally important [68]. As the work by Matheson et al. highlights, humans and robots increasingly share the same physical space in HRI, which makes communicating *robot motion intent* a particularly relevant aspect for safe and effective collaboration and a prerequisite for *explainable robotics* [78].

However, *robot motion intent* is a rather vague term and lacks a clear definition. Further, it is not consistently used by researchers in the field. Instead, similar underlying concepts have been investigated under terms such as situational awareness [74], forthcoming operation [80], or robot signaling system [117]. Suzuki et al., as part of their extensive literature review covering the relationship between AR and robotics, emphasize the potential of AR-based visualizations for communicating movement trajectories or the internal state of the robot [113]. However, as their literature review extends beyond intent communication, they do not further discuss or define different types of intent, nor do they provide a deeper understanding of intent properties.

Our work presents a systematic overview of the field and addresses the current issues by conducting a scoping review. Such a review or survey contribution helps to organize the published research of the field and enables reflection on previous findings after the field has reached a level of maturity [129]. The goal of our review is to provide a clear understanding and definition of *robot motion intent*, its properties, and its relationships within HRI. Furthermore, our work provides a first discussion to relate our HRI findings to the growing domain of Automated Vehicles (AVs), so-called external Human-machine interfaces (eHMIs), which have identified similar research and design challenges [11, 28, 32, 33, 100].

3 METHOD

Scoping reviews provide an overview of the extent, range, and nature of evolving research areas. They help to summarize research findings and identify research opportunities [5, 123]. Our approach is in line with previous work by Ghafurian et al. [48], Muñoz et al. [85], and Wallkötter et al. [125]. We applied *Preferred Reporting Items for Systematic Reviews (PRISMA)* [94] guidelines, focusing on the *Meta-Analyses Extension for Scoping Reviews (PRISMA-ScR)* [119].

For an overview of each step in our paper selection process, please refer to Figure 2. We will discuss specific details of the individual steps in the following subsections. (1) Based on an initial screening of relevant literature, potential search terms were identified to perform a systematic query using three primary databases in the field of HRI (ACM Digital Library, IEEE Xplore, and ScienceDirect; see Section 3.1). (2) A filtering step was applied based on an algorithmic analysis of the total corpus to identify the most relevant terms related to the topic (see Section 3.2). (3) The resulting set of 822 papers was manually screened in a two-step process, and eventually, additional sources were found through a cross-check of the references in selected papers (see Section 3.3). The final corpus consists of 77 papers.

3.1 Initial Query

We explored a variety of query terms and their combinations because, as discussed, the field currently lacks a coherent and established terminology. In addition, we found several terms to be used in ambiguous ways, in particular terms such as *communication* and *motion*. Therefore, we decided on a broad search in this first step to increase recall and reduce the risk of overlooking relevant literature. We aimed to encompass a variety of different robot technologies while still focusing on the concept of intent, even though the word may be used in a variety of circumstances. We searched the titles, abstracts, and keywords of the databases' full-text collections with the following combined terms[2]:

$$(robot^* \textbf{ OR } cobot^* \textbf{ OR } drone^*) \textbf{ AND } (intent^* \textbf{ OR } intend^*) \tag{1}$$

3.2 Algorithmic Filtering

Due to our initial search being quite broad, further filtering was required to identify relevant papers. The initial set allowed us to apply an algorithmic approach similar to that of previous research done by O'Mara-Eves et al. [92]. Specifically, we applied the Term Frequency-Inverse Document Frequency (TF-IDF) [102] method to identify frequently used terminology within our corpus. TF-IDF has been shown to be suitable for information retrieval in literature reviews [73, 112]. First, we preprocessed the entries by a) combining each paper's title, keywords, and abstract into one field, b) fixing encoding issues such as *&* (and), ° (degree), and — (emdash), and c) converting the strings to lowercase as well as removing punctuation, numbers, symbols, and standard English stop-words from the corpus and replacing tokens with their lemmatizations [77]. For the creation of the TF-IDF-weighted document-term matrix, we calculated the Term Frequency (TF) for each term of our corpus, taking the static Inverse Document Frequency (IDF) into account, and computed the TF-IDF for each term over all documents. The resulting TF-IDF-weighted document-term matrix is shown in Table 1.
From the first 150 entries of the TF-IDF sorted list of tokens, three researchers independently qualified related terms to *communication* and *motion* – two terms we had decided to leave out of the initial broad query due to word ambiguity. During the following consensus process, we excluded related terms that were too general and ambiguous (e.g., "show" is frequently used in "Our results show[...]," "present" in "In this work we present[...]," "demonstrate" in "We demonstrate in our results[...]," or "perform" in "We performed a study[...]"). All identified terms were then used in the filtering step by applying the following logic to the title, keywords, or abstract of each paper in our corpus:

$$\begin{aligned}&(communicat^* \textbf{ OR } visual^* \textbf{ OR } feedback^*)\\&\textbf{AND } (motion^* \textbf{ OR } movement^* \textbf{ OR } interaction^*)\end{aligned} \tag{2}$$

For a paper to be accepted, a term from the cluster "communication" and another from "motion" (both OR operation) had to appear in the title, keywords, or abstract (AND operation). As a result, 822 papers remained in our corpus.

[2] ScienceDirect does not support the wildcard "*" but uses stemming and lemmatization techniques. In order to achieve search results based on wildcards "*," we modified the combined term to: *(robot **OR** cobot **OR** drone) **AND** (intent **OR** intention **OR** intend **OR** intended).*

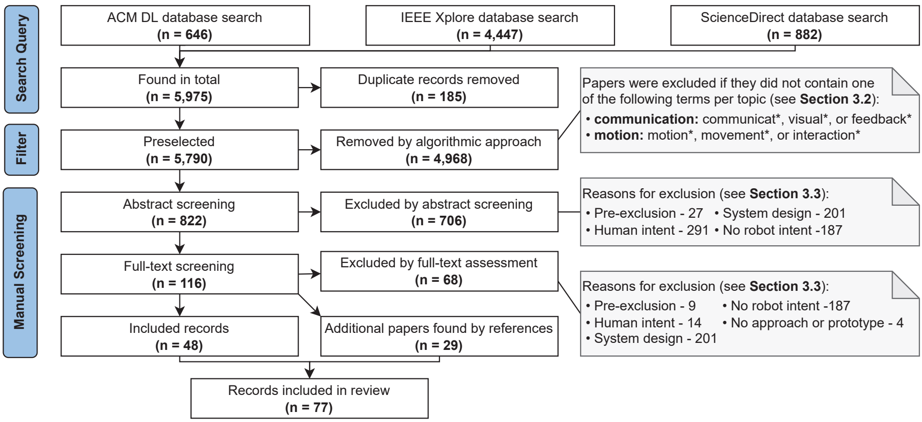


Figure 2: Flow chart of the corpus selection process with the identification of publishers and the initial search query (see Section 3.1), the reduction of the set by algorithmic filtering (see Section 3.2), and the manual screening (see Section 3.3), which resulted in 77 papers.

Table 1: Sorted list of terms from the TF-IDF-weighted document-term matrix. The selected terms are highlighted in bold.

Rank	Term	TF	IDF	TF-IDF	Rank	Term	TF	IDF	TF-IDF
1	human	6,547	0.92	6,052.89	7	**interaction**	3,383	1.33	4,515.61
2	control	6,769	0.87	5,902.24	15	**movement**	1,920	1.88	3,606.34
3	system	7,612	0.69	5,218.61	61	**communicat**	1,059	2.32	2,455.03
4	**motion**	3,640	1.42	5,154.59	140	**feedback**	665	2.74	1,820.08
5	model	3,978	1.24	4,938.74	143	**visual**	674	2.67	1,802.90

3.3 Manual Screening

The final phase of our paper selection process required manual screening, following an approach similar to that of Doherty and Doherty [34]. The process involved *abstract screening*, *full-text screening*, and *reference screening*. During the screening of all abstracts, we identified 706 out of 822 papers as not fitting into the scope of this review. The full-text analysis of the remaining 116 papers reduced the set to 48 papers. In addition, we screened the references cited by the set of 116 papers that were assessed for full-text screening. We identified 29 further relevant references, which we then included. This led to a final set of 77 papers, which were examined in the following. During the abstract and full-text screening, we **pre-excluded** 36 papers in unfitting paper formats still in the corpus, such as proceedings front matter, workshop calls, survey papers, or semi-duplicates – when two papers essentially presented the same contribution, due to one being a work in progress and the other a full paper. We also excluded 305 papers that aimed to convey the **human's intent** (to the robot) but not the robot's intent (e.g., Kurylo and Wilson [70]). Similarly, we removed another 210 papers where the research did not focus on the intention of robot motion (**no robot intent**). For example, 1:1 teleoperated devices (e.g., van Waveren et al. [121]), or work focusing on AVs and eHMIs. We excluded another 220 **system design** papers that focused on aspectus such as aesthetics, mathematical models of motion planning, or definitions (e.g., Girard et al. [50]). Eventually, we removed four papers where no approach or prototype was developed and reported (e.g., Thellman and Ziemke [118]).

4 INTENT COMMUNICATION MODEL

Through our literature review, we aim to improve understanding of the communication of *robot motion intent* by analyzing previous research. To that end, each author analyzed our literature corpus (n=77) in a multi-step process. It was discovered that several papers presented, combined, or empirically compared multiple intents (on average, more than two per paper). Therefore, we first systematically extracted all individual intents, resulting in a total of 172 intents. By screening these intents, we identified the primary entities (*robot*, *intent*, and *human*) as well as a communication flow between these entities that parallels that of the HCI model from Schomaker [104]. However, in contrast to the HCI model, we focus

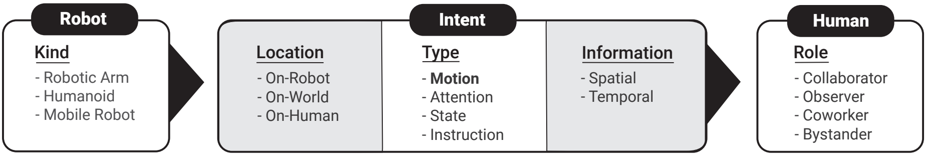


Figure 3: Overview of the intent communication model from robot to human. The three entities (i.e., robot, intent, human) and their dimensions are derived from our literature corpus. The flow of communication parallels the human-computer interaction model from Schomaker [104]. The main dimensions (i.e., kind, type, role) are discussed in Section 4, while a focused analysis of intent information and location is presented in Section 5.

solely on the communication of *intent* from *robot* to *human*, as previous research has already covered the inverse [62]. Furthermore, we identified a top-level entity, *goal*, which describes the motivation to communicate intent, as well as a low-level entity, *context*, which describes the situation in which the intent is communicated. Reflecting on all entities, we analyzed the intents by asking 1) *why* they were communicated (*goal*), 2) *who* communicated them (*robot*), 3) *what* they communicated (*intent*), 4) *to whom* they were communicated (*human*), and 5) *in which* circumstances they were communicated (*context*). Dimensions, categories, and properties emerged from the data through an open coding process of the extracted answers; specifically, we identified *kind of robot*, *location*, *type of intent*, *information* of *intent*, and *role of human* as our dimensions. The resulting *intent communication model* is shown in Figure 3. In the following, we present our findings for the three primary entities (*robot*, *intent*, and *human*), which we define and support by giving examples. We also discuss the *context* of communicating *robot motion intent*.

4.1 Human

In HRI, we can distinguish between different scenarios based on how involved a human is in the task performed by the robot. For the entity *human*, we utilize these levels of closeness between robot and human to define the different *roles of human*. Moreover, all four *roles of human* are illustrated in Figure 4.

4.1.1 Definition. The human has a crucial role during HRI, which strongly impacts which intents need to be communicated. From the analyzed intents of our corpus, we derived four different *roles of human* (*collaborator*, *observer*, *coworker*, and *bystander*). The roles are ordered by the degree of human collaboration and involvement with the robot, starting with the most involved (see Figure 4). These roles are also closely connected to the overarching goal of the HRI. Here, we found *supporting collaboration*, *oversight*, and *coexistence* to be of primary importance. In the following, we define the different roles, discuss their relationships to overarching goals, and support them with examples.

Collaborator. When in the role of a *collaborator*, a human works with a robot on a shared task in the same space and at the same time [78]. Thus, communication of *robot motion intent* in this context is for *supporting collaboration*. It aims to foster the coordination of robot and human actions regarding space and time to allow them to work together on a shared task (e.g., a human-robot assembly team in a manufacturing scenario [3]). The action of one of the two (i.e., robot or human) has immediate consequences for the other. For example, consider the scenario of a robot handing an object to a human [36, 89]. Here, the human has to precisely anticipate and coordinate with the time and place the object will be positioned to enable efficient handover. To that end, Dragan et al. propose a robotic arm that applies so-called *legible motion*, allowing the human to infer the goal of motion quickly and with certainty [36]. The role of a *collaborator* represents the closest degree of HRI, as they form a team in which both depend on each other. In our literature corpus, a *collaborator* is described in 18 papers and is the recipient of 37 different intents.

Observer. A human functions as an *observer* when their main job is to supervise the task that is being carried out by the robot. Although they mostly just watch, an *observer* must be ready to intervene and take control of the robot. In this context, communication of *robot motion intent* is for the goal of *supporting oversight*. Here, the robot has to provide information to the human to allow effective intervention when needed. Fundamentally, supporting oversight refers to the ability of a human to judge and evaluate if a robot is operating within its intended parameters. For example, in work by Hetherington et al., the robot communicates its movement paths to an *observer*, which enables the *observer* to foresee and prevent potential collisions of the robot with obstacles [60]. Others communicate the inner state of the robot, allowing an *observer* to anticipate potential task failures that may occur due to problems with the robot itself, e.g., faulty sensor information [8, 57]. An *observer* is described in 47 papers and is the recipient of 94 intents.

Coworker. In the *coworker* role, the human works next to the robot but handles their own task. While these tasks may be part of a shared overarching effort or entirely disconnected, they take place in the same shared workspace (e.g., a robotic arm that picks up one out of two objects and leaves the other one for the human [71]). In the *coworker* context, communication of *robot motion intent* is for the purpose of *supporting coexistence*. Here, the human needs to understand the robot's motion to avoid safety-critical situations (e.g., colliding with the robot). In Aubert et al., a robot and human pick up objects from a shared bin for their individual tasks [6]. Here, communication of *robot motion intent* can help the human to coordinate their actions and avoid collisions with the robot. Chadalavada

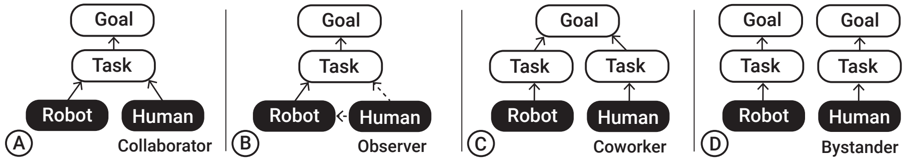


Figure 4: Comparison of the four human roles. The goals are further broken down by tasks to illustrate the relationships between the three entities (human, robot, and goals). A) The human (collaborator) and robot work on the same task. B) The human (observer) observes the robot and task but does not directly contribute. C) The human (coworker) and robot work on different tasks that contribute to the same goal. D) The human (bystander) and robot work on different tasks that contribute to different goals.

et al. showed that communication of *motion intent* through Spatial Augmented Reality (SAR) can improve perceived safety with mobile robots [20]. In their study, it meant that participants could choose safer walking paths and get closer to the robot without subsequent safety shutdowns. In our literature corpus, a *coworker* is described in six papers and is the recipient of 18 intents.

Bystander. The human is a *bystander* when they do not share the same task or the same task goal with the robot but still occupy an area overlapping the robot's physical workspace. Like the *coworker* role, the *bystander* role involves communication of *robot motion intent* to support the goal of *supporting coexistence*. A *bystander* needs motion information to avoid collision and feel safe. For example, imagine a human and a robot encountering each other in a corridor. To allow the human to choose a walking path that avoids collision, the robot can move to one side and communicate its intended movement path in advance [83, 127]. A *bystander* is described in 17 papers and is the recipient of 23 intents.

4.2 Intent

We identified four different types of *intent* that the *robot* can communicate to the *human* to express its intentions, contributing to increased transparency. We consider these types to be the main dimension for classifying *intent* in the following text. In addition, we identified the dimensions *location* and *information*, as shown in Figure 3, which help to further classify and describe *intent*. Given their great importance, they are discussed separately in Section 5.

4.2.1 Definition.

As our literature review focused on communicating *robot motion intent*, a majority of the corpus (69% of all papers; 54% of all unique intents) deals with *motion intent*. Nevertheless, we identified additional intent types that are related to *motion intent* and of equal importance (i.e., *attention*, *state*, and *instruction*). All *types of intent* are described below and the relationship of each to motion is explained. Furthermore, we found that for each *type of intent*, we can further distinguish between an *intent* that is *related to the robot* and one that is *related to the world* (more details can be found in the individual paragraphs below). An overview of all *types of intent* and associated papers can be found in Table 2.

Motion. These intents are the main *type of intent*. *Motion intent* is concerned with explicitly communicating future motions (i.e., actions that the robot will perform). As our survey is focused on *robot motion intent*, it encompasses more than 50% of the identified unique intents in our corpus. Most of the described intents deal with *robot self-actions*, aiming to indicate future robot movement. Thereby, users may be able to improve the coordination of their actions in concert with the robot's behavior to avoid collisions and improve safety. For example, Chadalavada et al. employed SAR to communicate future movement direction as well as the specific path the robot will take, which helped *bystanders* feel safe around a robotic forklift [20]. *World actions* are activities that manipulate the world around the robot. Again, this may help the *bystander* to coordinate their activities, but it also helps the *observer* to understand when to take over control from the robot. Psarakis et al. applied this concept of *world actions* in a VR simulation to visually augment the nearby objects that the robot planned to grasp [98].

Attention. Intents that communicate the need for attention are a supportive element. They precede a *motion intent* to shift human attention toward the robot or process, especially when the humans' attention is not guaranteed (e.g., because they focus on their own tasks). For example, Bolano et al. used acoustic feedback to alert the human and shift their attention toward the robot whenever it detected a possible collision [14]. An example of *robot-focused attention* was presented by Furuhashi et al., who designed an assistive robot based on the commercial Roomba device as a hearing dog that can notify deaf users of important events [45]. Here, the system uses physical touch to gain the human's attention by gently bumping into their body. As an example of *world-focused attention*, Mutlu et al. had a humanoid robot quickly look at an object of interest. They studied whether collaborators were able to understand the robot's gaze cues and correctly identify the object (among several others) that the robot had chosen as its object of interest [88]).

State. A robot communicating its state allows a *human* to deduce potential future motions and identify conflicts before they occur. For example, a *robot* could collide with nearby objects due to errors in its sensor system. However, robot communication of the detected objects enables a *human* to take over control and mitigate the issue. For *state* intents, we distinguish between *robot self-perception*, meaning the state the *robot* communicates about itself (e.g., simple text feedback presented on a display that indicates states such as "stop" or "moving" [80]), and *robot world perception*, meaning the

Table 2: Overview of different intent types, sorted by their categories and subcategories, with their counts (and percentages) of identified relevant papers (max. 77) and unique intents (max. 172). Note: Papers may include multiple unique intents and can therefore appear in multiple categories and subcategories.

Category	Subcategory	Number of Papers (%)	Number of Intents (%)	References
Motion	Robot Self-Actions	38 (49.35%)	75 (43.60%)	[3, 12–14, 16, 17, 20, 21, 23, 27, 30, 31, 35–37, 42, 44, 49, 54, 55, 58, 60, 63, 72, 79–83, 99, 101, 115, 120, 124, 127, 128, 130, 132]
	World Actions	15 (19.48%)	18 (10.47%)	[3, 6, 21, 25, 40, 41, 57, 61, 64, 66, 71, 84, 89, 95, 98]
Attention	Robot-Focused Attention	6 (7.79%)	8 (4.65%)	[6, 14, 19, 24, 45, 67]
	World-Focused Attention	4 (5.19%)	5 (2.91%)	[74, 88, 109, 111]
State	Robot Self-Perception	23 (29.87%)	27 (15.70%)	[3, 7, 8, 18, 29, 31, 38, 43, 55, 63, 74, 79, 80, 91, 105, 110, 114, 116, 117, 124, 128, 131, 132]
	Robot World Perception	8 (10.39%)	12 (6.98%)	[3, 21, 30, 31, 57, 101, 128, 132]
Instruction	Robot-Centered Instructions	10 (12.99%)	16 (9.30%)	[8, 19, 39, 45, 51, 67, 74, 86, 108, 117]
	World-Centered Instructions	9 (11.69%)	11 (6.40%)	[3, 8, 13, 16, 21, 22, 84, 98, 128]

communication of the perceived state of the world (e.g., visually highlighting objects in the environment that the sensor system has successfully detected, allowing the user to predict and understand subsequent robot movement [57]).

Instruction. In several papers, we identified *instruction* intents that accompany robot motion. For example, if a *robot* is blocked by an obstacle, it can instruct a *human* to remove the obstacle so it can continue its motion. *Instructions* can be *robot-centered instructions* when they stand in relation to the robot itself (e.g., Moon et al. applied head gaze cues to communicate instructions to the user to complete the handover of an object from the robot's gripper [84]). Or, in contrast, *instructions* can be *world-centered instructions* when they stand in relation to the world (e.g., a robot instructing a human to push a button on a wall to open an elevator so that it can continue its movement [128]).

4.2.2 Relationship to Human. Communicating a robot's intended motion to the human helps to improve the perception and understanding of the robot's behavior. However, humans that are, for example, not involved in the robots' task – perhaps because they are focusing on their own tasks (*coworker*) or are just uninvolved in general (*bystander*) – often need an additional cue to be able to read *robot motion intent*, which makes the intent type *attention* necessary (e.g., by an acoustic prompt [6]). *State* intents enable a human to see not only the next *motion* but also the internal state and planning, enabling them to understand actions ahead of time. Such intents also support *observers* in their task of supervising the robot. Finally, collaboration means a constant shifting of who is in charge when humans and robots work together on a shared task. Therefore, *motion*, *state*, *attention*, and *instructions* are all necessary intents for providing a baseline for collaboration (*collaborator*).

4.3 Robot

In our corpus, we identified three different *kinds of robot*, which together form the *robot* entity.

4.3.1 Definition. We identified three main *kinds of robots*: *robotic arm*, *humanoid*, and *mobile robot*. These, in order, represent a spectrum of increasing mobility and flexibility based on the area of deployment, starting with stationary robots (still with many DoF) and ending with robots that are inherently mobile (which also includes mobile arms with many DoF on a platform). Based on different robots, researchers have investigated different intents with varying frequencies. In the following, we illustrate each *kind of robot* with examples from our literature corpus.

Robotic Arm. Robotic arms can be described as a chain of axis links. They are typically fixed to one place and can have a manipulator [47]. Nowadays, they are the industry standard in production lines of factories [15] and work alongside humans in HRI environments [35]. *Robotic arms* are described in 13 papers and send 22 intents.

Humanoid. Humanoids have two robotic arms with manipulators, a torso, a head, eyes, and, often, basic facial expressions. Due to the two robotic arms, *humanoids* have more DoF than single robotic arms. Still, *humanoids* are often fixed to one place and lack mobility. Nonetheless, they are an important part of HRI when working with humans in a shared workspace [72, 99]. In rare cases, they can move in space, imitating human movement. Here, anthropomorphic features of the robots – such as gaze or certain gestures – can decrease the time required to predict the robot's intent [49]. *Humanoids* are described in 11 papers and send 21 intents.

Mobile Robot. With the addition of mobility comes increased flexibility. *Mobile robots* can be deployed in the air, on the ground, or in water. For this kind of robot, we have actively chosen to define them more broadly to include robots that appear only once in the corpus. For *mobile robots* (also referred to as drones), we distinguish between *ground drones without a manipulator* that move between locations, *ground drones with a manipulator* that can also manipulate the world, *flying drones* that maneuver through the air, and *water drones* that operate on water or underwater. Communicating *motion intent* helps *ground drones without a manipulator* to, for example, lead or follow a human to a specific place [51]. It can help *ground drones with a manipulator* to, for example, communicate which object they intend to pick up [21]. *Flying drones* or *water drones*, on the other hand, can communicate their *motion intent* by flying or driving in a pattern [91, 114]. All kinds of drones can appear alone [27] or as a swarm of drones [17]. *Mobile robots* are described in 53 papers and send 129 intents.

4.3.2 Relationship to Intent. As *mobile robots* move around more freely, they frequently encounter human *bystanders* who cross their paths. Consequently, *mobile robots* often have to first shift the *human's* attention toward the robot's display, preparing them for the communication of the robot's intended *motion*. For example, a projection in front of the robot can catch the attention of a bystander while simultaneously informing about the direction of driving [79]. At the same time, *mobile robots* need to communicate their *state* and planning of actions ahead of time, either the inner state (e.g., what is the current mission status [74]) or the perceived world state (e.g., which objects are detected [31]). *Humanoids* and *robotic arms*, on the other hand, are often deployed in collaborative scenarios, teaming up with humans. Here, robots need to communicate their intended *motion* to coordinate their actions with a human collaborator (e.g., which items the robot intents to pick next from a shared bin [6] or when objects are to be handed over to the collaborator [89]).

4.4 Context

The *context* describes the setting of the HRI scenario. While the location is an essential part of the context, there is more: for example, the social environment [103]. Nonetheless, we consider the location helpful to define HRI scenarios. In our analysis, we found various types of locations, including *workplace*, *domestic*, and *outdoor*. In *workplace* settings, the robot is frequently part of an assembly line or, more generically, a manufacturing process (e.g., collaborating with a human worker [117]). However, *workplace* locations also include industrial settings, offices, or generic work rooms. In total, 42 papers took place at a *workplace* location. In *domestic* environments, robots support a task at home (e.g., by picking cups up off a kitchen table [36]). Here, we found five relevant papers. Finally, in two papers the robot could move freely outside (e.g., fulfilling a mission and communicating its status [38]). Apart from these, 28 papers had no particular location specified. Instead, the authors of these papers investigate more generic scenarios of *robot motion intent* (e.g., by stating that a robot moves between two locations but without fine details of these locations [80]). For these scenarios, it is unclear which locations are most relevant.

5 ANALYSIS OF INTENT INFORMATION AND LOCATION

In addition to the different *types of intent* discussed in the previous section, two other dimensions of intent emerged from the data: *Intent information* (which refers to the data communicated by the *robot*) and *intent location* (which describes from where the intent is communicated to the *human*). In this section, we define these dimensions, illustrate their application with examples, and present a summary of empirical findings concerning their usage.

5.1 Intent Information

Based on our analysis of *how* the intent is communicated as well as *what* is communicated, we derived two main properties for categorizing *intent information*: *spatial* and *temporal*.

5.1.1 Spatial Property. The primary approach to convey spatial information is to embed it directly into the environment, i.e., have it **registered in space**. We identified 105 matching intents. We can further classify such intents as conveying *local* information (74 intents) or *directional* information (31 intents). *Local* information aims to precisely relate the information to the surrounding space by showing an exact position that naturally may contain orientation information as well. Han et al., as an example, convey *local* information by using SAR polygon visualizations to frame and highlight detected objects on a table, allowing a human observer to supervise the robot's intended movement and manipulation actions [57]. In contrast, *directional* information aims to communicate the explicit direction of movement (e.g., an arrow pointing in the direction of movement [20] or toward an object or person of interest [61]).

Information that is **unregistered in space**, however, employs an abstract encoding of the spatial property. In total, we identified 67 matching intents. This category includes the following *types of intent*: *Description*, *symbol*, and *signal*. *Description* (11 intents) applies to scenarios in which textual or verbal information is used (e.g., the robot informs the human verbally before initiating a movement to perform a touch [25]). *Symbol* (25 intents) applies to cases in which a symbolic representation is used to form the intent communication (e.g., a mobile robot that nods its head to request a human follow before moving toward its destination [39]). *Signal* (31 intents) applies when components are turned on/off to indicate a change (e.g., an acoustic prompt is turned on to gain attention for the upcoming communication of *motion intent* [6]). Mini maps provide an abstract but geographical encoding that includes the relationships among different objects in the environment [22, 124, 132].

Empirical Implications. While information *registered in space* provides a direct link between real-world objects and the displayed information, information *unregistered in space* lacks this connection and requires an additional mental step to create this link. Consequently, information *unregistered in space* may be less intuitive, and thus researchers have explored different combinations of information to mitigate that. Andersen et al. as well as Wengefeld et al. showed that combining multiple types of intent information that are *unregistered in space* (e.g., text *description* and *symbol* icons) helps to effectively communicate *motion intent* to the user [3, 128]. On the other hand, Staudte and Crocker found that combining both categories (*registered & unregistered*), which in their case involved a robot gazing at a specific object while a verbal description of the object played, leads to successful perception and understanding by the user [111]. Similarly, Bolano et al. later showed that a verbal description of the target can be combined with visual feedback of the motion endpoint to achieve the same improvement [14].

5.1.2 Temporal Property. The temporal property of *intent information* is about the distinction between having a *discrete* or *continuous* information flow. **Discrete** information has a fixed, distinct appearance in time and is beneficial for communicating *robot motion intent* because it enables the human to detect a change (i.e., the information appears) and it signals at which point the information loses its relevance (i.e., it disappears). For example, Aubert et al. equip their humanoid robot with a display that shows the number of the next bin it will approach, thereby allowing a human to avoid conflict with the robot [6]. Overall, we identified 89 intents that communicate *discrete* information. **Continuous** information, as has been provided in 83 intents, is available throughout the whole task or over several task phases (i.e., it is visible independent of its relevance to the current task). It enables the human to observe the robot, compare it with the world, and evaluate the correct task execution. Tsamis et al., for example, implemented AR visualizations for a Head-Mounted Display (HMD) to continuously communicate the intended movement space of a robotic arm by placing a semitransparent red sphere around the robotic arm [120].

Empirical Implications. Faria et al. showed that both *discrete* and *continuous* information are effective for communicating a *follow me* intent with spherical robots [39]. Koay et al. also evaluated both temporal properties using a robot dog that guides people living with hearing loss. However, they found that a motion-based approach (*continuous*), in which the robot's head movements request users to follow, is more successful than using a flashing Light-Emitting Diode (LED) stripe (*discrete*). They attribute this to the fact that head movements are more straightforward to interpret [67]. The findings of Aubert et al. suggest that combining *discrete* and *continuous* information is the most effective method. They showed that the combination of a motion-based approach (*continuous*) and a display approach (*discrete*) to communicate the robot motion end-point outperformed both uni-modal intent communication conditions [6].

5.1.3 Cross Relations. Inherently, the information of every intent has *spatial* and *temporal* properties. In the following, we describe the relationships between these properties of intent information.

For *unregistered in space*, the temporal property is almost evenly distributed between *discrete* and *continuous* information. Here, *signal* is an exception, as *discrete* (23 intents; e.g., having flashing lights attached to a mobile robot to indicate a discrete change of movement direction, similar to a car [60]) is used more often than *continuous* (eight intents; e.g., an LED stripe attached to the robot to continuously communicate the remaining distance to the target position through a color-coded progress bar [8]). *Signals* are primarily used to communicate sudden changes. Accordingly, such *discrete* events are naturally communicated as *discrete* intent information.

For *registered in space*, we see an uneven distribution for both subcategories. Intent information classified as *local* is mostly communicated as *continuous* information (50 intents; e.g., using SAR to continuously highlight an area in a workplace where the robot will be active during its movements and action [3]) instead of *discrete* (24 intents; e.g., using SAR to highlight a button on a wall that must be pushed by a human for the robot to continue its movement [128]). We think that robot *motion* likely relates to a continuous event because it is meant to happen over time and takes place continuously. Intent information classified as *directional* is mostly communicated as *discrete* information (23 intents; e.g., a display is attached to the top of a mobile robot, communicating the intended movement direction with an arrow [80]) and only seldom as *continuous* (8 intents; e.g., a drone is visualized as an eye in AR, constantly looking in the direction of movement [124]). The reason is that *directions* are primarily used to communicate an updated movement direction to the human; therefore, it makes sense that they are most often given as *discrete* information.

5.2 Intent Location

Various technologies can enable the communication of *robot motion intent*. We found that, in particular, the placement of these technologies (*on-robot*, *on-world*, and *on-human*) can help to classify the different approaches in the literature, as there is often a relationship between the placement and specific types of technology.

On-Robot can be further divided into *robot-only* technology or additional *robot-attached* devices. We identified 114 intents communicated through *on-robot* technology. As an example for the subcategory *robot-only*, Moon et al. utilize the head orientation of the robot, mimicking a gaze cue, to communicate mid-air locations for its intended movement as an instruction to the user [84]. Nearly half of all categorized intents that utilize *on-robot* technology fall into that subcategory, which is of particular interest because it limits the need for additional technology and often involves imitation of human-to-human behavior. The *robot-attached* subcategory requires some additional hardware to be mounted to the robot (e.g., SAR, LED, or displays). For example, Wengefeld et al. attach a laser projection system to the robot and thereby communicate various types of intents, including *state*, *motion*, and *instruction* [128].

On-World has received relatively little attention in the literature. It includes, for example, small displays attached to the workspace at object bins [6], or a desktop display (to visualize *motion intent*) with speakers (to gain *attention*) next to the robot's workspace [14]. While the inability to change the environment may be less desirable from a generalizability perspective, for some technology, it adds significant benefits. In particular, SAR would be easier to realize with a fixed projector position *on-world* and it would allow for larger projection areas. We identified eight different intents *on-world*.

On-Human includes *head-attached* technologies, which primarily refers to HMD devices, which allow more complex visualizations. Gruenefeld et al., for example, experimented with different spatial visualizations, such as visualizing the intended movement path, previewing future locations of the robot arm, or visualizing the activity area as a whole [54]. In addition, some approaches rely on *hand-held* technologies. Correa et al., for example, used a tablet device displaying various types of information (map, live view, next steps) to support oversight and communicate *motion intent* [31]. We identified 50 intents *on-human*.

Empirical Implications. For the *intent location*, it is generally better to output information closer to the target. For example, LeMasurier et al. compared several motion-based and light-based approaches for *humanoids* to communicate an intended start of movement at an assembly workplace. They saw that an LED bracelet located closest to the workspace was the most noticeable and least confusing [72]. Furthermore, researchers found evidence that humans may prioritize *on-human* technology over *on-robot* technology. For example, Che et al. were able to show that the use of a vibrotactile bracelet worn by the user led to a better expression of the robot's *motion intent*, reduced users' effort, and increased users' trust in the robot during a collision-avoidance movement when compared to a solely robot-based approach using *legible motion* [23]. Finally, combining multiple output technologies can further increase performance. For example, Mullen et al. investigated a multi-modal approach for communicating robot interference in a sorting scenario that combined an AR-HMD visualization and active feedback via a vibrotactile bracelet. They found that combining both feedback types outperformed the single modality baselines. It allowed the human to more efficiently teach the robot and decreased the required interaction time. [86].

5.3 Relation between Location and Information

In the following, we provide insights into the relationship between *intent location* and *intent information* (cf. Table 3).

5.3.1 Registered in Space. To communicate location information registered in space, most researchers rely on *head-attached* technologies, such as AR-HMDs (*on-human*). For example, Tsamis et al. implemented AR visualizations to communicate an intended movement trajectory of a robotic arm [120]. They placed small spheres along a defined path in 3D space from the robot's end-manipulator to a specific destination. They found that using their system improved task completion and robot idle times, with fewer interruptions to the overall workflow. In addition, users reported increased feelings of safety and trust toward the robot. In contrast, Correa et al. proposed a tablet visualization that showed a live camera feed of the mobile robot highlighting recognized objects in its environment via a wireframe in the visualization [31]. In addition to intents displayed *on-human*, robots are often used to convey information directly through specific movements or pointing (*on-robot*). For example, Holladay et al. used a robotic arm and its end-effector to communicate a directional cue by pointing toward an object placed on a table [61]. The resulting pointing configurations were reported to make it easier for novice users to infer the target object. Another example for displaying information *on-robot* is provided by Hetherington et al. They used SAR to project an arrow in the intended movement direction of the mobile robot on the floor [60]. Their results show that projected arrows were more socially acceptable and more understandable than flashing lights. Finally, information *registered in space* can be outputted *on-world*. For example, Cleaver et al. used their web-based environment [26] to compare four different conditions of visualizing the intended movement trajectory of a mobile robot on a *world*-located display [27]. In contrast, Aubert et al. placed small displays on three bins and used bin numbers and progress bars to indicate from which bin the robot coworker would next withdraw an item. However, the display-based approach could not significantly reduce the number of physical conflicts [6].

5.3.2 Unregistered in Space. Interestingly, a relatively large number of *symbol* information is communicated through the robot itself (*on-robot*). Here, we found many approaches where the robot performs specific movement patterns that the human has to decode appropriately. A symbolic approach is shown by LeMasurier et al. [72]. They slightly move the robot's manipulator to the left and right to communicate an intended movement start. This approach received relatively high ratings on several measures; however, the authors recommend that the addition of light signals near the workspace and the origin of motion (like an LED bracelet) may provide a benefit to HRI in shared spaces. Song and Yamada provide an example of the type *symbol* by using different static and dynamic light patterns on a *robot-attached* colored LED stripe to illustrate different *states* of the robot [108]. Communication of *signal* information is mainly achieved through robot-attached technology, such as LED or audio speakers. Wearable technologies can also show spatially *unregistered* information (*on-human*). Che et al. propose a vibrotactile bracelet worn by the user to communicate an initiated collision-avoidance movement of a *mobile robot* [23]. This approach led to a better expression of the robot's *motion intent*, reduced users' effort, and increased users' trust in the robot. Furthermore, Walker et al. implemented a radar-like mini-map in the corner of an AR visualization to illustrate the relative position of the user to a drone [124]. Although the radar provides the user with the means to rapidly locate the robot relative to their own position, some participants mentioned that they did not need to use the radar much because they always faced the drone. Finally, unregistered information can also be presented *on-world*. Bolano et al. propose verbally describing the updated destination of the robot's end-manipulator via a speaker in addition to the screens placed in the shared workspace [14]. They found that users better understood the robot's intended motion, including when the robot had to reroute itself to avoid collision.

5.3.3 Discrete. Discrete information is usually presented directly *on-robot*. As an example of *robot-attached* technology, Domonkos et al. attached a colored LED stripe to the base of a robotic arm to communicate the intended direction of movement to a human *coworker* [35]. In contrast, Glas et al. proposed a *mobile robot* that performs head gestures to initiate either a follow-me or lead-me request to the human [51], relying on the robot itself as in *robot-only*. Gu et al. evaluated a visual feedback displayed through an AR-HMD (*on-human*), indicating the planned movement direction of the robot via an arrow visualization [55]. They found that the visualization improved perceived safety and task efficiency. Instead of relying on the visual modality, Mullen et al. proposed discrete feedback through a vibrotactile bracelet that is activated to communicate robot interference, triggering the human to move in order to allow the robot to continue its movement [86]. Their findings show that vibrational feedback can reduce the time required to notice and respond to an intent. Aubert et al. equipped bins (from which items could be chosen) in the environment with speakers to emit *discrete* auditory information *on world* [6]. They recommend not solely relying on auditory information, but using it in a multi-modal approach, which is further supported by Bolano et al. [14].

5.3.4 Continuous. Like *discrete* information, *continuous* information is primarily displayed *on-robot*. Matsumaru et al. *attached* an omnidirectional display *on-robot*, projecting an eyeball-like visualization that effectively communicates the direction of movement to a human [81]. In contrast, Dragan et al. propose performing legible motions with a robotic arm itself to communicate the next object it will grasp [36], which they found enabled fluent collaboration. As an example of communicating intents *on-human*, Walker et al.

Table 3: Overview of intents with different properties of *intent information* (by rows) in combination with *intent location* (by columns) – up to three example references are listed for each category. Please note that each intent has a spatial and a temporal property.

Category	Subcategory	On-Human		On-World	On-Robot	
		Head-Attached	Hand-Held		Robot-Only	Robot-Attached
(Spatial) Registered	Local	**35** [54, 99, 124]	**3** [31, 127]	**4** [6, 14, 27]	**22** [12, 16, 36]	**10** [30, 60, 128]
	Directional	**3** [55, 101, 124]	**0**	**0**	**14** [61, 83, 84]	**14** [20, 60, 80]
(Spatial) Unregistered	Description	**0**	**1** [31]	**1** [14]	**0**	**9** [79, 111, 128]
	Symbol	**5** [124, 132]	**0**	**1** [22]	**14** [51, 67, 72]	**5** [3, 7, 108]
	Signal	**0**	**3** [23, 24, 86]	**2** [6, 14]	**0**	**26** [35, 115, 117]
Total		**43** (25.00%)	**7** (4.07%)	**8** (4.65%)	**50** (29.07%)	**64** (37.21%)
(Temporal) Discrete		**15** [55, 89, 98]	**4** [23, 24, 86]	**5** [6, 14]	**19** [45, 49, 72]	**45** [19, 39, 130]
(Temporal) Continuous		**28** [21, 120, 132]	**3** [31, 127]	**3** [14, 22, 27]	**31** [17, 18, 36]	**19** [29, 57, 81]

display a symbolic representation of a focusing eye lens in an AR-HMD, encoding the relative distance to the next target [124]. Their results show a significant improvement in users' understanding of *robot motion intent*. Watanabe et al. proposed presenting *continuous* visual feedback via a tablet to inform a wheelchair passenger of a robot's intended motion path [127]. Lastly, *continuous* information can be displayed *on-world*. Chandan et al. proposed a map visualization for a stationary tablet display that continuously shows the locations of three mobile robots and other objects of interest [22]. They found this approach significantly improved the participants' ability to observe and assist the robot. Similarly, albeit only studied in a web-based experiment, Cleaver et al. proposed a 3D visualization displayed on a 2D screen to continuously communicate the intended path of a mobile robot [27].

6 DISCUSSION AND FUTURE RESEARCH

In the following, we discuss key findings of our literature survey and formulate future research directions as takeaway messages for the HCI community. The organization of the section follows the three entities *human*, *intent*, and *robot* from our intent communication model and concludes with a discussion of the overall model.

Human. From the analyzed intents of our corpus, we derived four different *roles of human* (*collaborator*, *observer*, *coworker*, and *bystander*). In our analysis, we found that the human role is strongly related to the overarching goals of communicating *motion intent* – a specific goal can be directly derived given a specific human role. For example, if the HRI scenario involves the human taking the role of an *observer*, the *motion intent* needs to help with fostering oversight. As a result, this indicates that practitioners and researchers should explicitly define the role and, thereby, the involved human stakeholders before settling on the robot or specific intents they may want to communicate. The human roles we found in a bottom-up process through our analysis align well with the previous work of Onnasch and Roesler [93]. In contrast to Onnasch and Roesler, the role of the *operator* did not show up in our analysis. We suggest this is because robots are not manually operated by humans in our corpus, as this would not require the robot to communicate any intent [53].

Future Research: Our analysis showed that nearly all papers a) investigate individual human roles, e.g., they (often implicitly) pick one and focus on that, and b) design and study only for a 1:1 relationship between human and robot. The only exceptions to this are Faria et al., Kirchner et al., and Palinko et al., who investigate the legibility of robot movement for a group of humans [41] or explore the use of gaze cues to allow the robot to choose their human collaboration partner from a group of humans [66, 95]. This limited involvement of multi-user groups is, of course, to be expected in an emerging field that first needs to establish certain ground truths. Involving multiple persons or even multiple robots and persons complicates HRI tremendously, yet we think this is the subsequent step research must take. In particular, it would be interesting to reflect on the suitability of specific technologies (e.g., SAR will likely be better suited to satisfy multi-user scenarios compared to HMD technology).

Intent Types. Through our scoping review of *robot motion intent*, we observed that communication of motion often requires additional intents that serve as pre- or post-cursors to the communicated *motion intent*. Furthermore, we found that robot motion can also be indirectly communicated: For example, by communicating only the robot's state (e.g., [8]) or by instructing a human to open a door so the robot can continue on its path (e.g., [127]). These various *types of intent* demonstrate the different facets of *robot motion intent*, which represent both actual intended movement trajectories and related communication. We see that as a key finding, distinguishing our work from previous research that focuses primarily on the communication of *motion intent* [99, 113, 124]. With our survey, we are confident that other researchers will start to adopt a more holistic and precise use of the term *robot motion intent* and, for example, start highlighting the need for related intents, as we found in our analysis.

Future Research: Researchers should investigate how the different *types of intent* may best be combined to achieve specific intent communication goals. Currently, there is little empirical knowledge about, for example, when and to what extent a robot may need to first communicate *attention* before effectively being able to communicate *motion intent*. Further research should also challenge our classification of *types of intent* and potentially extend them.

Intent Information and Location. We derived two main properties that categorize our identified *intent information* related to space: *registered in space* (61.05%) and *unregistered in space* (38.95%). This almost-even distribution reveals that a lot of relevant research not only focuses on information that aims to convey *local* or *directional* information (e.g., a resulting trajectory [27]), but also on more abstract representations, namely *description*, *symbol*, and *signal*. These are often much less complex and indicate that *robot motion intent* can be communicated without visual 3D representations of future movement. This shows that there are viable alternatives to wearing special *on-body* technology, resulting in fewer system costs and a decreased setup time. An alternative can be the *intent location on-robot*. In previous work, researchers have refined robots with anthropomorphic elements – such as eye-like features or certain movement gestures – to communicate motion intent. Our literature review identified 15 such instances, specifically applying eye- or head-gaze (e.g., looking at an object to indicate a handover between human and robot [84]). While anthropomorphic elements may not be as precise as digital representations through technology means (e.g., visualizations in AR), they share the same baselines as in Human-Human Collaboration (HHC). The general assumption is that, in turn, they can be easily understood by users and can mostly be integrated into the actual HRI. A possible combination with a verbal description provides a multi-modal output to the user, resulting in faster recognition of the specific object [111].

Future Research: While previous research has explored combinations of spatially registered and unregistered information [111], we are unaware of research that has contrasted their effectiveness. Therefore, current design decisions may be based more on the availability of particular technology and less on the intended outcome. Future research should explore this further so that practitioners can more accurately judge the potential trade-offs between simple or complex information and related technology use. Regarding the use of anthropomorphic features, the integration of such communication cues has been explored regarding their legibility and effectiveness in communicating *robot motion intent*. However, their implicit consequences (e.g., causing the human to ascribe human-like behavior to the robot) may still need to be fully explored. The means and cues of communication have significant consequences for the trust relationship between humans and robots [56].

Robot. When looking at the three *kinds of robots* and their usage in research, we can see that the physical properties of a robot have a large impact on communication means: In particular, the *on-robot* location for intent communication. Some robots come with pre-installed displays, while others have anthropomorphic features built in. *Flying drones*, on the contrary, require some kind of remote communication tool (often in the form of HMDs) to communicate over a larger distance. Robots are also an area of much technical experimentation, i.e., many researchers are building or customizing their own robots. For example, one may add anthropomorphic features to a robotic arm. As a result, researchers tend to use these built-in or customized features to communicate intent. They may often have only a particular kind of robot available; thus, they are limited to a certain way of communicating *robot motion intent*. Of course, this limits the generalizability of current findings, as each robot conveys unique features that can impact HRI.

Future Research: These findings show that many research endeavors explore only certain *kinds of robots*. A more systematic approach is called for to investigate the various kinds of robots and their impacts on communicating *robot motion intent*. We also found that more and more research applies simulation environments in Virtual Reality (VR) to explore HRI. Nevertheless, we need more studies to validate such findings and provide a broader foundation for their generalizability.

Context. Compared with previous research in AVs [28, 32] and eHMIs [33], we can identify several similarities, despite the substantial differences in the context of use and robot technology. Colley et al. found that visualizing internal information processed by an Augmented Virtuality (AV) could calibrate trust by enabling the perception of the vehicle's detection capabilities (and its failures) while only inducing a low cognitive load [28]. Currano et al. explored the interaction between complexity of head up displays, driving style, and situation awareness [32]. In the area of eHMIs, researchers have been able to distinguish between different *natures of message* (e.g., danger and safety zones) [33]. These correspond to our identified *types of intent*, highlighting different meanings for the user for the provided intent. In the context of AVs, the information used to formulate the actual intent is primarily unregistered in space. It uses text, symbols, and audio prompts. The intent primarily describes the vehicle's state (e.g., automated/manual, cruising, yielding) or advice/instructions to the pedestrian (e.g., to allow safe road crossing). The large differences between the fields of research result primarily from the standardizations in automotive research, such as roads, road signs, markings, and restrictions. Nevertheless, there are potential overlaps.

Future Research: The two fields have, from our perspective, not yet shared many cross-activities among researchers, which could lead, for example, to transferring those *motion intent* techniques that have shown to be effective in one field to the other. We could imagine that future research could benefit both sides if a more holistic perspective is applied. In particular, the research for eHMIs in AVs could benefit from more exploratory technological approaches in HRI, such as making use of AR-HMDs and applying more advanced visualization to communicate *motion intent*. While this may not be relevant for the near future, as such devices are not yet consumer-ready, this may change over the coming years.

The Model. The overall model is an abstract characterization of the current literature on *robot motion intent*. It may be seen as a summary of the current understanding of the design space for robot intent communication, where it illustrates all components and highlights their interconnection. Thereby, future researchers and practitioners should benefit from the model by using it as a guidance and checklist throughout the design phase of such Human-Robot scenarios; i.e., being guided to carefully think and decide upon different types of intents or whether intent information should be encoded spatially or temporally. In addition, the model can help to unify the language of *robot motion intent* and thereby support researchers and practitioners to find related work as well as help to identify research gaps.

Future Research: We invite researchers to actively challenge the model and thereby helping to develop the field even further. They should scrutinize whether the design space is sufficiently classified or how it can and needs to be extended to cover future work. As our model was derived from the analysis of our literature corpus, it is fitted to the gathered research. Nonetheless, one can utilize novel research contributions that will be published in the future to revisit and evaluate the model (i.e., to investigate if novel contributions can still be described by our model). Moreover, we imagine that a more thorough discussion in the context of eHMIs may benefit the model as well as incorporating other lines of research that are concerned with communicating intent, such as Sodhi et al. or Müller et al. [87, 107].

7 CONCLUSION

This paper provides two main contributions: 1) a survey contribution that includes an analysis and classification of previous literature as well as future research directions, and 2) a theoretical contribution that introduces an intent communication model and describes the relationships of its entities, dimensions, and underlying properties. In particular, our work highlights that *robot motion intent* requires a broader perspective on robot intent and that it includes intent types that may seem, at first glance, unrelated to motion. However, in our analysis, we found that *attention*, *state*, and *instruction* are important and often necessary pre- or post-cursors to communicate explicit *motion intent*. We also found that only a few papers explicitly discuss or present the type of intent they aim to communicate and they also lack clear descriptions of intent information or location. Our work aims to help researchers in the future to better align their work with the suggested dimensions, making it easier to assess and compare different studies. Therefore, we aim to provide a foundation for a unified language regarding *robot intent*, even beyond motion. From a practical perspective, the classification of the existing research literature along our *intent communication model* helps researchers and practitioners alike to understand the design space for communicating *robot motion intent*. As it is an emerging field, much work has focused on finding novel approaches and solutions to communicate *robot motion intent* in one way or another. We have identified multiple areas of need for future research directions. However, we would like to emphasize once more that, above all, the field needs more systematic analysis and comparison of different approaches to improve understanding of the influences of different intent dimensions and properties. We believe that the presented intent communication model provides an empirically deducted foundation to inspire and guide such work.

REFERENCES

References marked with • are in the set of reviewed papers.

[1] Arash Ajoudani, Andrea Maria Zanchettin, Serena Ivaldi, Alin Albu-Schäffer, Kazuhiro Kosuge, and Oussama Khatib. 2017. Progress and prospects of the human–robot collaboration. *Autonomous Robots* 42, 5 (Oct. 2017), 957–975. https://doi.org/10.1007/s10514-017-9677-2

[2] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi Iqbal, Paul N. Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz. 2019. Guidelines for Human-AI Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3290605.3300233

[3] • Rasmus S. Andersen, Ole Madsen, Thomas B. Moeslund, and Heni Ben Amor. 2016. Projecting robot intentions into human environments. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 294–301. https://doi.org/10.1109/ROMAN.2016.7745145

[4] Stephanie Arevalo Arboleda, Franziska Rücker, Tim Dierks, and Jens Gerken. 2021. Assisting Manipulation and Grasping in Robot Teleoperation with Augmented Reality Visual Cues. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 728, 14 pages. https://doi.org/10.1145/3411764.3445398

[5] Hilary Arksey and Lisa O'Malley. 2005. Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology* 8, 1 (Feb. 2005), 19–32. https://doi.org/10.1080/1364557032000119616

[6] • Miles C. Aubert, Hayden Bader, and Kris Hauser. 2018. Designing Multimodal Intent Communication Strategies for Conflict Avoidance in Industrial Human-Robot Teams. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1018–1025. https://doi.org/10.1109/ROMAN.2018.8525557

[7] • Alexandra Bacula, Jason Mercer, and Heather Knight. 2020. Legible Light Communications for Factory Robots. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, Tony Belpaeme, James Young, Hatice Gunes, and Laurel Riek (Eds.). ACM, New York, NY, USA, 119–121. https://doi.org/10.1145/3371382.3378305

[8] • Kim Baraka, Stephanie Rosenthal, and Manuela Veloso. 2016. Enhancing human understanding of a mobile robot's state and actions using expressive lights. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 652–657. https://doi.org/10.1109/ROMAN.2016.7745187

[9] Andrea Bauer, Dirk Wollherr, and Martin Buss. 2008. Human–Robot Collaboration: A Survey. *International Journal of Humanoid Robotics* 05, 01 (2008), 47–66. https://doi.org/10.1142/S0219843608001303

[10] Wilhelm Bauer, Manfred Bender, Martin Braun, and Rally, Peter und Scholtz, Oliver. 2016. Lightweight Robots in Manual Assembly – Best to Start Simply: Examining Companies' Initial Experiences with Lightweight Robots. https://www.edig.nu/assets/images/content/Studie-Leichtbauroboter-Fraunhofer-IAO-2016-EN.pdf

[11] Pavlo Bazilinskyy, Dimitra Dodou, and Joost de Winter. 2019. Survey on eHMI concepts: The effect of text, color, and perspective. *Transportation Research Part F: Traffic Psychology and Behaviour* 67 (2019), 175–194. https://doi.org/10.1016/j.trf.2019.10.013

[12] • Christopher Bodden, Daniel Rakita, Bilge Mutlu, and Michael Gleicher. 2018. A flexible optimization-based method for synthesizing intent-expressive robot arm motion. *The International Journal of Robotics Research* 37, 11 (2018), 1376–1394. https://doi.org/10.1177/0278364918792295

[13] • Gabriele Bolano, Yuchao Fu, Arne Roennau, and Ruediger Dillmann. 2021. Deploying Multi-Modal Communication Using Augmented Reality in a Shared Workspace. In *2021 18th International Conference on Ubiquitous Robots (UR)*. IEEE, 302–307. https://doi.org/10.1109/UR52253.2021.9494689

[14] • Gabriele Bolano, Arne Roennau, and Ruediger Dillmann. 2018. Transparent Robot Behavior by Adding Intuitive Visual and Acoustic Feedback to Motion Replanning. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1075–1080. https://doi.org/10.1109/ROMAN.2018.8525671

[15] Sara Bragança, Eric Costa, Ignacio Castellucci, and Pedro M. Arezes. 2019. *A Brief Overview of the Use of Collaborative Robots in Industry 4.0: Human Role and Safety*. Springer International Publishing, Cham, 641–650. https://doi.org/10.1007/978-3-030-14730-3_68

[16] • Maya Cakmak, Siddhartha S. Srinivasa, Min Kyung Lee, Sara Kiesler, and Jodi Forlizzi. 2011. Using spatial and temporal contrast for fluent robot-human hand-overs. In *Proceedings of the 6th international conference on Human-robot interaction - HRI '11*, Aude Billard, Peter Kahn, Julie A. Adams, and Greg Trafton (Eds.). ACM Press, New York, New York, USA, 489. https://doi.org/10.1145/1957656.1957823

[17] • Beatrice Capelli, Cristian Secchi, and Lorenzo Sabattini. 2019. Communication Through Motion: Legibility of Multi-Robot Systems. In *2019 International Symposium on Multi-Robot and Multi-Agent Systems (MRS)*. IEEE, 126–132. https://doi.org/10.1109/MRS.2019.8901100

[18] • Jessica R. Cauchard, Kevin Y. Zhai, Marco Spadafora, and James A. Landay. 2016. Emotion encoding in Human-Drone Interaction. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 263–270. https://doi.org/10.1109/HRI.2016.7451761

[19] • Elizabeth Cha and Maja Mataric. 2016. Using nonverbal signals to request help during human-robot collaboration. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 5070–5076. https://doi.org/10.1109/IROS.2016.7759744

[20] • Ravi Teja Chadalavada, Henrik Andreasson, Maike Schindler, Rainer Palm, and Achim J. Lilienthal. 2020. Bi-directional navigation intent communication using spatial augmented reality and eye-tracking glasses for improved safety in human–robot interaction. *Robotics and Computer-Integrated Manufacturing* 61 (2020), 101830.

[21] • Tathagata Chakraborti, Sarath Sreedharan, Anagha Kulkarni, and Subbarao Kambhampati. 2018. Projection-Aware Task Planning and Execution for Human-in-the-Loop Operation of Robots in a Mixed-Reality Workspace. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 4476–4482. https://doi.org/10.1109/IROS.2018.8593830

[22] • Kishan Chandan, Vidisha Kudalkar, Xiang Li, and Shiqi Zhang. [n.d.]. Negotiation-based Human-Robot Collaboration via Augmented Reality. https://doi.org/10.48550/arXiv.1909.11227

[23] • Yuhang Che, Allison M. Okamura, and Dorsa Sadigh. 2020. Efficient and Trustworthy Social Navigation via Explicit and Implicit Robot–Human Communication. *IEEE Transactions on Robotics* 36, 3 (2020), 692–707. https://doi.org/10.1109/TRO.2020.2964824

[24] • Yuhang Che, Cuthbert T. Sun, and Allison M. Okamura. 2018. Avoiding Human-Robot Collisions Using Haptic Communication. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 5828–5834. https://doi.org/10.1109/ICRA.2018.8460946

[25] • Tiffany L. Chen, Chih-Hung King, Andrea L. Thomaz, and Charles C. Kemp. 2011. Touched by a robot. In *Proceedings of the 6th international conference on Human-robot interaction - HRI '11*, Aude Billard, Peter Kahn, Julie A. Adams, and Greg Trafton (Eds.). ACM Press, New York, New York, USA, 457. https://doi.org/10.1145/1957656.1957818

[26] Andre Cleaver, Darren Tang, Victoria Chen, and Jivko Sinapov. 2020. HAVEN: A Unity-based Virtual Robot Environment to Showcase HRI-based Augmented Reality. https://doi.org/10.48550/ARXIV.2011.03464

[27] • Andre Cleaver, Darren Vincent Tang, Victoria Chen, Elaine Schaertl Short, and Jivko Sinapov. 2021. Dynamic Path Visualization for Human-Robot Collaboration. In *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, Cindy Bethel, Ana Paiva, Elizabeth Broadbent, David Feil-Seifer, and Daniel Szafir (Eds.). ACM, New York, NY, USA, 339–343. https://doi.org/10.1145/3434074.3447188

[28] Mark Colley, Benjamin Eder, Jan Ole Rixen, and Enrico Rukzio. 2021. Effects of Semantic Segmentation Visualization on Trust, Situation Awareness, and Cognitive Load in Highly Automated Vehicles. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM. https://doi.org/10.1145/3411764.3445351

[29] • Emily C. Collins, Tony J. Prescott, Ben Mitchinson, and Sebastian Conran. 2015. MIRO. In *Proceedings of the 12th International Conference on Advances in Computer Entertainment Technology*, Adrian David Cheok, Mohd Shah Sunar, Ken Neo, and Yoram Chisik (Eds.). ACM, New York, NY, USA, 1–4. https://doi.org/10.1145/2832932.2832978

[30] • Michael D. Coovert, Tiffany Lee, Ivan Shindev, and Yu Sun. 2014. Spatial augmented reality as a method for a mobile robot to communicate intended movement. *Computers in Human Behavior* 34 (2014), 241–248. https://doi.org/10.1016/j.chb.2014.02.001

[31] • Andrew Correa, Matthew R. Walter, Luke Fletcher, Jim Glass, Seth Teller, and Randall Davis. 2010. Multimodal interaction with an autonomous forklift. In *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 243–250. https://doi.org/10.1109/HRI.2010.5453188

[32] Rebecca Currano, So Yeon Park, Dylan James Moore, Kent Lyons, and David Sirkin. 2021. Little Road Driving HUD: Heads-Up Display Complexity Influences Drivers' Perceptions of Automated Vehicles. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*. ACM. https://doi.org/10.1145/3411764.3445575

[33] Debargha Dey, Azra Habibovic, Andreas Löcken, Philipp Wintersberger, Bastian Pfleging, Andreas Riener, Marieke Martens, and Jacques Terken. 2020. Taming the eHMI jungle: A classification taxonomy to guide, compare, and assess the design principles of automated vehicles' external human-machine interfaces. *Transportation Research Interdisciplinary Perspectives* 7 (2020), 100174. https://doi.org/10.1016/j.trip.2020.100174

[34] Kevin Doherty and Gavin Doherty. 2018. The construal of experience in HCI: Understanding self-reports. *International Journal of Human-Computer Studies* 110 (2018), 63–74. https://doi.org/10.1016/j.ijhcs.2017.10.006

[35] • Mark Domonkos, Zoltan Dombi, and Janos Botzheim. 2020. LED Strip Based Robot Movement Intention Signs for Human-Robot Interactions. In *2020 IEEE 20th International Symposium on Computational Intelligence and Informatics (CINTI)*. IEEE, 121–126. https://doi.org/10.1109/CINTI51262.2020.9305854

[36] • Anca D. Dragan, Shira Bauman, Jodi Forlizzi, and Siddhartha S. Srinivasa. 2015. Effects of Robot Motion on Human-Robot Collaboration. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, Julie A. Adams, William Smart, Bilge Mutlu, and Leila Takayama (Eds.). ACM, New York, NY, USA, 51–58. https://doi.org/10.1145/2696454.2696473

[37] • Anca D. Dragan and Siddhartha Srinivasa. 2013. Generating Legible Motion. In *Robotics: Science and Systems*.

[38] • Brittany A. Duncan, Evan Beachly, Alisha Bevins, Sebasitan Elbaum, and Carrick Detweiler. 2018. Investigation of Communicative Flight Paths for Small Unmanned Aerial Systems * This work was supported by NSF NRI 1638099. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 602–609. https://doi.org/10.1109/ICRA.2018.8462871

[39] • Miguel Faria, Andrea Costigliola, Patricia Alves-Oliveira, and Ana Paiva. 2016. Follow me: Communicating intentions with a spherical robot. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 664–669. https://doi.org/10.1109/ROMAN.2016.7745189

[40] • Miguel Faria, Francisco S. Melo, and Ana Paiva. 2021. Understanding Robots: Making Robots More Legible in Multi-Party Interactions. In *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*. IEEE, 1031–1036. https://doi.org/10.1109/RO-MAN50785.2021.9515485

[41] • Miguel Faria, Rui Silva, Patricia Alves-Oliveira, Francisco S. Melo, and Ana Paiva. 2017. "Me and you together" movement impact in multi-user collaboration tasks. In *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2793–2798. https://doi.org/10.1109/IROS.2017.8206109

[42] • Kerstin Fischer, Lars C. Jensen, Stefan-Daniel Suvei, and Leon Bodenhagen. 2016. Between legibility and contact: The role of gaze in robot approach. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 646–651. https://doi.org/10.1109/ROMAN.2016.7745186

[43] • Paul Fletcher, Angeline Luther, Brittany Duncan, and Carrick Detweiler. 2021. Investigation of Unmanned Aerial Vehicle Gesture Perceptibility and Impact of Viewpoint Variance *. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 3531–3537. https://doi.org/10.1109/ICRA48506.2021.9561094

[44] • Morten Roed Frederiksen and Kasper Stoey. 2019. Augmenting the audio-based expression modality of a non-affective robot. In *2019 8th International Conference on Affective Computing and Intelligent Interaction (ACII)*. IEEE, 144–149. https://doi.org/10.1109/ACII.2019.8925510

[45] • Michihiko Furuhashi, Tsuyoshi Nakamura, Masayoshi Kanoh, and Koji Yamada. 2015. Touch-based information transfer from a robot modeled on the hearing dog. In *2015 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*. IEEE, 1–6. https://doi.org/10.1109/FUZZ-IEEE.2015.7337981

[46] Rinat R. Galin and Roman V. Meshcheryakov. 2020. *Human-Robot Interaction Efficiency and Human-Robot Collaboration*. Springer International Publishing, Cham, 55–63. https://doi.org/10.1007/978-3-030-37841-7_5

[47] Rahul Gautam, Ankush Gedam, Ashish Zade, and Ajay Mahawadiwar. 2017. Review on development of industrial robotic arm. *International Research Journal of Engineering and Technology (IRJET)* 4, 03 (2017).

[48] Moojan Ghafurian, Jesse Hoey, and Kerstin Dautenhahn. 2021. Social Robots for the Care of Persons with Dementia: A Systematic Review. *J. Hum.-Robot Interact.* 10, 4, Article 41 (sep 2021), 31 pages. https://doi.org/10.1145/3469653

[49] • Michael J. Gielniak and Andrea L. Thomaz. 2011. Generating anticipation in robot motion. In *2011 RO-MAN*. IEEE, 449–454. https://doi.org/10.1109/ROMAN.2011.6005255

[50] Alexandre Girard, Jean-Philippe Lucking Bigué, Benjamin M. O'Brien, Todd A. Gisby, Iain A. Anderson, and Jean-Sébastien Plante. 2015. Soft Two-Degree-of-Freedom Dielectric Elastomer Position Sensor Exhibiting Linear Behavior. *IEEE/ASME Transactions on Mechatronics* 20, 1 (2015), 105–114. https://doi.org/10.1109/TMECH.2014.2307006

[51] • Dylan F. Glas, Takahiro Miyashita, Hiroshi Ishiguro, and Norihiro Hagita. 2007. Robopal: Modeling Role Transitions in Human-Robot Interaction. In *Proceedings 2007 IEEE International Conference on Robotics and Automation*. IEEE, 2130–2137. https://doi.org/10.1109/ROBOT.2007.363636

[52] Michael A. Goodrich and Alan C. Schultz. 2008. Human–Robot Interaction: A Survey. *Foundations and Trends® in Human–Computer Interaction* 1, 3 (2008), 203–275. https://doi.org/10.1561/1100000005

[53] Jonathan Grudin. 2017. From tool to partner: The evolution of human-computer interaction. *Synthesis Lectures on Human-Centered Interaction* 10, 1 (2017), i–183.

[54] • Uwe Gruenefeld, Lars Prädel, Jannike Illing, Tim Stratmann, Sandra Drolshagen, and Max Pfingsthorn. 2020. Mind the ARm. In *Proceedings of the Conference on Mensch und Computer*, Bernhard Preim, Andreas Nürnberger, and Christian Hansen (Eds.). ACM, New York, NY, USA, 259–266. https://doi.org/10.1145/3404983.3405509

[55] • Morris Gu, Akansel Cosgun, Wesley P. Chan, Tom Drummond, and Elizabeth Croft. 2021. Seeing Thru Walls: Visualizing Mobile Robots in Augmented Reality. In *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*. IEEE, 406–411. https://doi.org/10.1109/RO-MAN50785.2021.9515322

[56] Adriana Hamacher, Nadia Bianchi-Berthouze, Anthony G. Pipe, and Kerstin Eder. 2016. Believing in BERT: Using expressive communication to enhance trust and counteract operational error in physical Human-robot interaction. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. 493–500. https://doi.org/10.1109/ROMAN.2016.7745163

[57] • Zhao Han, Alexander Wilkinson, Jenna Parrillo, Jordan Allspaw, and Holly A. Yanco. 2020. Projection Mapping Implementation: Enabling Direct Externalization of Perception Results and Action Intent to Improve Robot Explainability. *Proceedings of the AI-HRI Symposium at AAAI-FSS 2020* (2020). https://doi.org/10.48550/arXiv.2010.02263

[58] • Jinying He, Anouk van Maris, and Praminda Caleb-Solly. 2020. Investigating the Effectiveness of Different Interaction Modalities for Spatial Human-robot Interaction. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, Tony Belpaeme, James Young, Hatice Gunes, and Laurel Riek (Eds.). ACM, New York, NY, USA, 239–241. https://doi.org/10.1145/3371382.3378273

[59] Abdelfetah Hentout, Mustapha Aouache, Abderraouf Maoudj, and Isma Akli. 2019. Human–robot interaction in industrial collaborative robotics: a literature review of the decade 2008–2017. *Advanced Robotics* 33, 15-16 (2019), 764–799. https://doi.org/10.1080/01691864.2019.1636714

[60] • Nicholas J. Hetherington, Elizabeth A. Croft, and H. MachielF. van der Loos. 2021. Hey Robot, Which Way Are You Going? Nonverbal Motion Legibility Cues for Human-Robot Spatial Interaction. *IEEE Robotics and Automation Letters* 6, 3 (2021), 5010–5015. https://doi.org/10.1109/LRA.2021.3068708

[61] • Rachel M. Holladay, Anca D. Dragan, and Siddhartha S. Srinivasa. 2014. Legible robot pointing. In *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 217–223. https://doi.org/10.1109/ROMAN.2014.6926256

[62] Siddarth Jain and Brenna Argall. 2019. Probabilistic Human Intent Recognition for Shared Autonomy in Assistive Robotics. *J. Hum.-Robot Interact.* 9, 1, Article 2 (dec 2019), 23 pages. https://doi.org/10.1145/3359614

[63] • Gunnar Johannsen. 2002. Auditory display of directions and states for mobile systems. In *Proceedings of the 2002 International Conference on Auditory Display*.

[64] • Doğancan Kebüde, Cem Eteke, Tevfik Metin Sezgin, and Barş Akgün. 2018. Communicative Cues for Reach-to-Grasp Motions: From Humans to Robots. In *Proceedings of the 17th International Conference on Autonomous Agents and MultiAgent Systems (AAMAS '18)*. International Foundation for Autonomous Agents and Multiagent Systems, Richland, SC, 874–882.

[65] Lawrence H. Kim, Daniel S. Drew, Veronika Domova, and Sean Follmer. 2020. User-Defined Swarm Robot Control. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376814

[66] • Nathan Kirchner, Alen Alempijevic, and Gamini Dissanayake. 2011. Nonverbal robot-group interaction using an imitated gaze cue. In *Proceedings of the 6th international conference on Human-robot interaction - HRI '11*, Aude Billard, Peter Kahn, Julie A. Adams, and Greg Trafton (Eds.). ACM Press, New York, New York, USA, 497. https://doi.org/10.1145/1957656.1957824

[67] • K. L. Koay, G. Lakatos, D. S. Syrdal, M. Gacsi, B. Bereczky, K. Dautenhahn, A. Miklosi, and M. L. Walters. 2013. Hey! There is someone at your door. A hearing robot using visual communication signals of hearing dogs to communicate intent. In *2013 IEEE Symposium on Artificial Life (ALife)*. IEEE, 90–97. https://doi.org/10.1109/ALIFE.2013.6602436

[68] Danica Kragic, Joakim Gustafson, Hakan Karaoguz, Patric Jensfelt, and Robert Krug. 2018. Interactive, Collaborative Robots: Challenges and Opportunities. In *Proceedings of the 27th International Joint Conference on Artificial Intelligence* (Stockholm, Sweden) *(IJCAI'18)*. AAAI Press, Palo Alto, California, USA, 18–25.

[69] Thibault Kruse, Amit Kumar Pandey, Rachid Alami, and Alexandra Kirsch. 2013. Human-aware robot navigation: A survey. *Robotics and Autonomous Systems* 61, 12 (2013), 1726–1743. https://doi.org/10.1016/j.robot.2013.05.007

[70] Ulyana Kurylo and Jason R. Wilson. 2019. Using Human Eye Gaze Patterns as Indicators of Need for Assistance from a Socially Assistive Robot. In *Social Robotics*, Miguel A. Salichs, Shuzhi Sam Ge, Emilia Ivanova Barakova, John-John Cabibihan, Alan R. Wagner, Álvaro Castro-González, and Hongsheng He (Eds.). Springer International Publishing, Cham, 200–210.

[71] • Mai Lee Chang, Reymundo A. Gutierrez, Priyanka Khante, Elaine Schaertl Short, and Andrea Lockerd Thomaz. 2018. Effects of Integrated Intent Recognition and Communication on Human-Robot Collaboration. In *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 3381–3386. https://doi.org/10.1109/IROS.2018.8593359

[72] • Gregory LeMasurier, Gal Bejerano, Victoria Albanese, Jenna Parrillo, Holly A. Yanco, Nicholas Amerson, Rebecca Hetrick, and Elizabeth Phillips. 2021. Methods for Expressing Robot Intent for Human–Robot Collaboration in Shared Workspaces. *ACM Transactions on Human-Robot Interaction* 10, 4 (2021), 1–27. https://doi.org/10.1145/3472223

[73] Ivan Lerner, Perrine Créquit, Philippe Ravaud, and Ignacio Atal. 2019. Automatic screening using word embeddings achieved high sensitivity and workload reduction for updating living network meta-analyses. *Journal of Clinical Epidemiology* 108 (April 2019), 86–94. https://doi.org/10.1016/j.jclinepi.2018.12.001

[74] • Florent Levillain, David St-Onge, Giovanni Beltrame, and Elisabetta Zibetti. 2019. Towards situational awareness from robotic group motion. In *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 1–6. https://doi.org/10.1109/RO-MAN46459.2019.8956381

[75] Kexi Liu, Daisuke Sakamoto, Masahiko Inami, and Takeo Igarashi. 2011. Roboshop: Multi-Layered Sketching Interface for Robot Housework Assignment and Management. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 647–656. https://doi.org/10.1145/1978942.1979035

[76] Hamza Mahdi, Sami Alperen Akgun, Shahed Saleh, and Kerstin Dautenhahn. 2022. A survey on the design and evolution of social robots — Past, present and future. *Robotics and Autonomous Systems* 156 (2022), 104193. https://doi.org/10.1016/j.robot.2022.104193

[77] Christopher D. Manning, Prabhakar Raghavan, and Hinrich Schütze. 2008. *Introduction to Information Retrieval -*. Cambridge University Press, Cambridge.

[78] Eloise Matheson, Riccardo Minto, Emanuele G. G. Zampieri, Maurizio Faccio, and Giulio Rosati. 2019. Human–Robot Collaboration in Manufacturing Applications: A Review. *Robotics* 8, 4 (2019). https://doi.org/10.3390/robotics8040100

[79] • Takafumi Matsumaru. 2006. Mobile Robot with Preliminary-announcement and Display Function of Forthcoming Motion using Projection Equipment. In *ROMAN 2006 - The 15th IEEE International Symposium on Robot and Human Interactive Communication*. IEEE, 443–450. https://doi.org/10.1109/ROMAN.2006.314368

[80] • Takafumi Matsumaru. 2007. Mobile Robot with Preliminary-announcement and Indication Function of Forthcoming Operation using Flat-panel Display. In *Proceedings 2007 IEEE International Conference on Robotics and Automation*. IEEE, 1774–1781. https://doi.org/10.1109/ROBOT.2007.363579

[81] • Takafumi Matsumaru, Kazuya Iwase, Kyouhei Akiyama, Takashi Kusada, and Tomotaka Ito. 2005. Mobile Robot with Eyeball Expression as the Preliminary-Announcement and Display of the Robot?s Following Motion. *Autonomous Robots* 18, 2 (2005), 231–246. https://doi.org/10.1007/s10514-005-0728-8

[82] • Takafumi Matsumaru, Takashi Kusada, and Kazuya Iwase. 2006. Mobile Robot with Preliminary-Announcement Function of Forthcoming Motion using Light-ray. In *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 1516–1523. https://doi.org/10.1109/IROS.2006.281981

[83] • Masahiko Mikawa, Yuriko Yoshikawa, and Makoto Fujisawa. 2018. Expression of intention by rotational head movements for teleoperated mobile robot. In *2018 IEEE 15th International Workshop on Advanced Motion Control (AMC)*. IEEE, 249–254. https://doi.org/10.1109/AMC.2019.8371097

[84] • AJung Moon, Daniel M. Troniak, Brian Gleeson, Matthew K.X.J. Pan, Minhua Zheng, Benjamin A. Blumer, Karon MacLean, and Elizabeth A. Croft. 2014. Meet me where i'm gazing. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, Gerhard Sagerer, Michita Imai, Tony Belpaeme, and Andrea Thomaz (Eds.). ACM, New York, NY, USA, 334–341. https://doi.org/10.1145/2559636.2559656

[85] John Edison Muñoz and Kerstin Dautenhahn. 2021. Robo Ludens: A Game Design Taxonomy for Multiplayer Games Using Socially Interactive Robots. *J. Hum.-Robot Interact.* 10, 4, Article 32 (jul 2021), 28 pages. https://doi.org/10.1145/3451343

[86] • James Mullen, Josh Mosier, Sounak Chakrabarti, Anqi Chen, Tyler White, and Dylan Losey. 2021. Communicating Inferred Goals With Passive Augmented Reality and Active Haptic Feedback. *IEEE Robotics and Automation Letters* 6, 4 (2021), 8522–8529. https://doi.org/10.1109/LRA.2021.3111055

[87] Florian Müller, Martin Schmitz, Daniel Schmitt, Sebastian Günther, Markus Funk, and Max Mühlhäuser. 2020. Walk The Line: Leveraging Lateral Shifts of the Walking Path as an Input Modality for Head-Mounted Displays. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–15. https://doi.org/10.1145/3313831.3376852

[88] • Bilge Mutlu, Fumitaka Yamaoka, Takayuki Kanda, Hiroshi Ishiguro, and Norihiro Hagita. 2009. Nonverbal leakage in robots. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction - HRI '09*, Matthias Scheutz, François Michaud, Pamela Hinds, and Brian Scassellati (Eds.). ACM Press, New York, New York, USA, 69. https://doi.org/10.1145/1514095.1514110

[89] • Rhys Newbury, Akansel Cosgun, Tysha Crowley-Davis, Wesley P. Chan, Tom Drummond, and Elizabeth Croft. [n.d.]. Visualizing Robot Intent for Object Handovers with Augmented Reality. https://doi.org/10.48550/arXiv.2103.04055

[90] Yohei Noguchi and Fumihide Tanaka. 2020. OMOY: A Handheld Robotic Gadget that Shifts its Weight to Express Emotions and Intentions. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. ACM. https://doi.org/10.1145/3313831.3376775

[91] • Michael Novitzky, Charles Pippin, Thomas R. Collins, Tucker R. Balch, and Michael E. West. 2012. Bio-inspired multi-robot communication through behavior recognition. In *2012 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 771–776. https://doi.org/10.1109/ROBIO.2012.6491061

[92] Alison O'Mara-Eves, James Thomas, John McNaught, Makoto Miwa, and Sophia Ananiadou. 2015. Using text mining for study identification in systematic reviews: a systematic review of current approaches. *Systematic Reviews* 4, 1 (Jan. 2015). https://doi.org/10.1186/2046-4053-4-5

[93] Linda Onnasch and Eileen Roesler. 2020. A Taxonomy to Structure and Analyze Human–Robot Interaction. *International Journal of Social Robotics* 13, 4 (June 2020), 833–849. https://doi.org/10.1007/s12369-020-00666-5

[94] Matthew J. Page, Joanne E. McKenzie, Patrick M. Bossuyt, Isabelle Boutron, Tammy C. Hoffmann, Cynthia D. Mulrow, Larissa Shamseer, Jennifer M. Tetzlaff, Elie A. Akl, Sue E. Brennan, Roger Chou, Julie Glanville, Jeremy M. Grimshaw, Asbjørn Hróbjartsson, Manoj M. Lalu, Tianjing Li, Elizabeth W. Loder, Evan Mayo-Wilson, Steve McDonald, Luke A. McGuinness, Lesley A. Stewart, James Thomas, Andrea C. Tricco, Vivian A. Welch, Penny Whiting, and David Moher. 2021. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Systematic Reviews* 10, 1 (March 2021). https://doi.org/10.1186/s13643-021-01626-4

[95] • Oskar Palinko, Kerstin Fischer, Eduardo Ruiz Ramirez, Lotte Damsgaard Nissen, and Rosalyn M. Langedijk. 2020. A Drink-Serving Mobile Social Robot Selects who to Interact with Using Gaze. In *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, Tony Belpaeme, James Young, Hatice Gunes, and Laurel Riek (Eds.). ACM, New York, NY, USA, 384–385. https://doi.org/10.1145/3371382.3378339

[96] Max Pascher, Annalies Baumeister, Stefan Schneegass, Barbara Klein, and Jens Gerken. 2021. Recommendations for the Development of a Robotic Drinking and Eating Aid - An Ethnographic Study. In *Human-Computer Interaction – INTERACT 2021*, Carmelo Ardito, Rosa Lanzilotti, Alessio Malizia, Helen Petrie, Antonio Piccinno, Giuseppe Desolda, and Kori Inkpen (Eds.). Springer, Cham, Switzerland, 331–351. https://doi.org/10.1007/978-3-030-85623-6_21

[97] Max Pascher, Stefan Schneegass, and Jens Gerken. 2019. SwipeBuddy: A Teleoperated Tablet and Ebook-Reader Holder for a Hands-Free Interaction. In *Human-Computer Interaction – INTERACT 2019*, David Lamas, Fernando Loizides, Lennart Nacke, Helen Petrie, Marco Winckler, and Panayiotis Zaphiris (Eds.), Vol. 11749. Springer, Cham, Switzerland, 568–571. https://doi.org/10.1007/978-3-030-29390-1_39

[98] • Loizos Psarakis, Dimitris Nathanael, and Nicolas Marmaras. 2022. Fostering short-term human anticipatory behavior in human-robot collaboration. *International Journal of Industrial Ergonomics* 87 (2022), 103241. https://doi.org/10.1016/j.ergon.2021.103241

[99] • Eric Rosen, David Whitney, Elizabeth Phillips, Gary Chien, James Tompkin, George Konidaris, and Stefanie Tellex. 2019. Communicating and controlling robot arm motion intent through mixed-reality head-mounted displays. *The International Journal of Robotics Research* 38, 12-13 (2019), 1513–1526. https://doi.org/10.1177/0278364919842925

[100] Alexandros Rouchitsas and Håkan Alm. 2019. External Human–Machine Interfaces for Autonomous Vehicle-to-Pedestrian Communication: A Review of Empirical Work. *Frontiers in Psychology* 10 (Dec. 2019). https://doi.org/10.3389/fpsyg.2019.02757

[101] • Emanuele Ruffaldi, Filippo Brizzi, Franco Tecchia, and Sandro Bacinelli. 2016. Third Point of View Augmented Reality for Robot Intentions Visualization. In *Augmented Reality, Virtual Reality, and Computer Graphics*, Lucio Tommaso de Paolis and Antonio Mongelli (Eds.). Lecture Notes in Computer Science, Vol. 9768. Springer International Publishing, Cham, 471–478. https://doi.org/10.1007/978-3-319-40621-3{_}35

[102] Gerard Salton and Christopher Buckley. 1988. Term-weighting approaches in automatic text retrieval. *Information Processing & Management* 24, 5 (Jan. 1988), 513–523. https://doi.org/10.1016/0306-4573(88)90021-0

[103] Albrecht Schmidt, Michael Beigl, and Hans-W Gellersen. 1999. There is more to context than location. *Computers & Graphics* 23, 6 (1999), 893–901. https://doi.org/10.1016/S0097-8493(99)00120-X

[104] Lambert Schomaker. 1995. A taxonomy of multimodal interaction in the human information processing system. (1995).

[105] • Megha Sharma, Dale Hildebrandt, Gem Newman, James E. Young, and Rasit Eskicioglu. 2013. Communicating affect via flight path Exploring use of the Laban Effort System for designing affective locomotion paths. In *2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 293–300. https://doi.org/10.1109/HRI.2013.6483602

[106] Ben Shneiderman. 2022. *Human-Centered AI*. Oxford University Press.

[107] Rajinder Sodhi, Hrvoje Benko, and Andrew Wilson. 2012. LightGuide: Projected Visualizations for Hand Movement Guidance. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 179–188. https://doi.org/10.1145/2207676.2207702

[108] • Sichao Song and Seiji Yamada. 2018. Bioluminescence-Inspired Human-Robot Interaction. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, Takayuki Kanda, Selma Ŝabanović, Guy Hoffman, and Adriana Tapus (Eds.). ACM, New York, NY, USA, 224–232. https://doi.org/10.1145/3171221.3171249

[109] • Sichao Song and Seiji Yamada. 2018. Designing LED Lights for Communicating Gaze with Appearance-Constrained Robots. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 94–97. https://doi.org/10.1109/ROMAN.2018.8525661

[110] • Sichao Song and Seiji Yamada. 2018. Effect of Expressive Lights on Human Perception and Interpretation of Functional Robot. In *Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems*, Regan Mandryk, Mark Hancock, Mark Perry, and Anna Cox (Eds.). ACM, New York, NY, USA, 1–6. https://doi.org/10.1145/3170427.3188547

[111] • Maria Staudte and Matthew W. Crocker. 2009. Visual attention in spoken human-robot interaction. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction - HRI '09*, Matthias Scheutz, François Michaud, Pamela Hinds, and Brian Scassellati (Eds.). ACM Press, New York, New York, USA, 77. https://doi.org/10.1145/1514095.1514111

[112] Didi Surian, Florence T. Bourgeois, and Adam G. Dunn. 2021. The automation of relevant trial registration screening for systematic review updates: an evaluation study on a large dataset of ClinicalTrials.gov registrations. *BMC Medical Research Methodology* 21, 1 (Dec. 2021). https://doi.org/10.1186/s12874-021-01485-6

[113] Ryo Suzuki, Adnan Karim, Tian Xia, Hooman Hedayati, and Nicolai Marquardt. 2022. Augmented Reality and Robotics: A Survey and Taxonomy for AR-Enhanced Human-Robot Interaction and Robotic Interfaces. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 553, 33 pages. https://doi.org/10.1145/3491102.3517719

[114] • Daniel Szafir, Bilge Mutlu, and Terrence Fong. 2014. Communication of intent in assistive free flyers. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*, Gerhard Sagerer, Michita Imai, Tony Belpaeme, and Andrea Thomaz (Eds.). ACM, New York, NY, USA, 358–365. https://doi.org/10.1145/2559636.2559672

[115] • Daniel Szafir, Bilge Mutlu, and Terry Fong. 2015. Communicating Directionality in Flying Robots. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, Julie A. Adams, William Smart, Bilge Mutlu, and Leila Takayama (Eds.). ACM, New York, NY, USA, 19–26. https://doi.org/10.1145/2696454.2696475

[116] • Leila Takayama, Doug Dooley, and Wendy Ju. 2011. Expressing thought. In *Proceedings of the 6th international conference on Human-robot interaction - HRI '11*, Aude Billard, Peter Kahn, Julie A. Adams, and Greg Trafton (Eds.). ACM Press, New York, New York, USA, 69. https://doi.org/10.1145/1957656.1957674

[117] • Gilbert Tang, Phil Webb, and John Thrower. 2019. The development and evaluation of Robot Light Skin: A novel robot signalling system to improve communication in industrial human–robot collaboration. *Robotics and Computer-Integrated Manufacturing* 56 (2019), 85–94. https://doi.org/10.1016/j.rcim.2018.08.005

[118] Sam Thellman and Tom Ziemke. 2021. The Perceptual Belief Problem. *ACM Transactions on Human-Robot Interaction* 10, 3 (2021), 1–15. https://doi.org/10.1145/3461781

[119] Andrea C. Tricco, Erin Lillie, Wasifa Zarin, Kelly K. O'Brien, Heather Colquhoun, Danielle Levac, David Moher, Micah D.J. Peters, Tanya Horsley, Laura Weeks, Susanne Hempel, Elie A. Akl, Christine Chang, Jessie McGowan, Lesley Stewart, Lisa Hartling, Adrian Aldcroft, Michael G. Wilson, Chantelle Garritty, Simon Lewin, Christina M. Godfrey, Marilyn T. Macdonald, Etienne V. Langlois, Karla Soares-Weiser, Jo Moriarty, Tammy Clifford, Özge Tunçalp, and Sharon E. Straus. 2018. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Annals of Internal Medicine* 169, 7 (Oct. 2018), 467–473. https://doi.org/10.7326/m18-0850

[120] • Georgios Tsamis, Georgios Chantziaras, Dimitrios Giakoumis, Ioannis Kostavelis, Andreas Kargakos, Athanasios Tsakiris, and Dimitrios Tzovaras. 2021. Intuitive and Safe Interaction in Multi-User Human Robot Collaboration Environments through Augmented Reality Displays. In *2021 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN)*. IEEE, 520–526. https://doi.org/10.1109/RO-MAN50785.2021.9515474

[121] Sanne van Waveren, Linnéa Björklund, Elizabeth J. Carter, and Iolanda Leite. 2019. Knock on Wood: The Effects of Material Choice on the Perception of Social Robots. In *Social Robotics*, Miguel A. Salichs, Shuzhi Sam Ge, Emilia Ivanova Barakova, John-John Cabibihan, Alan R. Wagner, Álvaro Castro-González, and Hongsheng He (Eds.). Springer International Publishing, Cham, 211–221.

[122] Ana M Villanueva, Ziyi Liu, Zhengzhe Zhu, Xin Du, Joey Huang, Kylie A Peppler, and Karthik Ramani. 2021. RobotAR: An Augmented Reality Compatible Teleconsulting Robotics Toolkit for Augmented Makerspace Experiences. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems* (Yokohama, Japan) *(CHI '21)*. Association for Computing Machinery, New York, NY, USA, Article 477, 13 pages. https://doi.org/10.1145/3411764.3445726

[123] Erik von Elm, Gerhard Schreiber, and Claudia Cornelia Haupt. 2019. Methodische Anleitung für Scoping Reviews (JBI-Methodologie). *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen* 143 (June 2019), 1–7. https://doi.org/10.1016/j.zefq.2019.05.004

[124] • Michael Walker, Hooman Hedayati, Jennifer Lee, and Daniel Szafir. 2018. Communicating Robot Motion Intent with Augmented Reality. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, Takayuki Kanda, Selma Ŝabanović, Guy Hoffman, and Adriana Tapus (Eds.). ACM, New York, NY, USA, 316–324. https://doi.org/10.1145/3171221.3171253

[125] Sebastian Wallkötter, Silvia Tulli, Ginevra Castellano, Ana Paiva, and Mohamed Chetouani. 2021. Explainable Embodied Agents Through Social Cues: A Review. *J. Hum.-Robot Interact.* 10, 3, Article 27 (jul 2021), 24 pages. https://doi.org/10.1145/3457188

[126] Lihui Wang, Robert Gao, Jozsef Váncza, Jörg Krüger, Xi Vincent Wang, Sotiris Makris, and George Chryssolouris. 2019. Symbiotic human-robot collaborative assembly. *CIRP Annals* 68, 2 (2019), 701–726. https://doi.org/10.1016/j.cirp.2019.05.002

[127] • Atsushi Watanabe, Tetsushi Ikeda, Yoichi Morales, Kazuhiko Shinozawa, Takahiro Miyashita, and Norihiro Hagita. 2015. Communicating robotic navigational intentions. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 5763–5769. https://doi.org/10.1109/IROS.2015.7354195

[128] • Tim Wengefeld, Dominik Hochemer, Benjamin Lewandowski, Mona Kohler, Manuel Beer, and Horst-Michael Gross. 2020. A Laser Projection System for Robot Intention Communication and Human Robot Interaction. In *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 259–265. https://doi.org/10.1109/RO-MAN47096.2020.9223517

[129] Jacob O. Wobbrock and Julie A. Kientz. 2016. Research Contributions in Human-Computer Interaction. *Interactions* 23, 3 (apr 2016), 38–44. https://doi.org/10.1145/2907069

[130] • Lisa Zahray, Richard Savery, Liana Syrkett, and Gil Weinberg. 2020. Robot Gesture Sonification to Enhance Awareness of Robot Status and Enjoyment of Interaction. In *2020 29th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 978–985. https://doi.org/10.1109/RO-MAN47096.2020.9223452

[131] • Allan Zhou, Dylan Hadfield-Menell, Anusha Nagabandi, and Anca D. Dragan. 2017. Expressive Robot Motion Timing. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*, Bilge Mutlu, Manfred Tscheligi, Astrid Weiss, and James E. Young (Eds.). ACM, New York, NY, USA, 22–31. https://doi.org/10.1145/2909824.3020221

[132] • Mark Zolotas and Yiannis Demiris. 2019. Towards Explainable Shared Control using Augmented Reality. In *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 3020–3026. https://doi.org/10.1109/IROS40897.2019.8968117

Adaptive DoF: Concepts to Visualize AI-generated Movements in Human-Robot Collaboration

Max Pascher
max.pascher@w-hs.de
Westphalian University of Applied Sciences
Gelsenkirchen, Germany

Kirill Kronhardt
kirill.kronhardt@studmail.w-hs.de
Westphalian University of Applied Sciences
Gelsenkirchen, Germany

Til Franzen
til.franzen@studmail.w-hs.de
Westphalian University of Applied Sciences
Gelsenkirchen, Germany

Jens Gerken
jens.gerken@w-hs.de
Westphalian University of Applied Sciences
Gelsenkirchen, Germany

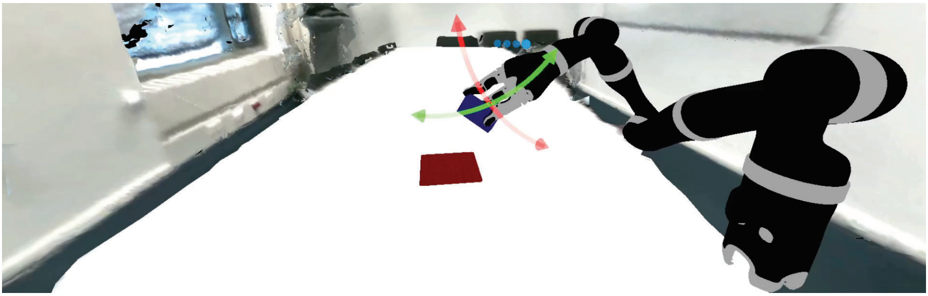


Figure 1: Communicating Cobot's Motion Intent Feedback via Gizmo Approach

ABSTRACT

Nowadays, robots collaborate closely with humans in a growing number of areas. Enabled by lightweight materials and safety sensors, these cobots are gaining increasing popularity in domestic care, supporting people with physical impairments in their everyday lives. However, when cobots perform actions autonomously, it remains challenging for human collaborators to understand and predict their behavior. This, however, is crucial for achieving trust and user acceptance. One significant aspect of predicting cobot behavior is understanding their motion intent and comprehending how they "think" about their actions. We work on solutions that communicate the cobots AI-generated motion intent to a human collaborator. Effective communication enables users to proceed with the most suitable option. We present a design exploration with different visualization techniques to optimize this user understanding, ideally resulting in increased safety and end-user acceptance.

AVI 2022, June 6–10, 2022, Frascati, Rome, Italy

ACM ISBN 978-1-4503-9719-3/22/06.
https://doi.org/10.1145/3531073.3534479

CCS CONCEPTS

• **Computer systems organization** → *Robotic control*; *Robotic autonomy*; • **Human-centered computing** → **Visualization techniques**.

KEYWORDS

cobot, human-robot collaboration, visualization techniques, neural network, intention feedback

ACM Reference Format:
Max Pascher, Kirill Kronhardt, Til Franzen, and Jens Gerken. 2022. Adaptive DoF: Concepts to Visualize AI-generated Movements in Human-Robot Collaboration. In *Proceedings of the 2022 International Conference on Advanced Visual Interfaces (AVI 2022), June 6–10, 2022, Frascati, Rome, Italy*. ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3531073.3534479

1 INTRODUCTION

Robotic solutions are becoming increasingly prevalent in our personal and professional lives, and have started to evolve into close collaborators [3, 7, 10]. These so-called cobots support humans in various ways that were unimaginable just a few years ago. Enabled by technological advances, newer lightweight materials, and improved safety sensors, they are gaining increasing popularity in domestic care, supporting people with disabilities in their everyday lives [11].

However, new potential issues arise when cobots are tasked with (semi-)autonomous actions, resulting in added stress for end-users [13]. Particularly close proximity collaboration between humans and cobots remains challenging [8]. These challenges include effective communication to the end-user of (a) motion intent and (b) the spatial perception of the cobot's vicinity [12].

2 RELATED WORK

In recent years, Augmented Reality (AR) technology has been frequently used for human-robot collaboration [2, 6]. Previous work focused primarily on the use of Head-Mounded Displays, Mobile Augmented Reality, and Spatial Augmented Reality for the visualization of cobot motion intent [8, 14, 16]. Rosen et al. showed that AR is an improvement compared to traditional desktop interfaces when visualizing the intended motion of robots [14]. Previous literature has focused mainly on visualizations of motion intent for autonomous robotic systems [1, 4, 5, 8, 15, 17], communicating recommended cobot intention and its control methods has however not attracted as much attention.

3 TESTBED ENVIRONMENT

In earlier work, we developed an adaptive control interaction method based on a recommendation system generated by a Convolutional Neural Network [9]. From the cobot's seven Degrees of Freedom (DoF), the adaptive control combined several DoFs to provide a more straightforward control to the user with fewer necessary mode-switches.

The virtual environment, including a virtual model of the *Kinova Jaco*[1] robot arm was developed to be compatible with the *Oculus Quest 2*[2] VR headset (see Figure 1). This provided us with a VR testbed environment for developing and evaluating further feedback techniques.

4 VISUALIZATION CONCEPTS

Our proposed concepts fall into a spectrum with two extremes — indicative and explanatory. **Indicative:** Focus on crucial information only, quick and easy solution, suitable for experienced cobot users. **Explanatory:** Movements are shown in great detail, high level of information, especially helpful for new users.

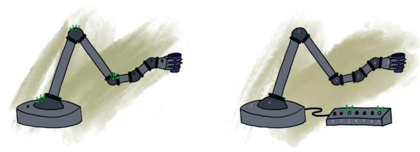
DoF-Indicator: LEDs attached to the cobot's axis and joints - or mounted on a bar in front of it - communicate active and nonactive DoFs (see Figure 2). Likely more suitable for experienced users, allows understanding of current DoF mapping by the recommendation system plus resulting movement abilities.

DoF-Combination-Indicator: Movement ability is communicated by a simplified representation of the cobot only showing two modalities, e. g. rotating and extending (see Figure 3). The AR representation (aka "fake joint") either overlays the real cobot or can be displayed separately in the corner of the AR screen.

Gizmo Visualisation: Arrows, planes and point clouds communicate the current movement ability of the cobot (see Figure 4). This allows for several different design options. A first arrow-based approach was already successfully evaluated in a previous study [9].

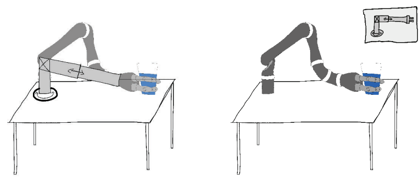
[1]Kinova Jaco robot arm: https://assistive.kinovarobotics.com/product/jaco-robotic-arm, last retrieved April 29, 2022

[2]Oculus Quest 2: https://www.oculus.com/quest-2/, last retrieved April 29, 2022



(a) (b)

Figure 2: DoF-Indicator: (a) LEDs attached to the cobot; (b) LEDs mounted on a bar.



(a) (b)

Figure 3: DoF-Combination-Indicator: (a) as an AR overlay; (b) as an icon in the screen corner.

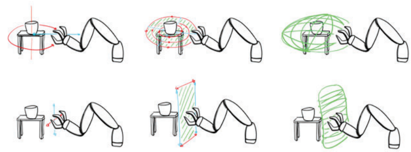


Figure 4: Gizmo Visualization: (left) simple: straight and curved arrows; (center) planar: planes of movement; (right) cloud: 3D-cloud of possible boundary positions.

Demonstration: Current movement possibilities are demonstrated through either the actual cobot or an AR representation. With both options a quick movement indicates the intended motion.

Future work will see the implementation of the various visualization options. Through this, we expect to gain a number of valuable insights regarding the explainability of AI behavior in the context of robotic movements.

REFERENCES

[1] Rasmus S. Andersen, Ole Madsen, Thomas B. Moeslund, and Heni Ben Amor. 26.08.2016 - 31.08.2016. Projecting robot intentions into human environments. In *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 294–301. https://doi.org/10.1109/ROMAN.2016.7745145

[2] Stephanie Arévalo-Arboleda, Franziska Ruecker, Tim Dierks, and Jens Gerken. 2021. Assisting Manipulation and Grasping in Robot Teleoperation with Augmented Reality Visual Cues. In *CHI Conference on Human Factors in Computing Systems (CHI '21)* (2021-05-03). ACM. https://doi.org/10.1145/3411764.3445398

[3] Andrea Bauer, Dirk Wollherr, and Martin Buss. 2008. Human–Robot Collaboration: A Survey. *International Journal of Humanoid Robotics* 05, 01 (2008), 47–66. https://doi.org/10.1142/S0219843608001303

[4] Ravi Teja Chadalavada, Henrik Andreasson, Robert Krug, and Achim J. Lilienthal. 02.09.2015 - 04.09.2015. That's on my mind! robot to human intention communication through on-board projection on shared floor space. In *2015 European Conference on Mobile Robots (ECMR)*. IEEE, 1–6. https://doi.org/10.1109/ECMR.2015.7403771

[5] Michael D. Coovert, Tiffany Lee, Ivan Shindev, and Yu Sun. 2014. Spatial augmented reality as a method for a mobile robot to communicate intended movement. *Computers in Human Behavior* 34 (2014), 241–248. https://doi.org/10.1016/j.chb.2014.02.001

[6] Morteza Dianatfar, Jyrki Latokartano, and Minna Lanz. 2021. Review on existing VR/AR solutions in human–robot collaboration. *Procedia CIRP* 97 (2021), 407–411. https://doi.org/10.1016/j.procir.2020.05.259 8th CIRP Conference of Assembly Technology and Systems.

[7] Terrence Fong, Illah Nourbakhsh, and Kerstin Dautenhahn. 2003. A survey of socially interactive robots. *Robotics and Autonomous Systems* 42, 3 (2003), 143–166. https://doi.org/10.1016/S0921-8890(02)00372-X Socially Interactive Robots.

[8] Uwe Gruenefeld, Lars Prädel, Jannike Illing, Tim Stratmann, Sandra Drolshagen, and Max Pfingsthorn. 2020. Mind the ARm: Realtime Visualization of Robot Motion Intent in Head-Mounted Augmented Reality. In *Proceedings of the Conference on Mensch Und Computer* (Magdeburg, Germany) *(MuC '20)*. Association for Computing Machinery, New York, NY, USA, 259–266. https://doi.org/10.1145/3404983.3405509

[9] Kirill Kronhardt, Stephan Rübner, Max Pascher, Felix Goldau, Udo Frese, and Jens Gerken. 2022. Adapt or Perish? Exploring the Effectiveness of Adaptive DoF Control Interaction Methods for Assistive Robot Arms. *Technologies* 10, 1 (2022). https://doi.org/10.3390/technologies10010030

[10] Max Pascher, Annalies Baumeister, Barbara Klein, Stefan Schneegass, and Jens Gerken. 2019. Little Helper: A Multi-Robot System in Home Health Care Environments. In *Proceedings of the 2019 International workshop on Human-Drone Interaction (iHDI) as part of the ACM Conference on Human Factors in Computing Systems* (2019-05-04). ACM. https://hal.archives-ouvertes.fr/hal-02128382

[11] Max Pascher, Annalies Baumeister, Stefan Schneegass, Barbara Klein, and Jens Gerken. 2021. Recommendations for the Development of a Robotic Drinking and Eating Aid - An Ethnographic Study. In *Human-Computer Interaction – INTERACT 2021* (2021-09-01), Carmelo Ardito, Rosa Lanzilotti, Alessio Malizia, Helen Petrie, Antonio Piccinno, Giuseppe Desolda, and Kori Inkpen (Eds.). Springer, Cham. https://doi.org/10.1007/978-3-030-85623-6_21

[12] Max Pascher, Kirill Kronhardt, Til Franzen, Uwe Gruenefeld, Stefan Schneegass, and Jens Gerken. 2022. My Caregiver the Cobot: Comparing Visualization Techniques to Effectively Communicate Cobot Perception to People with Physical Impairments. *Sensors* 22, 3 (2022). https://doi.org/10.3390/s22030755

[13] Anita Pollak, Mateusz Paliga, Matias M. Pulopulos, Barbara Kozusznik, and Malgorzata W. Kozusznik. 2020. Stress in manual and autonomous modes of collaboration with a cobot. *Computers in Human Behavior* 112 (2020), 106469. https://doi.org/10.1016/j.chb.2020.106469

[14] Eric Rosen, David Whitney, Elizabeth Phillips, Gary Chien, James Tompkin, George Konidaris, and Stefanie Tellex. 2019. Communicating and controlling robot arm motion intent through mixed-reality head-mounted displays. *The International Journal of Robotics Research* 38, 12-13 (2019), 1513–1526. https://doi.org/10.1177/0278364919842925

[15] Freek Stulp, Jonathan Grizou, Baptiste Busch, and Manuel Lopes. 28.09.2015 - 02.10.2015. Facilitating intention prediction for humans by optimizing robot motions. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 1249–1255. https://doi.org/10.1109/IROS.2015.7353529

[16] Michael Walker, Hooman Hedayati, Jennifer Lee, and Daniel Szafir. 2018. Communicating Robot Motion Intent with Augmented Reality. In *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction - HRI '18*, Takayuki Kanda, Selma Ŝabanović, Guy Hoffman, and Adriana Tapus (Eds.). ACM Press, New York, New York, USA, 316–324. https://doi.org/10.1145/3171221.3171253

[17] Atsushi Watanabe, Tetsushi Ikeda, Yoichi Morales, Kazuhiko Shinozawa, Takahiro Miyashita, and Norihiro Hagita. 28.09.2015 - 02.10.2015. Communicating robotic navigational intentions. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 5763–5769. https://doi.org/10.1109/IROS.2015.7354195

sensors

MDPI

Article

My Caregiver the Cobot: Comparing Visualization Techniques to Effectively Communicate Cobot Perception to People with Physical Impairments

Max Pascher [1,2,*], Kirill Kronhardt [1], Til Franzen [1], Uwe Gruenefeld [2], Stefan Schneegass [2] and Jens Gerken [1]

1 Human-Computer Interaction Group, Department of Media Informatics and Communication, Westphalian University of Applied Sciences, 45897 Gelsenkirchen, Germany; kirill.kronhardt@studmail.w-hs.de (K.K.); til.franzen@studmail.w-hs.de (T.F.); jens.gerken@w-hs.de (J.G.)

2 Human-Computer Interaction Group, Paluno—The Ruhr Institute for Software Technology, Faculty of Business Administration and Economics, University of Duisburg-Essen, 45127 Essen, Germany; uwe.gruenefeld@uni-due.de (U.G.); stefan.schneegass@uni-due.de (S.S.)

* Correspondence: max.pascher@w-hs.de

Abstract: Nowadays, robots are found in a growing number of areas where they collaborate closely with humans. Enabled by lightweight materials and safety sensors, these cobots are gaining increasing popularity in domestic care, where they support people with physical impairments in their everyday lives. However, when cobots perform actions autonomously, it remains challenging for human collaborators to understand and predict their behavior, which is crucial for achieving trust and user acceptance. One significant aspect of predicting cobot behavior is understanding their perception and comprehending how they "see" the world. To tackle this challenge, we compared three different visualization techniques for Spatial Augmented Reality. All of these communicate cobot perception by visually indicating which objects in the cobot's surrounding have been identified by their sensors. We compared the well-established visualizations *Wedge* and *Halo* against our proposed visualization *Line* in a remote user experiment with participants suffering from physical impairments. In a second remote experiment, we validated these findings with a broader non-specific user base. Our findings show that *Line*, a lower complexity visualization, results in significantly faster reaction times compared to *Halo*, and lower task load compared to both *Wedge* and *Halo*. Overall, users prefer *Line* as a more straightforward visualization. In Spatial Augmented Reality, with its known disadvantage of limited projection area size, established off-screen visualizations are not effective in communicating cobot perception and *Line* presents an easy-to-understand alternative.

Keywords: cobot; human–robot collaboration; visualization techniques; projection; virtual reality

check for updates

Citation: Pascher, M.; Kronhardt, K.; Franzen, T.; Gruenefeld, U.; Schneegass, S.; Gerken, J. My Caregiver the Cobot: Comparing Visualization Techniques to Effectively Communicate Cobot Perception to People with Physical Impairments. *Sensors* **2022**, *22*, 755. https://doi.org/10.3390/s22030755

Academic Editors: Abolfazl Zaraki and Hamed Rahimi Nohooji

Received: 30 December 2021
Accepted: 15 January 2022
Published: 19 January 2022

1. Introduction

While robots were previously taught to perform simple repetitive tasks, they have started to evolve into collaborators in our professional and personal lives [1,2]. As a result, these so-called cobots support humans in various ways that were unimaginable just a few years ago. One area that has seen drastic advances in human–robot collaboration is domestic care, with cobots supporting people with physical impairments [3]. These assist people in various ways [4], from activities of daily living (ADLs), including basic tasks such as drinking, eating, and grooming, to leisure-time activities [5,6]. In domestic care, cobots reduce the need for the constant presence of caregivers, empowering people previously reliant on others for help to regain their independence. Our previous research on the needs of people with physical impairments showed a strong desire for privacy and alone time, which can undoubtedly be achieved with reliable robotic support [7].

However, new challenges arise when cobots are tasked with autonomous or semi-autonomous actions, resulting in additional stress for end-users [8]. Close proximity

collaboration between humans and cobots remains particularly challenging [9]. These challenges include effective communication to the end-user of (a) motion intent and (b) the spatial perception of the cobot's vicinity [10]. Accurate communication increases our understanding of the cobot while avoiding the unpredictability regarding impending steps, motions, and sensed environment parameters. While visualizations of motion intent have been extensively studied [9–14], communicating cobot perception has received less attention [15]. We define cobot perception as the sensory information acquired to computationally understand the surroundings, including the detection and identification of objects of interest in the physical vicinity. In our work, we communicate these sensory information acquired by the cobot using three different visualization techniques. Users benefit from receiving information about and understanding a cobot's spatial perception as perception failures, including errors in computer vision and object perception, can occur. Without communication, these are otherwise difficult to predict and to understand [16,17]. Accordingly, there is a clear need to accurately express cobot perception to their human collaborators to improve the correct prediction of the cobot behavior [18].

Augmented reality (AR) technology is a promising medium to communicate cobot perception, with the possibility to directly show relevant perceptual information in the user's line-of-sight whilst linking 3D with the physical world. In previous work, AR technology has shown encouraging results for the visualization of motion intent [9,12]. Any visualization technique aiding users in understanding the cobot's perception of its surroundings needs to effectively communicate all objects both within the visible area of the user and outside (or "off-screen"). The off-screen area is defined by the field of view of the user but, more importantly, limited by the means of the AR systems spatial visualization capabilities.

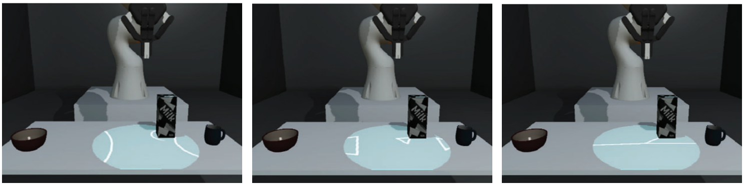
The release of the first Microsoft HoloLens resulted in an increased focus on approaches relying on Head-Mounted Displays (HMDs) [9]. However, even state-of-the-art HMD-AR such as the Microsoft HoloLens 2 (https://www.microsoft.com/de-de/hololens, last retrieved 30 December 2021) have a restricted display area which limits the field of view of the user [19]. Recent studies on the design preferences of people living with physical impairments also revealed that these displays are often impractical or not usable at all for the target population [20]. In addition, HMDs prevent direct information exchange with secondary users such as caregivers, thereby excluding them from providing necessary support. Similarly to HMDs, approaches using Mobile-AR (MAR) also limit the field of view through their display size and orientation, rendering them potentially unusable for people with physical impairments [21].

Spatial Augmented Reality (SAR) is another approach using projection techniques to augment the surface in the environment [22]. While essentially limited to 2D, research for motion intent has shown that SAR can be adapted to cover a dynamic workspace that encompasses multiple surface areas [10,11], e.g., in our scenario, this refers to interacting with objects "on a table" and "retrieving objects from shelves". Due to the significant decreases in the cost of projection technology and advances in pico-sized projectors, SAR has garnered increased interest in recent years [23]. SAR can augment larger areas of the surroundings, exceeding even the physical field of view of the user, and unlike HMDs, can be observed by secondary users. However, the possible field of view depends on the mounting position of the projection technique. While SAR may increase the visible augmentation area, the problem for effectively communicating off-screen objects still exists and is currently unsolved.

We investigated the potential visualization approaches that communicate the cobot's perception and particularly the information about detected objects in its physical surroundings. Information about physical objects are critical, as any breakdown in the successful detection of such objects by a (semi-) autonomous cobot can result in errors in behavior with the potential to harm the user, such as knocking over objects or even destroying them in the process. We applied this scenario to a breakfast situation in which a cobot supports a

person with physical impairments in performing basic tasks such as picking up a bottle and pouring a glass of water.

We used two established off-screen visualization techniques from research on small screens, namely *Wegde* [24] and *Halo* [25], where off-screen guidance is a well-explored topic. We added a third visualization technique *Line* (see Figure 1), which aims to reduce the potential drawback of visual clutter of *Wegde* and *Halo* as well as reduce the level of detail encoded in the visualization. All three approaches are used to continuously communicate the position of each object as they are perceived by the cobot. The user should immediately recognize a failure in object detection as the visualization for the lost object ceases.



(**a**) (**b**) (**c**)

Figure 1. The compared visualization techniques to communicate the cobot's perception are (**a**) *Halo*; (**b**) *Wedge*; and (**c**) *Line*.

We conducted two remote user studies exploring the efficiency, effectiveness, and task load of all three off-screen visualizations when communicating robot perception. First, we provided an exploratory experiment with 12 participants from our target user group of people with physical impairments. Second, we followed up with a validation experiment with 116 participants without physical impairments. Both studies show that a simple but clear visualization approach such as *Line* provides advantages for robot perception communication, both in terms of user preferences as well as objective measures. The remote nature of our study was adapted (a) to accommodate for social distancing guidelines during the SARS-CoV-2 pandemic; and (b) to allow for a more controlled and risk-reduced setup for target group participants.

2. Related Work

Previous literature has focused on (a) the usage of cobots for care support; (b) AR in human–robot collaborations; and (c) visualization techniques for target localization. We focus on ways cobots can effectively communicate their perception.

2.1. Cobots for Care

In 2021, the World Health Organization estimated that 15% of people live with some form of disability. (WHO. Disability & Health Report. https://www.who.int/news-room/fact-sheets/detail/disability-and-health, last retrieved 30 December 2021). Building on this, 7.9 million people are classed as severely disabled in Germany alone. (DESTATIS. Disability Facts and Figures—Brief Report 2019. https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Behinderte-Menschen/Publikationen/Downloads-Behinderte-Menschen/sozial-schwerbehinderte-kb-5227101199004.html, last retrieved 30 December 2021). Over 58% of these cases cover people with physical impairments and therefore we focused on this group for our study. In particular, we concentrated on people with a permanent and significant degree of compromised mobility of the extremities. Ample literature has examined the impact of assistive robotic systems in supporting people with motor impairments. The works of Chen et al. [5] for the *Robots for Humanity* project and Fattal et al. [6] looked into the feasibility and acceptance of robotic systems as assistive technologies. Both found that robotic devices are often designed to assist with several different activities of daily living, often resulting in larger robotic devices that frequently require a robotic arm mounted on a mobile unit. Drolshagen et al. investigated the acceptance of robots in sheltered workshops,

finding that robots are quickly accepted and close proximity is preferred [26]. Currently, a trend can be observed towards the research into cobots in domestic care to support people in their everyday lives [27–32]. One elementary part of everyday tasks is the consumption of food and drinks [7]. Based on this, the present study investigates cobot assistance for people with physical impairments during a standard breakfast scenario.

2.2. Augmented Reality in Human–Robot Collaboration

In recent decades, AR technology has been frequently used for human–robot collaboration [33]. Previous work has mainly focused on the use of HMDs, MAR, SAR and the visualization of the robot motion intent [9,34,35]. Rosen et al. showed that AR is an improvement compared to classical desktop interfaces when visualizing the intended motion of robots [34]. However, while visualizations of motion intent have been studied extensively in previous work [9–14], communicating cobot perception remains an open challenge. It is vital that the human user can recognize what the semi-autonomous robotic system perceives explicitly (e.g., objects such as a glass or a bottle) because this enables users to recognize any occurred error in the robot's perception [16,17]. The non-perception of objects can have especially drastic consequences, as often demonstrated in autonomous driving. (Guardian. https://www.theguardian.com/technology/2018/mar/19/uber-self-driving-car-kills-woman-arizona-tempe, last retrieved 30 December 2021.) In this paper, we focused on the communication of cobot perception, and in particular, different methods to make the cobot's sensor-based detection of objects in its surroundings visible and clear to the user.

2.3. Visualization Techniques for Object Localization

As discussed in Section 1, AR, and in particular HMD-AR or MAR, reduce the field of view of the user as they either only display information in a small part in front of the user's eye (HMD) or on the available screen area. This means that they often require guiding the users' attention to an off-screen object of interest. For example, Biocca et al. proposed the *Attention Funnel* to achieve this attention shift [36]. However, these mostly depend on the possibility of MAR and HMD approaches to easily adjust the field of view by turning the device or head. Adapting them to SAR might be difficult as the projection is often fixed. In addition, these approaches are not usually meant to highlight and identify multiple objects in the surroundings. They would likely overwhelm the user with too much visual clutter.

However, there is a large body of research in the context of off-screen visualization techniques, originally addressing the challenge of small-screen devices, which could pose a promising approach for this particular challenge. *Halo* is an early off-screen visualization method proposed by Baudisch et al. and initially intended for small, rectangular screens [25]. It uses circles with their center around off-screen objects and their radius just large enough to cut the screen's border. Using *Halo*, the distance information is encoded in the arcs themselves and directly incorporates the scale of the scene, which was preferred by the users. Furthermore, *Halo* can be extended to on-screen objects by drawing the circle around the object. *Wedge* is another frequently used off-screen visualization technique proposed by Gustafson et al. [24]. It visualizes off-screen objects by attaching isosceles triangles to them. Two corners of the triangle are always on screen; the third is fixed to the point of interest. This leads to an encoded distance information by an amodal completion as with *Halo*. Similarly to *Halo*, *Wedge* can also be used to visualize in-view objects. Gruenefeld et al. already demonstrated in two studies that both *Halo* and *Wedge* are transferable to AR; however, they did not investigate SAR [37,38]. To address this, we applied off-screen visualization techniques to SAR and investigated their effectiveness for conveying cobot perception by visualizing all objects currently detected by the sensor system.

3. Experimental Approach

In this paper, we investigated how to communicate cobot perception in a scenario related to activities of daily living (ADL). Our main target group are people with physical

impairments. Our previous work—an ethnographic study to establish recommendations for the development of a robotic drinking and eating aids—has shown a clear need for (semi-) autonomous assistive technology during meal time [7]. Hence, we focused on a breakfast situation at a kitchen table (see Figure 2a). Our goal is to help users understand the cobot and its actions so that users are able to understand how the cobot works and predict potential failures. Overall, this should contribute to better collaboration and foster trust and acceptance.

Our research investigates the subjective experience, effectiveness, and efficiency of different visualization approaches. We conducted two independent remote studies with 12 participants in the first and 116 participants in the second experiment, analyzing a total of three different visualizations. In the first experiment (see Section 4), 12 people with physical impairments—the target group—participated and delivered valuable quantitative and qualitative insights. After this, we conducted a second experiment (see Section 5 Experiment II: Validation Study) with 116 participants without impairments to verify our findings. We selected the established visualizations *Halo* and *Wedge* (see Section 3.2.3) and compared them to a simplified line-based visualization—*Line*.

All visualizations served the purpose of optically highlighting and indicating each object on a kitchen table that the sensory system of the cobot is currently detecting. The challenge for the user is to understand a potential failure of the system, indicated by the ceased highlighting of a previously perceived object. The studies required users to recognize these failures and indicate the object no longer detected by the cobot.

We applied a SAR solution using a projector to display visualizations on the table surface, as detailed in Section 3.2. This visualization technology enables a dynamic workspace of the cobot with visual cues directly projected in the working area of the kitchen table.

3.1. Experimental Task

We wanted to determine which visualization technique allows users to recognize the cobot's perception errors quickly, accurately and with minimal effort. Therefore, users were presented with a simple task (see Sections 4.4 and 5.4). They had to observe a virtual scene where a robot arm was moving across a breakfast table containing multiple items. Initially the current visualization technique shows each object as detected and perceived by the cobot. After a randomized time in an interval of 5–15 seconds in *experiment I* (see Section 4) and an interval of 3–15 seconds in *experiment II* (see Section 5), the cobot ceased to detect a random object; indicated by a vanished visualization. The user had to a) recognize this situation as quickly as possible and b) identify the no longer perceived object.

3.2. Apparatus

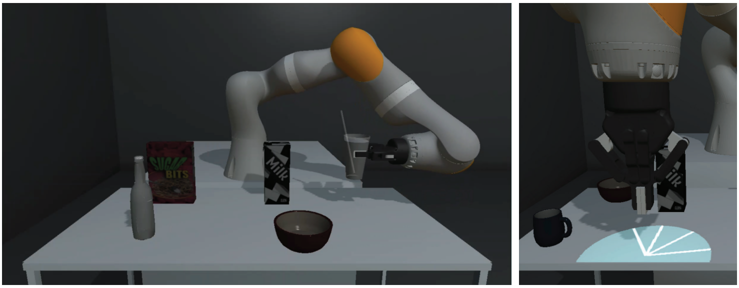
Here, we describe the developed apparatus of both experiments. In particular, we (a) describe our 3D testbed environment; (b) compare different mounting settings of the projector and report the concluding setting; and (c) introduce the selected visualization techniques.

3.2.1. 3D Testbed Environment

We developed a simulation of the robot setup present in our laboratory using the Unity3D Game Engine. (https://unity.com/, last retrieved 30 December 2021) We used Bio IK (https://assetstore.unity.com/packages/tools/animation/bio-ik-67819, last retrieved 30 December 2021) to simulate the robot's inverse kinematics. The project was exported as a WebGL application and hosted online for easy access by the participants within their particular web-browser environment. Any user interaction with the prototype was performed through mouse clicks, enabling participants with motor impairments to use their respective pointing devices.

For the virtual robot, we used a simulated KUKA LBR iiwa 7 R800 robot with a Robotiq 2-Finger 85 gripper module attached to the robot's flange. A simulated projector connected to the virtual flanch of the gripper/robot points towards the object of interest. Alternative mounting positions of the projector are discussed in Section 3.2.2. A virtual plane with a cir-

cular cutout restricts the simulated projection radius, creating a circular shape of projection to ensure the same size of projection to every site (projection distance: 50 cm; projection radius: 15 cm). The robot is located in front of a table (dimensions: 120 cm × 60 cm × 75 cm) with five items one might find in a hypothetical breakfast scenario (a box of cereals, a carton of milk, a plate of fruits, a bowl and a mug). See Figure 2 for a glance at the setup and a closeup on the projection.



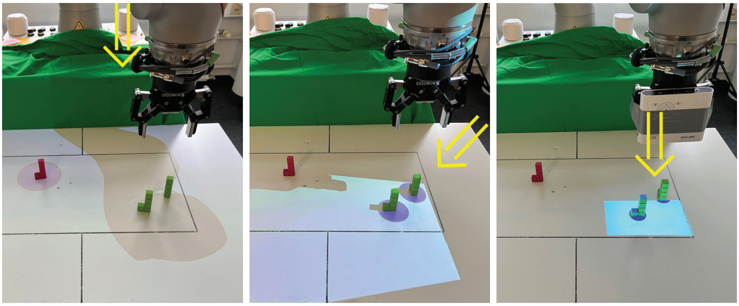
(**a**) 3D testbed environment showing the breakfast setup. (**b**) *Line* visualization.

Figure 2. Screenshots of the 3D testbed environment. (**a**): showing the complete setup with the five items placed on the table; (**b**): showing the *Line* visualization with one object on the table not perceived by the cobot.

3.2.2. Different Mounting Settings of the Projector

As part of the projection-based cobot perception visualization development process, we compared different potential mounting options for a pico-sized projector in a real-world setting with a real robot. As illustrated in Figure 3, we compared a top-mounted projection, e.g., from the ceiling, with a side-mounted projection, e.g., by using a tripod, with a cobot flanch-mounted projection by attaching it next to the gripper.

Top-mounted projection: Because of the large distance between the projector and table surface, a top-mounted projection has a large area. It can cover the whole workspace, e.g., the surface of the table, to visualize cobot perception. However, as shown in Figure 3a, objects in the vicinity of the cobot arm are not visualized. In addition, any visualization trying to highlight objects directly beneath the gripper is not visible due to the shadow that is cast by the cobot device itself.
Side-mounted projection: Attaching the projector at one side of the table tackles this issue of visualizing the objects of interest beneath the gripper and also enables a quite large projection area (see Figure 3b). However, a shadow can still hide the visualization related to objects in the vicinity of the cobot's arm due to the same reasons.
Cobot flanch-mounted projection: By mounting the pico-projector to the cobot's flanch next to the gripper, the cobot or its gripper do not cast a shadow within the projection area. Because the light comes from above, the size of the objects' shadows is reduced in contrast to the other projector settings, as shown in Figure 3c. As a drawback, the projection area is limited, and therefore, increases the need for visualization which can also highlight off-screen objects, which are currently not placed within the projection area.



(**a**) (**b**) (**c**)

Figure 3. For our work, we compared three different mounting-settings of a projector in the cobot's workspace to communicate cobot perception. The compared settings are (**a**) top-mounted; (**b**) side-mounted; and (**c**) cobot flanch-mounted projection. The direction of the projection is indicated by an arrow.

A top- or side-mounted setting leads to a large projection area but casts a shadow that a visualization cannot overcome. Here, it is especially hard to take into account the shadow cast by the robot arm and compensate for its' impact on a presented visualization. Whereas a flanch-mounted setting enables visualizing the objects that need the most attention in the case of cobot failure—the objects right under the gripper—but reduces the projection area. However, as discussed, we aim to explore whether this issue can be tackled by using off-screen visualization techniques.

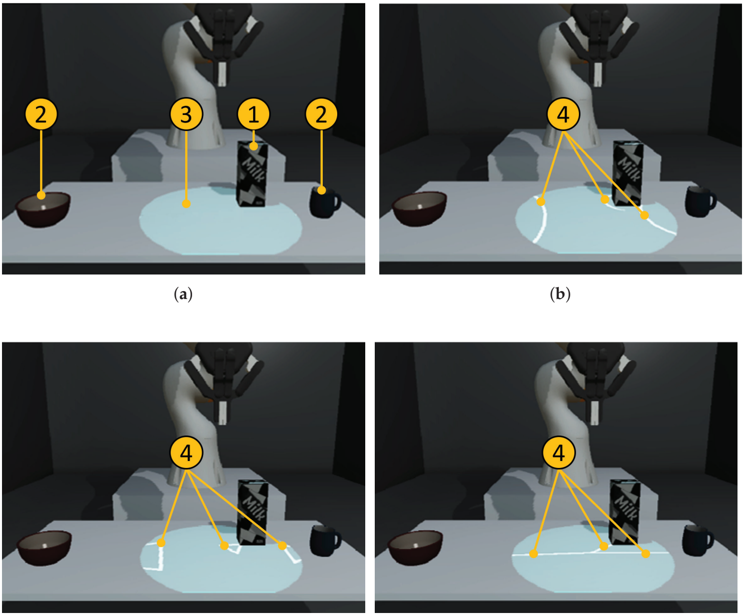
3.2.3. Selected Visualization Techniques

In our user study, we compare three different visualization methods: (a) *Halo*; (b) *Wedge*; and (c) *Line*. *Halo* and *Wedge* are well-established off-screen visualization techniques taken from previous work (see Section 2.3), while *Line* is proposed by us (see Figure 4).

Halo: Off-screen objects are visualized by attaching circles around objects, which any person around the table can see as well. These circles are always drawn with a radius as large as necessary to visualize part of the circle in the on-screen area. This means that the user can (a) understand the direction of the target object in the off-screen area; and (b) determine the distance, as this is encoded through the radius of the circle. For on-screen objects, we kept the circle visualization and show the radius to be 5 cm larger than the radius of the object's footprint.

Wedge: While the approach works similarly to *Halo*, here, off-screen objects are visualized by attaching isosceles triangles to them. Two corners of the triangle are always on-screen; the third is fixed to the point of interest. The distance is encoded via an amodal completion of the triangle, which avoids overlapping and leads to a reduced visual clutter. This allows a more accurate determination of the object's distance compared to *Halo*. For on-screen objects, we decided to keep the triangle attached to the corner of the object, pointing towards its center. The on-screen triangles point from the projection center to the object's center, comparable to arrow-based techniques [37].

Line: While *Halo* and *Wedge* try to encode distance information quite accurately, they also lead to visual clutter when many off-screen objects are visualized at the same time. Therefore, as a baseline, we propose a reduction to a simple line-based visualization technique. Here, a line connects the center of the projection to each object's center. Several lines—for each object one—are "shining" in a manner resembling a beam from the center in every direction (see Figure 2b). We still encode some distance information through the light intensity of the *Line* visualization.



(**a**) (**b**)

(**c**) (**d**)

Figure 4. A detailed overview of (**a**) the setup highlights the different parts as (1) on-screen object; (2) off-screen object; and (3) the projection area. The (4) main features are highlighted for the selected visualization techniques (**b**) *Halo*, (**c**) *Wedge*, and (**d**) *Line*.

As previous research shows that both *Wedge* and *Halo* work well within the realm of small-screen devices, we wanted to explore whether this can be adapted to the presented SAR off-screen problem. Visual complexity, in general, could cause problems when rapid judgments are necessary. Given that users need to recognize errors quickly and accurately, we wanted to add a visually less complex visualization method with *Line*. Still, all visualizations allow the user (a) to see that an object is recognized and (b) to infer its position relative to the projection center.

4. Experiment I: Target Group

The first experiment compared different visualization techniques following our experimental approach (see Section 3). We involved participants of the target group—people with physical impairments. Our goal was to explore how to best communicate cobot perception feedback to potential users for such essential tasks such as having breakfast to enable a more independent and self-determined life.

4.1. Study Design

To evaluate the performance of different visualization techniques for conveying cobot perception, we conducted a within-subjects remote user study with an counterbalancing order of the visualization techniques. Our independent variable was the visualization technique with three levels (*Line* vs. *Wedge* vs. *Halo*).

As dependent variables, we used a mixed-methods approach. As quantitative measures to evaluate task performance, we took into account recognizability, accuracy, reaction time, task load and individual Likert-scaled items. Furthermore, we collected qualitative data in the form of subjective feedback from our participants.

The recognizability describes the percentage of how often the simulated cobot failure was correctly recognized. To analyze this, we counted the number of trials on which a participant clicked—to interrupt the trial—only after an actual cobot failure happened. We acknowledge that this measure only gives an indication of recognizability, as we cannot be sure whether the participants really recognized the failure or simply thought that it should already have happened. Still, it excludes those clicks that happened before a failure occurred, where we can be sure that participants did not correctly judge the situation.

The accuracy describes the percentage of how often the correct affected object was identified. To determine the accuracy, we compared the selected object by the participants that they thought was no longer perceived by the cobot with the correct one. The result could either be correct or incorrect (0, 1).

For the reaction time, we measured the time from which a cobot's perception error happened to the point in time when the participant performed a mouse click (or equivalent input device). To reduce the impact of individual differences on reaction time, which can be quite large given not just cognitive differences but also differences in input devices and physical abilities, we measured a baseline reaction time for each participant and subtracted the median of this testing from the individual measurement. The resulting reaction time is: $time_{reaction} = time_{clicked} - time_{failure} - median(t_{reactionPreTest})$.

We used the mean of the task load scores by dimensions as measured by the NASA Raw-Task Load Index (Raw-TLX) [39] to determine the participants' perceived task load during the trials.

After each visualization, we asked three 7-point Likert-items (1 = strongly agree, 7 = strongly disagree) to determine participants' ability to detect which objects were perceived by the cobot and which were not, that neither the cobot itself nor the number of objects made it hard to observe the scenario, and if visualizations were understandable. We reported the mean values of each 7-point Likert-item.

We asked our participants to sign up for post-test interviews if they were interested. Unfortunately, only two of the participants did so. We conducted a 25-min telephone interview with these two participants on the same day that they participated in the remote experiment. Here, eleven open-ended questions were asked about the following topics:

- Status quo and acceptance of technology support;
- Appearance and implications;
- Trust and understanding;
- Preference and reason;
- Importance of a perceptual feedback;

4.2. Research Questions

To explore the suitability of the three selected visualizations, our research was guided by the following set of research questions:

RQ_1 **Do *Wedge* and *Halo*—because of their more detailed integrated distance information—enable the more accurate identification of failure objects or does the extra visual clutter disturb the user?**

RQ_2 **Do the different visualization techniques have an influence on the reaction time, i. e., are certain visual features quicker to recognize, process and thereby identify when they vanish?**

RQ_3 **How do users recognize the task load of different visualizations? Is the extra visual clutter of *Wedge* and *Halo* considered a problem or does the integrated distance information actually help reduce the task load?**

4.3. Participants

Twelve volunteers participated in this experiment: three females, three males and six who preferred not to say. They fell into four age groups: two participants were aged between 30 and 39 years; three participants were aged between 40 and 49 years; two participants were aged between 50 and 59 years; and one participant was aged between 60 and 69 years. Four participants preferred not to state their age. All participants suffered from complex motor impairments caused by spinal cord injuries and required assistance in everyday life. Only one participant had prior experience with cobots, while three participants mentioned some experience with toy robots.

Participants were recruited via announcements in different social media communities regarding assistive technology (e.g., Paraplegie.ch, Assistive Technology Community. https://community.paraplegie.ch/de/forum/hilfsmittel-technologie, last retrieved 30 December 2021) and social media discussion communities for people suffering from multiple sclerosis (MS) (e.g., mein.ms-life, Community for people suffering from multiple sclerosis (MS). https://mein.ms-life.de/ms-community/, last retrieved 30 December 2021) among other more local announcements. Participants did not receive any monetary compensation.

4.4. Procedure

Before the experiment started, participants were informed about the study and the experimental setup. This was implemented as a landing page for the study's URL. Participants had to give their informed consent by enabling a checkbox. Through another checkbox, they gave us the permission to use their anonymized recorded data. After a short demographic questionnaire, participants performed a reaction time test. We measured their reaction time when clicking on a screen as soon as a change in display color occurred. Ten repetitions allowed us to define the median time needed for the participant to react to a stimulus. We used this datum to determine the actual reaction time after recognizing a cobot perception error, thus reducing variability between subjects because of individual differences (e.g., latency of input devices, differences in physical abilities).

Participants then viewed a screen describing the first visualization method. We used images highlighting and describing any part of the visualization and a full text which gave step-by-step instructions. In a subsequent trial run, they watched the cobot perform a set of movement paths—which differed from those in other trials. Participants were instructed to click anywhere on the screen as soon as they noticed the disappearance of a visualization connected to an object. Right after they did click on such a case, a screen appeared which showed all potential target objects next to each other. The participants could then choose the item they thought the cobot did no longer perceive without any time constraints. Once this trial run was completed, the cobot performed twelve different movement paths as repetitions of this task, counting towards the data analysis. Participants viewed the twelve pre-programmed paths in random order. Six paths had an on-screen object disappear and six paths had an off-screen object disappear. Objects disappeared after a random time of between 5 and 15 s.

Once they completed all twelve paths for one visualization, participants filled out a NASA Raw-TLX questionnaire to report their workload. They also answered three additional questions specifically tailored to the respective experiment to evaluate their preferred visualization. The entire process was repeated with the two remaining visualization methods. The order in which the three visualization types were shown was counterbalanced using a Latin-square design. This experiment lasted an average of 40 min. The two participants who volunteered for the post-test interview did take part in this, as stated, on the same day as participating in the online study.

4.5. Results

During the *Line* technique run, one participant did not generate valid reaction times, as they performed mouse clicks before the cobot failure actually happened in every single trial. While this may be caused by an ineffective visualization, the fact that this hap-

pened in each trial and usually instantaneously after the start led us to the conclusion that the participant did not follow the test protocol. Consequently, this participant was excluded, which resulted in 11 remaining valid participant responses. We did not assume normality for the statistical analysis of our quantitative data and therefore we relied on non-parametric tests. Given the within-subject design with three conditions, we first applied a Friedman test as omnibus test followed by Wilcoxon tests as post hoc pairwise analysis with Bonferroni–Holm correction applied. Overall, the experiment resulted in 396 (11 participants $\times$ 3 visualization techniques $\times$ 12 trials) measured trials excluding training trials. Used abbreviations and symbols are:

- SD: Standard deviation;
- $\chi^2(2)$: Chi-squared with two degrees of freedom;
- p: p-value as expression of the level of statistical significance (p: ≤ 0.05 *, ≤ 0.01 **, and ≤ 0.001 ***);
- N: Sample size;
- W: Minimum sum of ranks;
- Z: Normalized minimum sum of ranks;
- r: Effect size (r: >0.1 small, >0.3 medium, and >0.5 large effect).

4.5.1. Recognizability: Percentage of Correctly Recognized Cobot Failures

In each trial, the cobot failed after 5–15 seconds. Participants had to respond with a mouse click to verify that they recognized the failure. However, in certain trials, participants did not click at all (*Wedge* = 3 trials; *Halo* = 2 trials; *Line* = 0 trials) or clicked before the object disappeared (*Halo* = 11 trials; *Wedge* = 3 trials; *Line* = 1 trials). From the reaction test at the beginning of the experiment, we calculated a median of the reaction time for each participant. This was taken into account to count those trials as unsuccessful, when the individual reaction to a cobot failure was faster than the median reaction time (*Line* = 12 trials; *Halo* = 11 trials; *Wedge* = 8 trials). The mean percentage of correctly recognized trials per participant for each visualization are (in descending order): *Line* = 90.2% (SD = 19.7%); *Wedge* = 89.4% (SD = 10.6%); and *Halo* = 81.8% (SD = 17.8%). A Friedman test showed no significant main effect of percentage of recognized failures on visualization ($\chi^2(2) = 4.71$, $p = 0.095$, N = 11).

Moreover, we can distinguish between correctly recognized on-screen and off-screen objects. The mean percentage of correctly recognized on-screen objects for each visualization are (in descending order): *Line* = 92.4% (SD = 20.2%); *Wedge* = 87.9% (SD = 15.1%); and *Halo* = 78.9% (SD = 27.0%). A Friedman test showed no significant differences ($\chi^2(2) = 4.26$, $p = 0.119$, N = 11). The mean percentage of correctly recognized off-screen objects for each visualization are (in descending order): *Wedge* = 90.9% (SD = 11.5%); *Line* = 87.9% (SD = 19.8%); and *Halo* = 84.8% (SD = 13.9%). A Friedman test again showed no significant differences ($\chi^2(2) = 1.19$, $p = 0.552$, N = 11).

4.5.2. Accuracy: Percentage of Correctly Identified Failure Objects

For the percentage of correctly identified objects that the cobot failed to perceive during the trial, we only considered all trials for which participants responded after the cobot failure happened (n = 345) and therefore had a chance to select the correct object. The mean percentage per participant of correctly identified failure objects per visualization are (in descending order): *Line* = 94.4% (SD = 7.1%); *Wedge* = 77.8% (SD = 15.9%); and *Halo* = 72.5% (SD = 19.9%). A Friedman test showed a significant main effect ($\chi^2(2) = 8.72$, $p = 0.012$ *, N = 11). Post hoc pairwise comparisons using a Wilcoxon signed-rank with Bonferroni correction showed a significant difference between *Halo* and *Line*, but not between any other pairs (see Table 1).

Table 1. Pairwise comparisons of accuracy for the visualization techniques: *Wedge*, *Halo*, and *Line*.

Comparison	W	Z	*p*	r
Wedge vs. Halo	33	1.39	0.563	0.30
Wedge vs. Line	2	−2.25	0.070	0.48
Halo vs. Line	1	−2.55	0.023 *	0.54

* $p \leq 0.05$.

4.5.3. Reaction Time

For the reaction time, we only considered trials for which participants correctly responded, meaning that participants clicked after the failure happened, responded before the trial ended (10 seconds after the cobot failure happened) and clicked on the correct object (n = 282; *Line* = 112, *Wedge* = 92, *Halo* = 78). We measured the time from the failure of the cobot visualization to the participant's mouse click. Again, we considered the median reaction time from the reaction test during the beginning of the study. The mean reaction time per visualization without extreme outliers ($\geq 3 \times$ IQR) was calculated for each participant.The mean reaction times for each visualization calculated over the means of the participants (without values $\geq 3 \times$ IQR) are (in ascending order): *Line* = 1.11 s (SD = 1.05 s); *Wedge* = 2.51 s (SD = 1.47 s); and *Halo* = 4.26 s (SD = 2.63 s). The reaction times are plotted in Figure 5.

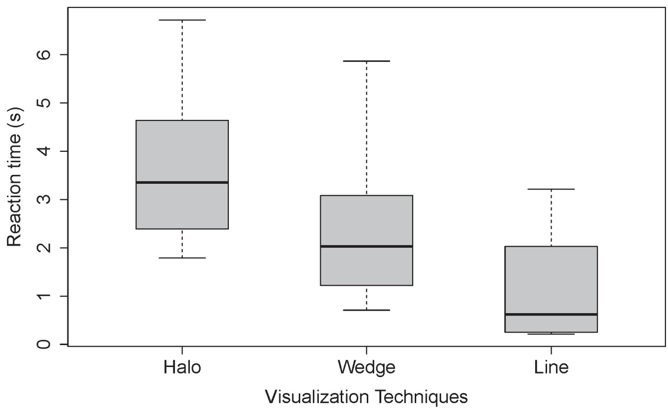


Figure 5. Comparison of the reaction times for the three different visualization techniques: *Wedge*; *Halo*; and *Line*.

A Friedman test revealed a significant main effect of reaction time on visualization ($\chi^2(2) = 11.64$, $p = 0.003$ **, N = 11). Post hoc pairwise comparisons using a Wilcoxon signed-rank with Bonferroni correction showed a significant difference between *Halo* and *Line*, but not between any other pairs (see Table 2). Concerning reaction times, we can conclude that *Line* has a significant lower reaction time than *Halo*.

Table 2. Pairwise comparisons of reaction times for the visualization techniques: *Wedge*; *Halo*; and *Line*.

Comparison	W	Z	*p*	r
Wedge vs. Halo	8	−2.22	0.073	0.47
Wedge vs. Line	59	2.31	0.056	0.49
Halo vs. Line	64	2.76	0.009 **	0.59

** $p \leq 0.01$.

4.5.4. Task Load

The mean of the task load ratings as measured by the NASA Raw-Task Load Index (Raw-TLX) [39] are (in ascending order): *Line* = 22.89 (SD = 16.47); *Halo* = 40.83 (SD = 17.28); and *Wedge* = 47.13 (SD = 22.89). A Friedman test revealed a significant main effect of task load on visualization ($\chi^2(2) = 7.09$, $p = 0.029$ *, N = 11). Post hoc pairwise comparisons using a Wilcoxon signed-rank with Bonferroni correction showed a significant difference between *Wedge* and *Line* (W = 60, Z = 2.40, $p = 0.041$ *, r = 0.51) and *Halo* and *Line* (W = 56, Z = 2.04, $p = 0.042$ *, r = 0.44), but not between *Wedge* and *Halo* (W = 44, Z = 0.98, $p = 0.730$, r = 0.21). Concerning the task load, we can conclude that *Line* has a significantly lower task load than *Wedge* and *Halo*. The resulting task load scores per individual dimension of the TLX are presented in Figure 6.

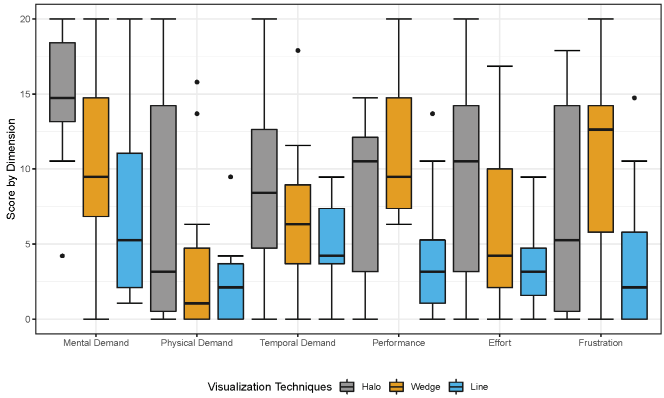


Figure 6. Comparison of the task load dimensions for the three different visualization techniques: *Wedge*; *Halo*; and *Line*.

Mental demand: A Friedman test revealed no significant main effect of mental demand on visualization ($\chi^2(2) = 15.69$, $p = 0.058$, N = 11);
Physical demand: A Friedman test revealed no significant main effect of physical demand on visualization ($\chi^2(2) = 5.43$, $p = 0.066$, N = 11);
Temporal demand: A Friedman test revealed no significant main effect of physical demand on visualization ($\chi^2(2) = 4.89$, $p = 0.087$, N = 11);
Performance: A Friedman test revealed a significant main effect of physical demand on visualization ($\chi^2(2) = 8.79$, $p = 0.012$ *, N = 11). Post hoc pairwise comparisons using a Wilcoxon signed-rank with Bonferroni correction showed a significant difference between *Halo* and *Line* (W = 66, Z = 2.93, $p = 0.003$ **, r = 0.63), but not between *Wedge* and *Halo* (W = 21.5, Z = -1.03, $p = 0.668$, r = 0.22) and *Halo* and *Line* (W = 39, Z = 1.20, $p = 0.744$, r = 0.26);
Effort: A Friedman test revealed no significant main effect of physical demand on visualization ($\chi^2(2) = 3.21$, $p = 0.201$, N = 11);
Frustration: A Friedman test revealed a significant main effect of physical demand on visualization ($\chi^2(2) = 7.39$, $p = 0.025$ *, N = 11). Post hoc pairwise comparisons using a Wilcoxon signed-rank with Bonferroni correction showed a significant difference between *Halo* and *Line* (W = 45, Z = 2.82, $p = 0.012$ *, r = 0.60), but not between *Wedge* and *Halo* (W = 17, Z = -0.76, $p = 0.977$, r = 0.16) and *Halo* and *Line* (W = 26, Z = 0.99, $p = 0.703$, r = 0.21).

4.5.5. Individual Likert-Items

After each visualization, we asked 3 7-point Likert-items (1 = strongly agree–7 = strongly disagree). Participants stated that the visualization helped them understand the position of the objects on the table for *Line* (Md = 2, IQR = 2.5) while they slightly disagreed for *Wedge* (Md = 5, IQR = 1.5) and *Halo* (Md = 5, IQR = 1.5). A Friedman test showed a significant main effect ($\chi^2(2) = 6.45$, $p = 0.040$ *, N = 11). Post hoc pairwise comparisons using a Wilcoxon signed-rank with Bonferroni correction showed a significant difference between *Wedge* and *Line* (W = 53.5, Z = 2.69, $p = 0.018$ *, r = 0.57, but not between *Wedge* and *Halo* (W = 20, Z = 0.18, $p = 0.891$, r = 0.04) and *Halo* and *Line* (W = 55.5, Z = 2.02, $p = 0.146$, r = 0.43).

Moreover, participants voiced that they could easily notice the cobot failure for *Line* (Md = 2, IQR = 2), while they slightly disagreed *Halo* (Md = 5, IQR = 2) and *Wedge* (Md = 6, IQR = 2). A Friedman test revealed a significant main effect ($\chi^2(2) = 9.5$, $p = 0.009$ **, N = 11). Post hoc pairwise comparisons using a Wilcoxon signed-rank with Bonferroni correction showed a significant difference between *Wedge* and *Line* (W = 45, Z = 2.83, $p = 0.012$, r = 0.60), but not between *Wedge* and *Halo* (W = 11.5, Z = 0.09, $p = 0.969$, r = 0.02) and *Halo* and *Line* (W = 49, Z = 2.19, $p = 0.094$, r = 0.47).

Furthermore, participants stated that for them, the cobot itself does not interfere with the recognition of the visualizations for *Line* (Md = 3, IQR = 2). In contrast, they slightly disagreed for *Wedge* (Md = 4, IQR = 1) and *Halo* (Md = 5, IQR = 1.5). A Friedman test revealed a significant main effect ($\chi^2(2) = 7.09$, $p = 0.029$ *, N = 11). Post hoc pairwise comparisons using a Wilcoxon signed-rank with Bonferroni correction showed a significant difference between *Halo* and *Line* (W = 43.5, Z = 2.55, $p = 0.035$ *, r = 0.54), but not between *Wedge* and *Halo* (W = 2.5, Z = -1.98, $p = 0.234$, r = 0.42) and *Wedge* and *Line* (W = 29.5, Z = 0.90, $p = 0.867$, r = 0.19).

4.5.6. Qualitative Insights

We applied open coding, followed by a thematic analysis of our interview data. We did this to find patterns of two participants' opinions and thoughts about the cobot assistance and presented visualizations. Once all the interviews were completed, two researchers transcribed all audio recordings and open coded the transcriptions. We then conducted an online affinity diagram of the open codes and organized the codes into groups, using Miro (https://miro.com, last retrieved 30 December 2021)—an online whiteboard [40]. During the telephone interview, we asked 11 open questions covering the status quo, a need for assistive technology, trust against such a cobot, if visualizations can increase this trust and understandability of cobot's perception, and which visualization technique they would or would not prefer and why. Because we asked open questions, we could identify further insights from the participants in addition to the one related to the visualizations. We identified four main themes, which we outline below.

Scenario and Technology Support

Both participants relied on assistance during breakfast from their caregivers and were interested in the concept of a cobot-supported breakfast routine. P2: "I am entirely open-minded and always interested in trying new things." However, several concerns were voiced, including the worry about the cost of a robotic aid and replacing the human caregiver, thus resulting in decreased social interaction. One additional design feature was frequently requested: the ability to mount the robotic arm to a wheelchair to increase flexibility.

Trust and Understanding

From the onset, the overall trust towards a robotic aid was high. However, the same principles as with humans apply; trust has to be earned. Participants indicated that their confidence in a cobot increased when they observed the cobot's perception and communicated with it. P1: "I understood the cobot's perception visualization, which helped

me trust the system because I could see which objects were perceived by the cobot". Easy-to-understand visualization methods can help address this concern by clearly displaying the cobot's perception, allowing for a greater level of user oversight.

Positive Feedback

Post-experiment feedback was positive. Participants were generally happy with the overall look and appearance of the cobot, a frequent concern of potential users. The off-screen visualization helped increase trust and user acceptance by increasing the collaborative effort between the human and cobot. Participants preferred the proposed new *Line* type over traditional visualization methods. P1: "I liked *Line* the most because there was always a clear reference, and I could see when something was out of order, even when I was looking somewhere else".

Problems and Drawbacks

When designing for non-tech-savvy users and physically vulnerable people, great care must be taken that the technology works consistently before releasing it for general use. P1: "I assume that any teething problems have been removed beforehand. That is why I already have a certain basic trust". This also increases end-user acceptance and addresses frequently mentioned reservations concerning the cobot making more mistakes than a human caregiver. P2: "I only have my caregivers as a reference. So if the cobot does not knock something over more often than my caregivers, I would be happy. However, even if the cobot would make a few more mistakes, I could forgive the cobot".

Participants voiced several issues concerning possible communication methods. Both *Wedge* and *Halo* were regarded as excessively complex and difficult to understand. P1: "I had problems recognizing *Wedge* because it was difficult to interpret the truncated arrows correctly". P2: "I did not prefer the visualization with the circles [*Halo*]. I really could not distinguish anything, and when something was gone, I could only guess which object it was".

4.6. Discussion

The results show a clear overall preference for a simple visualization technique such as *Line* to communicate which objects are recognized by the cobot. More complexly shaped visualizations, such as *Wedge* and *Halo*, lead to a comparatively higher task load in detecting unperceived objects.

4.6.1. Performance of Visualizations

No statistically significant results were found concerning correctly perceived cobot failures with overall high detection rates. This indicates that all three visualizations were effective in communicating the failure states.

However, there were differences in efficiency, with *Line* showing a significantly lower reaction time than *Halo* but not *Wedge*. Interestingly, the same results apply for accuracy, as *Line* shows a significantly higher accuracy compared to *Halo*. This indicates that the simple coding of *Line* has benefits even when the user has to understand which object is affected. The added information of the distance coding of Halo does not seem to overcome potential limitations due to visual clutter. While descriptive data do show differences between *Line* and *Wedge*, statistical tests do not confirm this, potentially due to low statistical power in a study with only twelve participants.

In our study, we were unable to confirm the advantages of *Wedge* in contrast to *Halo* in error rate and completion time as mentioned by Gustafson et al. [24]. While the descriptive data do show differences, another reason might be the round projection area. One potential advantage of *Wedge* compared to *Halo* is to overcome the corner-density problem of the latter [25], which is not applicable to round shaped-screens. This is in line with findings from Gruenefeld et al., who also found that with a round visualization area, the advantage of *Wedge* over *Halo* is less strong [37].

4.6.2. Task Load

We found that *Line* could significantly reduce participants' task load compared to the more complex *Wedge* and *Halo*. We believe this to be due to the simpler shape of *Line*, resulting in less visual clutter. In contrast to the other visualizations, with *Line*, users' focus of attention is on the gripper and center of the projection, which allows them to directly observe any changes. Using *Line* does not require the user to observe the periphery, making the visualization easier to interpret. Special consideration must be given to design an easy-to-view visualization, with objects and paths large enough to be identified at a glance.

4.6.3. Usefulness and Trust

Qualitative insights showed a clear preference for *Line*, mainly due to its avoidance of overlapping visualizations as with *Halo* and the amodal completion of the triangle as in *Wedge*. Although initial confidence in the robot's abilities is high, this needs to be maintained through constant and consistent correct behavior and clear communication. Our participants noted that higher levels of trust develop whereas failures happen rarely. Participants highlighted that a higher rate of mistakes compared to a human caregiver would result in lower acceptance and trust in the cobot.

4.6.4. Limitations

One of the main issues we faced when conducting this experiment was the small sample size. This limitation highlights the difficulty of designing for and involving people suffering from severe disabilities. The ongoing SARS-CoV-2 pandemic amplifies the problem as access to people is further restricted. Nonetheless, we believe that the remote nature of our study enabled us to gain valuable insights while granting access to participants from a wider geographic range. In addition to the low number of participants, even fewer participants were willing to participate in interviews via telephone, resulting in a limited number of qualitative data.

In our experiment, we did not measure trust with standardized questionnaires. Nonetheless, during our interviews, participants reported insight into how they would trust the cobot in this scenario. Hence, we did not address trust in-depth, i. e., with standardized measures, and thereby, our results can only be the foundation for further hypotheses and research concerning trust in cobots.

While this experiment was able to explore the potential benefits and drawbacks of the three selected visualization techniques, it did not provide clear statistical evidence in several cases. In particular, the differences between *Line* and *Wedge* did not show statistically significant effects, although *Wedge* is conceptually quite similar to *Halo*. Two difficulties can account for this, one being the overall small sample size and the second being the higher level of individual differences in the target group, potentially overshadowing smaller effects.

5. Experiment II: Validation Study

Based on the limitations of *experiment I* (see Section 4), we conducted a second experiment open to a non-specific user group, aiming to gain statistical evidence on particular hypotheses gained from this first experiment. Therefore, we also opted to exclude *Halo* from this second experiment, as results regarding this technique were already quite clear in the first. While the absolute results of such a second experiment with a non-specific user group may not be applicable to the target user group of people with physical impairments, we are confident that the relative results are. The main reasoning here is that the experimental task only requires very little physical interaction and the kind of physical interaction (mouse click) is kept constant for both tested visualizations *Line* and *Wedge*.

5.1. Study Design

Based on the study design of *experiment I* (see Section 4.1), we designed the second experiment as a within-subjects remote user study. Here, we changed, based on the results

of *experiment I* (see Section 4.5), our independent variable to visualization technique with two levels (*Line* vs. *Wedge*). The order of the visualization techniques was counterbalanced. We used the same quantitative measures to evaluate task performance (recognizability, accuracy, reaction time, task load, and individual Likert-scaled items).

5.2. Hypotheses

Based on the results of *experiment I*, which showed the potential advantages of *Line*, we developed the following set of hypotheses:

Hypothesis 1 (H1). *We hypothesize that Line can select the failure object with higher accuracy than Wedge. Results in* experiment I *already point in this direction. It seems that the user is mostly focused on the gripper and thereby the center of the projection, which benefits Line, as any change in the visualization can be directly seen in the center. While this should first benefit the recognizability of the failure, it also seems to have a positive effect on the accuracy, as the target object can usually be inducted from the vanishing line. Thereby, this should outweigh the better distance encoding of Wedge.*

Hypothesis 2 (H2). *We expect that Line allows quicker reaction times when the cobot's sensors lose track of an object compared to Wedge. We believe this to be the case because with Line, as before, the focus of attention is on the gripper and center of the projection, which allows the user to directly observe any changes in the Line visualization. In contrast, Wedge requires the user to observe the periphery.*

Hypothesis 3 (H3). *As a consequence of prior hypotheses, we also hypothesize that Line will lead to a lower task load. The Line visualizations that are characteristically displayed beneath the gripper and thereby in the center of the projection require less attention shifts from the user—which should be visible in task load measures.*

5.3. Participants

In total, we collected data from 209 participants. Since the experiment was conducted as an online study, the data of those participants were checked with regard to plausibility. We wanted to make sure that we did not include data from participants who simply "clicked through" the study without actually following the task protocol. Therefore, as a reasonable limit, we decided to remove participants whose median time of the mouse click ($time_{clicked}$) was less than three seconds. The limit of three seconds was chosen as the task scenario was designed in such a way, that it took at least three seconds for a cobot failure to happen.

This check led to the exclusion of 93 participants. The remaining 116 participants were categorized into four age groups: 85 of them were between 18 and 29 years old; 21 of them were aged between 30 and 39 years; one was between 40 and 49 years old; and four of them were 50–59 years old; and another three were 50–59 years old. Two participants preferred not to state their age.

In total, 63 participants had used a robot before the experiment, while 53 participants had no prior experience using robots. The remaining ten participants did not mention their prior experience using robots. Among all participants, 49 had previous experience using robots in the form of toy robots. In addition, 15 participants had used drones, eight had used service robots and another seven had used industrial or humanoid robots before. Furthermore, six participants had experiences with robots other than those mentioned.

Participants were recruited via SurveyCircle (https://www.surveycircle.com/, last retrieved 30 December 2021)—an open platform for survey submissions among other more local announcements. Participants did not receive any monetary compensation, but earned "survey ranking points" for their own study on SurveyCircle.

5.4. Procedure

The experiment followed the same procedure as described in Section 4.4, with the only difference being that, in all twelve pre-programmed paths, only off-screen objects disappeared after a random time between 3 and 15 seconds. We changed the focus onto off-screen targets, as here lie the main conceptual differences between the visualizations.

5.5. Results

For the analysis, we did not assume the normality of our quantitative data, especially since the measure reaction time was not normally distributed but generally right-skewed. Other measures, such as accuracy, are dichotomous by nature and therefore not on a metric scale. As a result, we applied non-parametric statistical tests. Given the within-subject design of our evaluation, we applied Wilcoxon signed-rank tests. Overall, we had 2784 (116 participants $\times$ 2 visualization techniques $\times$ 12 trials) measured trials, excluding training trials.

5.5.1. Recognizability: Percentage of Correctly Recognized Cobot Failures

The cobot failed after 3–15 seconds in each trial. Participants responded with a mouse click to verify that they recognized the failure. However, in certain trials, participants did not click at all ($n = 189$; *Line* = 105 trials and *Wedge* = 84 trials) or clicked before the visualization disappeared ($n = 1111$; *Wedge* = 601 trials and *Line* = 510 trials). From the reaction test at the beginning of the experiment, we calculated a median for the reaction time of each participant. This was taken into account to count those trials as unsuccessful, when the individual reaction to a cobot failure was faster than the median reaction time ($n = 22$; *Wedge* = 15 and *Line* = 7). The mean percentage of correctly recognized trials for each visualization are (in descending order): *Line* = 64.8% (SD = 29.8%) and *Wedge* = 59.5% (SD = 31.0%). A Wilcoxon signed-rank test showed no significant difference between the *Wedge* and *Line* ($W = 1559$, $Z = -1.56$, $p = 0.120$, $r = 0.10$, $N = 116$).

5.5.2. Accuracy: Percentage of Correctly Identified Failure Objects

For the percentage of correctly identified objects that the cobot failed to perceive during the trial, we again only considered all trials for which participants responded after the cobot failure happened ($n = 1462$). The mean percentage per participant of correctly identified failure objects per visualization are (in descending order): *Line* = 72.7% (SD = 27.9%) and *Wedge* = 64.4% (SD = 30.1%). A Wilcoxon signed-rank test showed a significant difference between *Wedge* and *Line* ($W = 1561.5$, $Z = -2.34$, $p = 0.019$ *, $r = 0.15$, $N = 116$).

5.5.3. Reaction Time

For the reaction time, we only considered all trials in which participants correctly responded, meaning participants clicked after the failure happened, responded before the trial ended (10 seconds after the cobot failure happened) and clicked on the correct object ($n = 1100$; *Line* = 608, *Wedge* = 492). We measured the time from the failure of the cobot visualization to the participant's mouse click. Again, we considered the median reaction time from the reaction test during the beginning of the experiment. The mean reaction time per visualization without extreme outliers ($\geq 3 \times$ IQR) was calculated for each participant. The mean reaction times for each visualization calculated over the means of the participants (without values $\geq 3 \times$ IQR) are (in ascending order): *Line* = 1.72 s (SD = 1.51 s) and *Wedge* = 2.07 s (SD = 1.81 s). The reaction times are plotted in Figure 7.

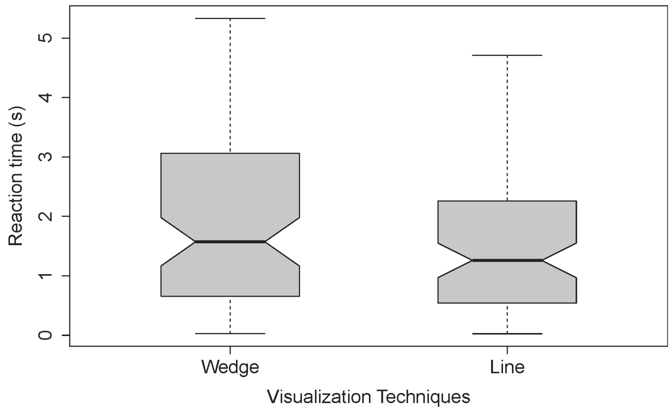


Figure 7. Comparison of reaction times for the two different visualization techniques: *Wedge* and *Line*.

A Wilcoxon signed-rank test showed no significant difference between *Wedge* and *Line* (W = 2254, Z = 1.44, $p = 0.151$, r = 0.11).

However, it appears that there is a potential interaction effect between the order variable (start visualization) and the independent variable (visualization). When participants started with *Line*, the means are (in ascending order): *Wedge* = 2.11 s (SD = 1.94 s) and *Line* = 3.32 s (SD = 3.11 s). A Wilcoxon signed-rank test showed a significant difference between *Wedge* and *Line* (W = 413, Z = −2.17, $p = 0.030$ *, r = 0.22).

Looking at *Wedge* as the start visualizations, the means are (in ascending order): *Line* = 1.24 s (SD = 1.02 s) and *Wedge* = 2.01 s (SD = 1.64 s). A Wilcoxon signed-rank test showed a significant difference between *Wedge* and *Line* (W = 904, Z = 2.57, $p = 0.009$ **, r = 0.25).

This effect shows that the mean of *Wedge* is relatively stable, independently of it being the first or second condition participants encountered (first condition: M = 2.11 s; SD = 1.94 s and second condition: M = 2.01 s; SD = 1.64 s). A Wilcoxon signed-rank test showed no significant difference (W = 725; Z = 0.84; $p = 0.404$; r = 0.08). However, the mean of *Line* depends on the ordering of the condition (first condition: M = 3.32 s; SD = 3.11 s and second condition: M = 1.24 s; SD = 1.02 s). A Wilcoxon signed-rank test showed this difference to be statistically significant (W = 1062, Z = 4.10, $p \leq 0.001$ ***, r = 0.41).

5.5.4. Task Load

The mean of the task load ratings as measured by the NASA Raw-Task Load Index (Raw-TLX) [39] are (in ascending order): *Line* = 42.39 (SD = 15.02) and *Wedge* = 47.19 (SD = 16.03). A Wilcoxon signed-rank test showed a significant difference between *Wedge* and *Line* (W = 4483.5, Z = 3.19, $p = 0.001$ ***, r = 0.21). Concerning the task load, we can conclude that *Line* has a significantly lower task load than *Wedge*. The resulting task load scores per dimension are presented in Figure 8.

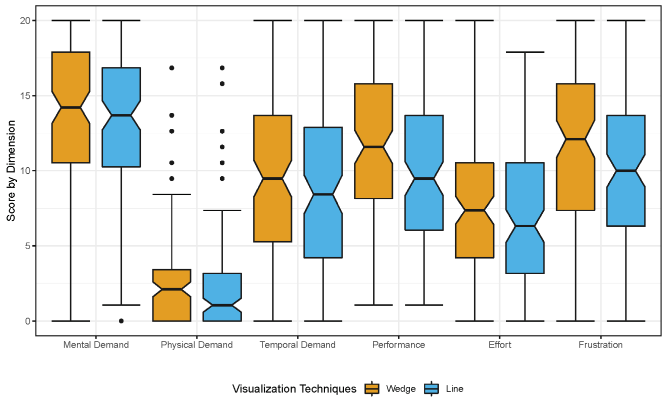


Figure 8. Comparison of task load dimensions for the two different visualization techniques: *Wedge* and *Line*.

Mental demand: A Wilcoxon signed-rank test showed a significant difference between *Wedge* and *Line* (W = 3068, Z = 2.52, $p = 0.011$ *, r = 0.17);
Physical demand: A Wilcoxon signed-rank test showed no significant difference between *Wedge* and *Line* (W = 1498.5, Z = 1.25, $p = 0.211$, r = 0.08);
Temporal demand: A Wilcoxon signed-rank test showed a significant difference between *Wedge* and *Line* (W = 2900, Z = 2.24, $p = 0.025$ *, r = 0.15);
Performance: A Wilcoxon signed-rank test showed a significant difference between *Wedge* and *Line* (W = 3618, Z = 2.35, $p = 0.018$ *, r = 0.15);
Effort: A Wilcoxon signed-rank test showed no significant difference between *Wedge* and *Line* (W = 3158.5, Z = 1.92, $p = 0.055$, r = 0.13);
Frustration: A Wilcoxon signed-rank test showed a significant difference between *Wedge* and *Line* (W = 3783.5, Z = 2.67, $p = 0.007$ **, r = 0.18).

5.5.5. Individual Likert-Items

After each visualization, we asked 3 7-point Likert-items (1 = strongly agree–7 = strongly disagree). Participants voiced that the visualization helped them to understand whether an object on the table was not detected by the cobot for *Line* (Md = 4, IQR = 2), while they slightly disagreed for *Wedge* (Md = 5, IQR = 3). Post hoc pairwise comparisons using a Wilcoxon signed-rank showed a significant difference (W = 2857, Z = 2.28, $p = 0.022$ *, r = 0.15).

Moreover, participants stated that the number of objects on the table did not disturb them for *Line* (Md = 3, IQR = 3) and *Wedge* (Md = 3, IQR = 3). Post hoc pairwise comparisons using a Wilcoxon signed-rank showed a significant difference (W = 1808, Z = 2.31, $p = 0.021$ *, r = 0.07).

Then, in the last question, participants mentioned that the visualization was always understandable for *Line* (Md = 3, IQR = 3), while they slightly disagreed *Wedge* (Md = 4, IQR = 3). Post hoc pairwise comparisons using a Wilcoxon signed-rank showed no significant difference (W = 2654, Z = 3.31, $p \leq 0.001$ ***, r = 0.22).

6. Discussion

Results from *experiment II* are in line with those from *experiment I* (see Section 4.6). Both highlight the advantages of using a straightforward visualization such as *Line* to show off-screen objects recognized by the cobot. Complex visualizations, such as *Wedge*,

appear to lead to a higher amount of errors in detecting perception failures, identifying the corresponding object, and a perceived higher task load.

6.1. Performance of Visualizations

Both experiments found no statistically significant results concerning correctly recognized cobot failures. All visualizations appear effective in communicating the cobots' failures. However, the recognition percentage dropped in the second experiment, as even with *Line*, participants were only able to correctly recognize cobot failures in 64.8% of the trials, compared to 90.2% in the first experiment. Potentially, the small but intrinsically motivated participant group in the first experiment did try to follow the protocol more closely. An alternative explanation might be an excessively conservative exclusion criteria (median response time smaller than 3 seconds) in *experiment II*.

When cobot failures were correctly recognized, a significant difference between *Line* and *Wedge* regarding accuracy became apparent. The percentage of correctly identified failure objects was highest with the *Line* (72.7%) visualization. This mirrors the results from *experiment I*, where the difference between *Halo* and *Line* was significant in favor of *Line*, with descriptive data showing an advantage also compared to *Wedge*. Hence, we can accept Hypothesis 1.

Regarding the reaction time, the overall trend resembles the first experiment, with *Line* having the lowest mean reaction time. However, the difference was again not statistically significant. Thus, we cannot accept Hypothesis 2. Interestingly, due to the simpler experimental design with only two conditions, we observed an interaction between the order of the conditions and the two different visualization techniques.

When participants worked with *Line* as the second visualization, the reaction time significantly improved compared to those cases in which participants started with *Line*—which, in turn, was not the case for *Wedge*). We concluded that *Line* might need a longer learning phase for participants to fully benefit from it. Why this is the case, however, remains an open question for future research.

6.2. Task Load

We found that *Line* significantly reduces participants' task load compared to the visually more complex *Wedge* technique. Therefore, we can accept our Hypothesis 3. We attribute this to *Line* not requiring attention shifts but rather allowing the user to focus on the gripper at all times. It also does not require amodal completion to decode the distance information, which, as the accuracy results show, is not necessary to understand and identify which object is affected.

6.3. Individual Likert-Items

The results of the individual Likert-items show that *Line* received better scores than *Wedge*. The simpler design of *Line* ensures that cobot failure and the corresponding object can be better detected, making this visualization more obvious than *Wedge*. For both visualizations, the number of objects in the experiment, five, did not disturb participants. This might, of course, change with a larger amount of objects, which could be necessary for more complex scenarios.

6.4. Limitations

The high number of excluded participants (93) infer that a remote study setup has less oversight, potentially enticing participants not to follow the study protocol. This issue required the need for careful data cleaning. One reason could be that further guiding of the participants was not possible.

Since participants could not be observed during the experiment, we cannot say whether they ran the experiment on a suitable device and whether their full attention was on the trials. Based on the reduced level of control, we expect additional noise in our data which may overshadow certain effects. This is a common problem with remote stud-

ies, as one cannot ensure, for example, identical technical conditions for each participant. Nonetheless, we do not believe that there has been a systematic bias in our data due to this reduced level of control.

Based on the experiences from *experiment I*, we deliberately did not include more open-ended questions or even an interview apart from the 7-point Likert-scale items. Experience shows that with a remote online study, the longer it takes, the higher the chance of participants dropping out. Therefore, we decided to keep the study as short and concise as possible and focused on quantitative data.

In addition to projection orientation, special considerations regarding ambient light and background effects have to be taken into account when using SAR. Our investigation neglected these external factors as they are present independently of the projection orientation. Nevertheless, more research is required to better understand their influence.

7. Conclusions

We investigated the performance differences of three visualization techniques in communicating cobot perception for a Spatial Augmented Reality setup, specifically focusing on people with physical impairments as potential end-users. We were particularly interested in comparing well-established off-screen visualization techniques to a reduced and straightforward line-based visualization for perceived objects inside and outside the projection area. The first experiment focused on 11 target group participants, while 116 non-specific respondents participated in the second experiment. Both experiments analyzed and compared the effectiveness, efficiency, subjective satisfaction and task load of the visualization techniques *Halo*, *Wedge* and *Line*. While the reaction times showed only minimal differences between *Line* and the established off-screen visualization techniques, *Line* did significantly improve the percentage of correctly identified failure objects and persistently lowered participants' task load. This result is mirrored by qualitative feedback from two target group participants, each highlighting the importance of an easy-to-understand visualization of the cobots' perception. Overall, our results stress that communicating the cobots' perception, including identification failures, is invaluable for assessing the overall situation and improving end-user trust. Our results generalize to similar pick-and-place workbench situations but may have limited applicability for more complex scenarios without a clearly defined environment. Overall, our findings add to a growing body of user-centered HRI literature with the overarching goal of increased user acceptance and confidence in cobots.

Author Contributions: Conceptualization, M.P.; methodology, M.P., U.G. and J.G.; software, M.P. and K.K.; validation, M.P.; formal analysis, M.P., U.G. and T.F.; investigation, M.P.; resources, M.P.; data curation, M.P.; writing—original draft preparation, M.P.; writing—review and editing, M.P., U.G., S.S. and J.G.; visualization, M.P.; supervision, U.G., S.S. and J.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the German Federal Ministry of Education and Research (BMBF, FKZ: 16SV7866K).

Institutional Review Board Statement: Not applicable. Ethical reviews by an Ethics Committee are waived in Germany as they are not compulsory for fields outside medicine. Only the participant's consent is requested.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Not applicable.

Acknowledgments: We would like to thank all study participants and their assistants for their valuable opinions and time. We acknowledge support by the Open Access Publication Fund of the Westphalian University of Applied Sciences.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bauer, A.; Wollherr, D.; Buss, M. Human–Robot Collaboration: A Survey. *Int. J. Humanoid Robot.* **2008**, *5*, 47–66. [CrossRef]
2. Fong, T.; Nourbakhsh, I.; Dautenhahn, K. A survey of socially interactive robots. *Robot. Auton. Syst.* **2003**, *42*, 143–166. [CrossRef]
3. Bemelmans, R.; Gelderblom, G.J.; Jonker, P.; de Witte, L. Socially assistive robots in elderly care: A systematic review into effects and effectiveness. *J. Am. Med. Dir. Assoc.* **2012**, *13*, 114–120. [CrossRef] [PubMed]
4. Colgate, J.E.; Edward, J.; Peshkin, M.A.; Wannasuphoprasit, W. *Cobots: Robots for Collaboration with Human Operators*; CiteSeer^{X}: University Park, PA, USA, 1996.
5. Chen, T.L.; Ciocarlie, M.; Cousins, S.; Grice, P.M.; Hawkins, K.; Hsiao, K.; Kemp, C.C.; King, C.H.; Lazewatsky, D.A.; Leeper, A.E.; et al. Robots for humanity: using assistive robotics to empower people with disabilities. *IEEE Robot. Autom. Mag.* **2013**, *20*, 30–39. [CrossRef]
6. Fattal, C.; Leynaert, V.; Laffont, I.; Baillet, A.; Enjalbert, M.; Leroux, C. SAM, an Assistive Robotic Device Dedicated to Helping Persons with Quadriplegia: Usability Study. *Int. J. Soc. Robot.* **2019**, *11*, 89–103. [CrossRef]
7. Pascher, M.; Baumeister, A.; Schneegass, S.; Klein, B.; Gerken, J. *Recommendations for the Development of a Robotic Drinking and Eating Aid—An Ethnographic Study*; Human-Computer Interaction—INTERACT 2021; Ardito, C., Lanzilotti, R., Malizia, A., Petrie, H., Piccinno, A., Desolda, G., Inkpen, K., Eds.; Springer: Cham, Swizerland, 2021. [CrossRef]
8. Pollak, A.; Paliga, M.; Pulopulos, M.M.; Kozusznik, B.; Kozusznik, M.W. Stress in manual and autonomous modes of collaboration with a cobot. *Comput. Hum. Behav.* **2020**, *112*, 106469. [CrossRef]
9. Gruenefeld, U.; Prädel, L.; Illing, J.; Stratmann, T.; Drolshagen, S.; Pfingsthorn, M. Mind the ARm: Realtime Visualization of Robot Motion Intent in Head-Mounted Augmented Reality. In Proceedings of the Conference on Mensch Und Computer—MuC '20, Magdeburg, Germany, 6–9 September 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 259–266. [CrossRef]
10. Chadalavada, R.T.; Andreasson, H.; Krug, R.; Lilienthal, A.J. That's on my mind! robot to human intention communication through on-board projection on shared floor space. In Proceedings of the 2015 European Conference on Mobile Robots (ECMR), 2–4 September 2015; pp. 1–6. [CrossRef]
11. Andersen, R.S.; Madsen, O.; Moeslund, T.B.; Amor, H.B. Projecting robot intentions into human environments. In Proceedings of the 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN), 26–31 August 2016; pp. 294–301. [CrossRef]
12. Coovert, M.D.; Lee, T.; Shindev, I.; Sun, Y. Spatial augmented reality as a method for a mobile robot to communicate intended movement. *Comput. Hum. Behav.* **2014**, *34*, 241–248. [CrossRef]
13. Stulp, F.; Grizou, J.; Busch, B.; Lopes, M. Facilitating intention prediction for humans by optimizing robot motions. In Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 28 September–2 October 2015; pp. 1249–1255. [CrossRef]
14. Watanabe, A.; Ikeda, T.; Morales, Y.; Shinozawa, K.; Miyashita, T.; Hagita, N. Communicating robotic navigational intentions. In Proceedings of the 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 28 September–2 October 2015; pp. 5763–5769. [CrossRef]
15. Bonci, A.; Cen Cheng, P.D.; Indri, M.; Nabissi, G.; Sibona, F. Human-Robot Perception in Industrial Environments: A Survey. *Sensors* **2021**, *21*, 1571. [CrossRef] [PubMed]
16. Honig, S.; Oron-Gilad, T. Understanding and Resolving Failures in Human-Robot Interaction: Literature Review and Model Development. *Front. Psychol.* **2018**, *9*, 861. [CrossRef] [PubMed]
17. Steinbauer, G. A Survey about Faults of Robots Used in RoboCup. In *RoboCup 2012: Robot Soccer World Cup XVI*; Chen, X., Stone, P., Sucar, L.E., van der Zant, T., Eds.; Springer: Berlin/Heidelberg, Germany, 2013; pp. 344–355.
18. Kragic, D.; Gustafson, J.; Karaoguz, H.; Jensfelt, P.; Krug, R. Interactive, Collaborative Robots: Challenges and Opportunities. In Proceedings of the 27th International Joint Conference on Artificial Intelligence (IJCAI'18), Stockholm, Sweden, 13–19 July 2018; AAAI Press: Palo Alto, CA, USA, 2018; pp. 18–25.
19. Trepkowski, C.; Eibich, D.; Maiero, J.; Marquardt, A.; Kruijff, E.; Feiner, S. The Effect of Narrow Field of View and Information Density on Visual Search Performance in Augmented Reality. In Proceedings of the 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), Osaka, Japan, 23–27 March 2019; pp. 575–584. [CrossRef]
20. Arévalo-Arboleda, S.; Pascher, M.; Baumeister, A.; Klein, B.; Gerken, J. Reflecting upon Participatory Design in Human-Robot Collaboration for People with Motor Disabilities: Challenges and Lessons Learned from Three Multiyear Projects. In Proceedings of the 14th PErvasive Technologies Related to Assistive Environments Conference—PETRA 2021, Corfu, Greece, 29 June–2 July 2021; ACM: New York, NY, USA, 2021. [CrossRef]
21. Chatzopoulos, D.; Bermejo, C.; Huang, Z.; Hui, P. Mobile Augmented Reality Survey: From Where We Are to Where We Go. *IEEE Access* **2017**, *5*, 6917–6950. [CrossRef]
22. Bimber, O.; Raskar, R. *Spatial Augmented Reality: Merging Real and Virtual Worlds*; A K Peters: Wellesley, MA, USA, 2005.
23. Malinverni, L.; Maya, J.; Schaper, M.M.; Pares, N. The World-as-Support. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems—CHI '17, Denver, CO, USA, 6–11 May 2017; Mark, G., Fussell, S., Lampe, C., Schraefel, M., Hourcade, J.P., Appert, C., Wigdor, D., Eds.; ACM Press: New York, NY, USA, 2017; pp. 5132–5144. [CrossRef]

24. Gustafson, S.; Baudisch, P.; Gutwin, C.; Irani, P. Wedge: Clutter-Free Visualization of off-Screen Locations. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI '08, Florence, Italy, 5–10 April 2008; Association for Computing Machinery: New York, NY, USA, 2008; pp. 787–796. [CrossRef]
25. Baudisch, P.; Rosenholtz, R. Halo: A Technique for Visualizing off-Screen Objects. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems—CHI '03, Ft. Lauderdale, FL, USA, 5–10 April 2003; Association for Computing Machinery: New York, NY, USA, 2003; pp. 481–488. [CrossRef]
26. Drolshagen, S.; Pfingsthorn, M.; Gliesche, P.; Hein, A. Acceptance of Industrial Collaborative Robots by People With Disabilities in Sheltered Workshops. *Front. Robot.* **2021**, *7*, 173. [CrossRef] [PubMed]
27. Canal, G.; Alenyà, G.; Torras, C. Personalization Framework for Adaptive Robotic Feeding Assistance. In *Social Robotics*; Springer International Publishing: Cham, Swizerland, 2016; Volume 9979, pp. 22–31. [CrossRef]
28. Hashimoto, K.; Saito, F.; Yamamoto, T.; Ikeda, K. A Field Study of the Human Support Robot in the Home Environment. In Proceedings of the 2013 IEEE Workshop on Advanced Robotics and Its Social Impacts (ARSO), Tokyo, Japan, 7–9 November 2013; IEEE: Piscataway, NJ, USA, 2013; pp. 143–150. [CrossRef]
29. Kim, D.J.; Hazlett-Knudsen, R.; Culver-Godfrey, H.; Rucks, G.; Cunningham, T.; Portee, D.; Bricout, J.; Wang, Z.; Behal, A. How Autonomy Impacts Performance and Satisfaction: Results from a Study with Spinal Cord Injured Subjects Using an Assistive Robot. *IEEE Trans. Syst. Man, Cybern. Part Syst. Humans* **2012**, *42*, 2–14. [CrossRef]
30. McColl, D.; Louie, W.Y.G.; Nejat, G. Brian 2.1: A Socially Assistive Robot for the Elderly and Cognitively Impaired. *IEEE Robot. Autom. Mag.* **2013**, *20*, 74–83. [CrossRef]
31. Pascher, M.; Baumeister, A.; Klein, B.; Schneegass, S.; Gerken, J. Little Helper: A Multi-Robot System in Home Health Care Environments. In Proceedings of the 1st International Workshop on Human-Drone Interaction, Glasgow, UK, 5 May 2019.
32. Tanaka, K.; Mu, S.; Nakashima, S. Meal-Assistance Robot Using Ultrasonic Motor with Eye Interface. *Int. J. Autom. Technol.* **2014**, *8*, 186–192. [CrossRef]
33. Dianatfar, M.; Latokartano, J.; Lanz, M. Review on existing VR/AR solutions in human–robot collaboration. *Procedia CIRP* **2021**, *97*, 407–411. [CrossRef]
34. Rosen, E.; Whitney, D.; Phillips, E.; Chien, G.; Tompkin, J.; Konidaris, G.; Tellex, S. Communicating and controlling robot arm motion intent through mixed-reality head-mounted displays. *Int. J. Robot. Res.* **2019**, *38*, 1513–1526. [CrossRef]
35. Walker, M.; Hedayati, H.; Lee, J.; Szafir, D. Communicating Robot Motion Intent with Augmented Reality. In Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction—HRI '18, Chicago, IL, USA, 5–8 March 2018; Kanda, T., Ŝabanović, S., Hoffman, G., Tapus, A., Eds.; ACM Press: New York, NY, USA, 2018; pp. 316–324. [CrossRef]
36. Biocca, F.; Owen, C.; Tang, A.; Bohil, C. Attention Issues in Spatial Information Systems: Directing Mobile Users' Visual Attention Using Augmented Reality. *J. Manag. Inf. Syst.* **2007**, *23*, 163–184. [CrossRef]
37. Gruenefeld, U.; Ali, A.E.; Heuten, W.; Boll, S. Visualizing Out-of-View Objects in Head-Mounted Augmented Reality. In Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services—MobileHCI '17, Vienna, Austria, 4–7 September 2017; Association for Computing Machinery: New York, NY, USA, 2017. [CrossRef]
38. Gruenefeld, U.; Ali, A.E.; Boll, S.; Heuten, W. Beyond Halo and Wedge: Visualizing out-of-View Objects on Head-Mounted Virtual and Augmented Reality Devices. In Proceedings of the 20th International Conference on Human-Computer Interaction with Mobile Devices and Services—MobileHCI '18, Barcelona, Spain, 3–6 September 2018; Association for Computing Machinery: New York, NY, USA, 2018. [CrossRef]
39. Hart, S.G. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2006**, *50*, 904–908. [CrossRef]
40. Harboe, G.; Huang, E.M. Real-World Affinity Diagramming Practices: Bridging the Paper-Digital Gap. In Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, Seoul, Korea, 18–23 April 2015; pp. 95–104. [CrossRef]

Exploring AI-Enhanced Shared Control for an Assistive Robotic Arm

Max Pascher[1,2]( ), Kirill Kronhardt[1], Jan Freienstein[3], and Jens Gerken[1]

[1] TU Dortmund University, Dortmund, NW, Germany
max.pascher@udo.edu
[2] University of Duisburg-Essen, Essen, NW, Germany
[3] Westphalian University of Applied Sciences, Gelsenkirchen, NW, Germany

Abstract. Assistive technologies and in particular assistive robotic arms have the potential to enable people with motor impairments to live a self-determined life. More and more of these systems have become available for end users in recent years, such as the *Kinova Jaco* robotic arm. However, they mostly require complex manual control, which can overwhelm users. As a result, researchers have explored ways to let such robots act autonomously. However, at least for this specific group of users, such an approach has shown to be futile. Here, users want to stay in control to achieve a higher level of personal autonomy, to which an autonomous robot runs counter. In our research, we explore how Artificial Intelligence (AI) can be integrated into a shared control paradigm. In particular, we focus on the consequential requirements for the interface between human and robot and how we can keep humans in the loop while still significantly reducing the mental load and required motor skills.

Keywords: Assistive Technologies · Human-Robot-Interaction Mixed Reality · Shared Control · Visual Cues

1 Introduction and Motivation

When controlling an assistive robotic arm, one major challenge from a human-robot interface perspective is the mapping between available input controls and resulting robot movements. Assistive robotic arms capable of performing arbitrary pick-and-place tasks in 3D-space require at least seven Degrees-of-Freedom (DoFs): x-, y- and z-translation, roll, pitch, yaw, and opening/closing the robot's fingers (*cardinal* DoFs). Therefore, there is no direct mapping possible with most input devices. As a result, the user needs to constantly switch between modes, i.e., flip through several pre-defined mappings between input and output space. Research has shown that such mode switching requires a considerable amount of time and cognitive demand [2,5,12].

This research is supported by the *German Federal Ministry of Education and Research* (BMBF, FKZ: 16SV8565).

L. Zaina et al. (Eds.): EICS 2023, LNCS 14517, pp. 1–14, 2024.
https://doi.org/10.1007/978-3-031-59235-5_10

Much research in assistive robotics is concerned with autonomous robotic functions [4,10,18,30]. From that perspective, the mode switching dilemma may seem moot with robots gaining more and more autonomy through advanced Artifical Intelligence (AI) and thereby reducing the need for such direct control altogether. However, studies have shown that people with motor impairments may prefer manual control over autonomous execution, as they see the robot not so much as another agent but as a tool to regain self-determination (e.g., [16,28]).

To tackle this dilemma, different approaches along the continuum of shared control methods were proposed (e.g., [8,11,12,15,17,27,29,33]). The concept of shared control has great potential to design communication and control between humans and robots [1]. For assistive robotic arms, however, the various approaches address the issue on different levels, sometimes reducing the involvement of the user to simply indicate an object to be picked [33]. Other approaches directly addresses the mode switching issue, with Herlant et al. suggesting time-optimal mode switching along the cardinal DoFs [12] and – based on Dijkstra's algorithm – to predict when the robot should switch modes.

Recently, Goldau & Frese proposed an approach integrating a Convolutional Neural Network (CNN), which interprets live camera data from the gripper to constantly describe the probabilistic distribution of intended DoF robot motion and, accordingly, optimal mapping of DoFs [11]. In principle, the idea is that the user gets a suggestion not just for when and how to switch mode but going beyond cardinal DoFs suggestions, allowing more flexible and adaptive DoF combinations.

In our work, we explored the feasibility of such a CNN-based approach, identified empirical implications for shared control systems, and what kind of human-robot interaction design is feasible [17,24,27]. For this paper, we present the main challenges we identified and how our work has aimed to address these. In summary, these challenges are:

- **AI legibility:** Given an AI which is able to automate mode switches, the user needs to understand the behavior and actions of the AI. A system which changes input mappings without notification or explanation would be perceived as unpredictable.
- **AI user control:** Given an AI which is able to automate mode switches, the user must stay in the loop and have a final say in making the choices.
- **AI intervention:** Given an AI which is able to automate mode switches, the user needs to expect and prepare for AI mistakes. Therefore, they need to interfere with the AI and, at best, make it reconsider the mode switching.

2 Engineering Context: Simulation Environment

In our research, we found that – given the complexity of robotic arms in combination with the limitations of current AI technologies – a testbed environment that allows the integration of different control mechanisms and user interface components facilitates engineering.

To that end, we developed *AdaptiX* [24], a transitional XR framework for shared control applications in assistive robotics. *AdaptiX* resembles a real-world scenario, where an assistive robotic arm (here a *Kinova Jaco 2*) is used to facilitate pick-and-place tasks as we observed that they are part of many Activities of Daily Living (ADLs). The system combines a physical robot implementation with a 3D simulation environment. This approach, reminiscent of simulations used in industrial contexts [19,22,32], helps to address challenges associated with bulky, expensive, and complex assistive robotic arms. Researchers are empowered to streamline their Design and Development (D&D) process, reducing complexity and enhancing efficiency. The system's integrated Robot Operating System (ROS) interface enables seamless connectivity to a physical robotic arm, supporting bidirectional interactions and data exchange through a *DigitalTwin* and *PhysicalTwin* approach.

In addition to Cartesian robot control, the framework includes Adaptive DoF Mapping Controls (ADMC), an initial shared control approach that employs AI-generated suggestions, subject to user approval and control. Figure 1 provides an overview of the framework's architecture.

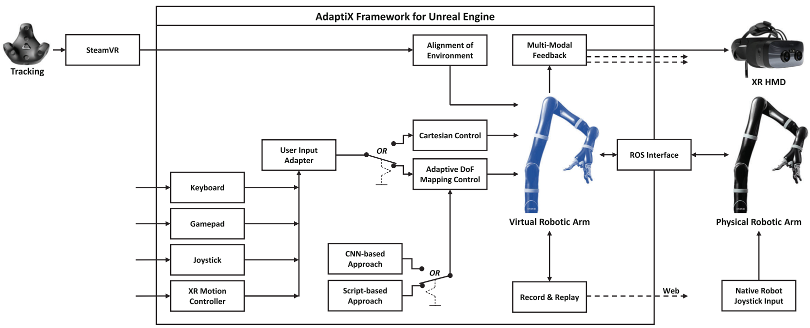


Fig. 1. Overview of *AdaptiX*' architecture, illustrating each component, their directional communication, and the crossover from and to the framework [24].

AdaptiX is built on the foundation of the game engine *Unreal Engine 4.27* [7]. This game engine is renowned for its advanced real-time 3D photorealistic visuals and immersive capabilities, making it an ideal choice for our framework. Furthermore, it offers a rich set of assets that can be readily used for future expansions. *Unreal Engine* is versatile and supports a variety of hardware configurations, allowing the framework to be deployed across different operating systems. It is compatible with a wide range of Virtual Reality (VR), Mixed Reality (MR), and Augmented Reality (AR) headsets, as well as gamepads and joysticks, making it suitable for development in both *C++* and *Blueprints*.

In the default scenario within *AdaptiX*, the focus is on a room that has been meticulously scanned using photogrammetry techniques. This room contains a table with an integrated virtual robotic arm, as depicted in Fig. 2. The simulation of the robotic arm has been optimized for operation via a VR motion controller, which features an analog stick, several functional buttons, and motion capture capabilities. An example of a compatible motion controller is the *Meta Quest 2* [20].

As most real-world scenarios will include pick-and-place operations [9,16,31,35], we designed a straightforward testbed scenario which requires to move a blue block to a red target area.

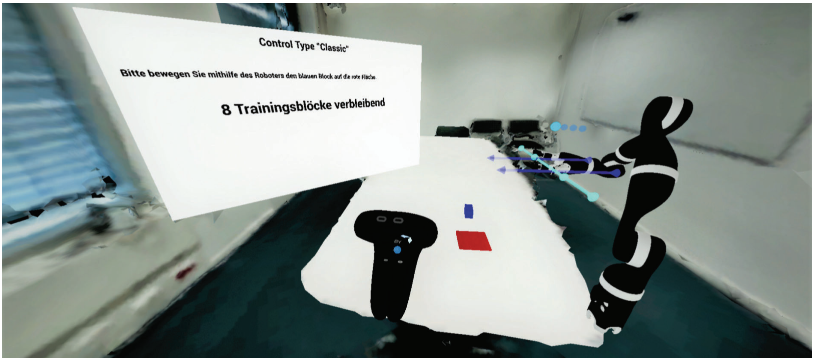
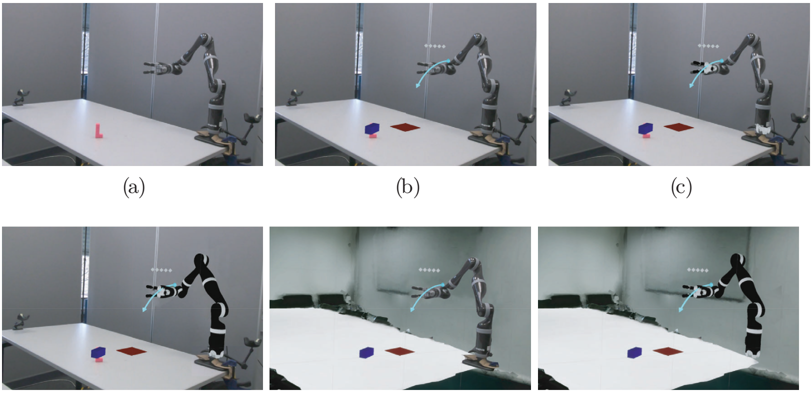


Fig. 2. Virtual environment consisting of (left to right): a virtual canvas, the motion controller, a table with a blue object and red target, and a *Kinova Jaco* with an arrow-based visualization [27].

We integrated the *Varjo XR-3* [34], a high-resolution XR Head-Mounted Displays (HMD), to create a seamless MR environment. By employing two *HTC VIVE* trackers [14], we synchronized the virtual and real worlds, ensuring that the operational spaces of the robots were perfectly aligned. This synchronization enables the adjustment of the MR level in multiple increments, as outlined in the *virtuality continuum* proposed by Milgram and Kishino [21]. A visual comparison between the user's perspective in the real world and the simulation is presented in Fig. 3.

The MR continuum comprises different levels. Level **(1)** serves as the study's baseline condition, offering no multi-modal feedback to the user. At level **(2)**, the system mimics an AR visualization technique, which overlays the entire physical setup with basic cues. Especially level **(3)** and **(4)** enable customizing either the robot itself or the environment to extent/exchange the physical setup but still not loosing the context. In **(3)** users can interact with a totally new or customized robot while being in a familiar environment. World's distractions can be excluded in **(4)** while the original robot is presented. Level **(5)** provides a VR environment that is entirely customizable.



(a) (b) (c) (d) (e) (f)

Fig. 3. MR continuum with (**a**) only the real robotic arm in real environment, (**b**) augmenting of directional cues in the real environment with the real robotic arm, (**c**) additional visualizing the gripper and base of the virtual robotic arm in the real environment, (**d**) visualizing the simulated robotic arm in the real environment, (**e**) visualizing the real robotic arm in the virtual environment, and (**f**) visualizing the simulated robotic arm in the virtual environment [24].

3 AI-Enhanced Shared Control

This section will provide more details about our previous research and how we addressed the three main challenges stated at the end of Sect. 1 – *AI legibility*, *AI user control*, and *AI intervention*. Therefore, we illustrate initial design concepts, work in progress, and prototypes that were evaluated in user studies.

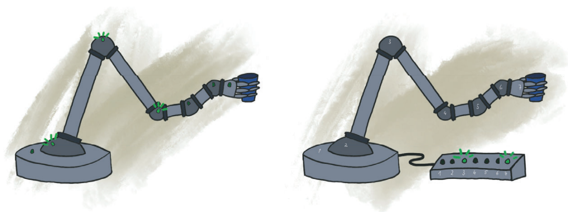
3.1 AI Legibility

In the context of our research, achieving a level of AI legibility is mostly concerned with making it easier to understand how the AI would reassign the input mapping and/or change the movement direction of the robot. In our recent survey on such robot motion intent approaches [25], we found that, for communicating location information (such as a movement direction), head-mounted technology such as AR HMDs allow to represent the robot movement visually and have shown to provide a potential fruitful approach [6]. Although research has explored robot motion intent, there needs to be more insight into what works best in various situations and for different user types. Customizing the visualization and feedback modality is crucial, as there is no "one size fits all" solution [13].

We proposed different design concepts that fall into a spectrum with two extremes – indicative and explanatory [26]. **Indicative:** Focus on crucial information only, quick and easy solution, suitable for experienced robot users.

Explanatory: Movements are shown in great detail, high level of information, especially helpful for new users.

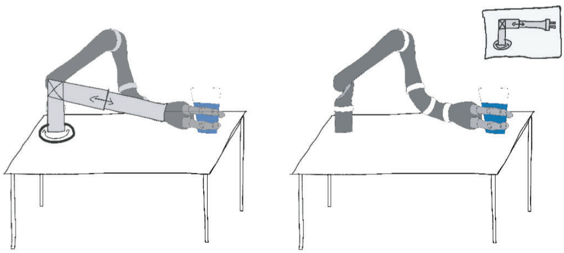
DoF-Indicator: LEDs attached to the robot's axis and joints – or mounted on a bar in front of it – communicate active and nonactive DoFs (see Fig. 4). Likely more suitable for experienced users because of indirect communication of movement, it communicates the DoF mapping and resulting movement abilities.



(a) (b)

Fig. 4. DoF-Indicator: **(a)** LEDs directly attached at each robot's joint; **(b)** LEDs mounted on a bar in front of the robot referring to each joint (1–7) [26].

DoF-Combination-Indicator: Movement ability is communicated by a simplified representation of the robot only able to move two DoFs simultaneously, for example, rotating and extending the arm (see Fig. 5). The AR representation either overlays the real robot or can be displayed separately in the corner of the AR screen. This visualization decreases the robot's complexity.



(a) (b)

Fig. 5. DoF-Combination-Indicator: **(a)** as an AR overlay, supporting robot and visualization in line of sight; **(b)** as an icon in the screen's corner [26].

Gizmo Visualization: Arrows, planes and point clouds communicate the current movement ability of the robot (see Fig. 6). This allows for several different

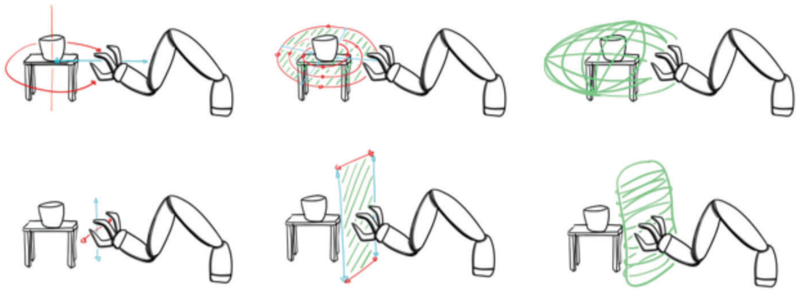


Fig. 6. Gizmo Visualization: **(left)** simple: straight and curved arrows; **(center)** planar: planes of movement; **(right)** cloud: 3D-cloud of possible positions [26].

design options. A arrow-based approach was already successfully evaluated in previous studies [17,27].

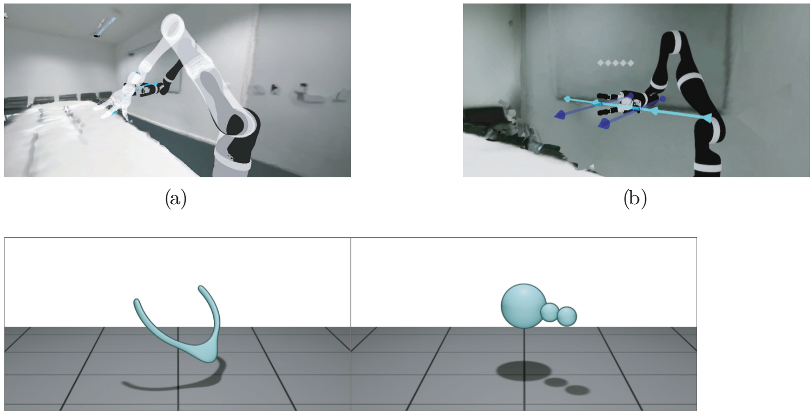
Demonstration: Current movement possibilities are demonstrated through either the actual robot or an AR ghost-representation. With both options a quick movement indicates the intended motion.

Visualization Approaches. Based on our initial concepts, we have explored augmenting the users' view with directional movement cues both in true – three-dimensional – AR (registered in 3D, [17,27]) as well as in 2D as symbolic representations on a data glass (for both refer to Fig. 7).

The former allows information-rich visualization and has shown in our studies to allow users to sufficiently anticipate a new suggested input mode mapping and corresponding movement direction [27].

The latter provides the advantage that the technology is already market ready and the devices are lightweight to carry and relatively easy to set up (compared to AR HMDs). However, without the ability to display directional visual cues registered in 3D space, the visual feedback is separated from the interaction space (robot) and may be more difficult to align with the current robot movement. In our work, we are currently exploring different visual forms, as can be seen in Fig. 7 c) and d).

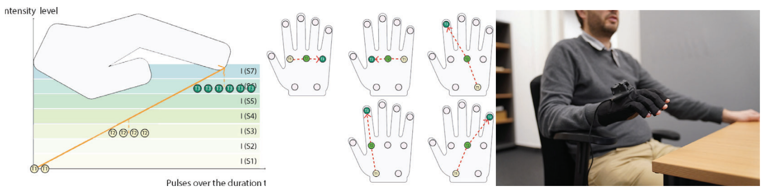
In addition, we have been exploring other modalities which could either replace or complement visual feedback to increase the legibility of the AI. In [23], we explored different designs for vibrotactile feedback to communicate three-dimensional motion directions. We developed two conditions based on the *Cutaneous Rabbit* illusion and one based on *Apparent Tactile Motion* to communicate 2D direction. The gradient of the overall 3D direction was then encoded by the number of discrete vibration pulses, the vibration intensity, or a combination of both. Our study showed that three-dimensional directional cues could be communicated with a high success rate for both the 2D direction and gradient, but



(a) (b) (c) (d)

Fig. 7. Visualization examples for directional cues of an AI-supported robotic control in a VR 3D environment: (**a**) Ghost & (**b**) Arrows [24], and a real-world setting via data glasses, e.g., *Google Glass EE2*: (**c**) Ring & (**d**) Points.

may benefit from dual-encoding of the gradient information as well as individual customization of the specific implementation of the vibrotactile feedback patterns (see Fig. 8).



(a) 3D Gradient (b) 2D Directions (c) Study Setup

Fig. 8. Vibrotactile directional cues **(a)** 3D Gradient encoding with pulses and intensity mapping **(b)** 2D direction encoding across the hand **(c)** illustrates the study setup, with the arm resting on the armrest while the hand is in the air [23].

Empirical Implications. Through related work and our own research, we identified that – while the robot system is designed to act in the user's best interest – the user still needs to build trust, which requires transparency and legibility that they can comprehend. They should also be able to interfere with the robot's control if the robot makes a mistake or gives inappropriate suggestions for interaction. Communicating intent further requires having the user pay attention or guiding the attention of the user, requiring multi-modal stimuli, depending on the situation and the capabilities of the user.

3.2 AI User Control

For the user to remain in control, automatic mapping of input modes may not be desirable. Instead, we have explored different ways which allow users to stay in control but also benefit from the potential increase in efficient task completion. In their original approach, Goldau & Frese [11] asked users to wait for five seconds without interacting with the robot to trigger a new mapping.

As this has shown to cause some level of frustration, we explored ways for the user to directly request a new mapping [17] as well as integrating a continuous or threshold-based feedforward visualization of an updated mapping [27]. The latter has shown positive effects, as the user thereby can always compare the current movement and mapping with an update from the AI, but only switch to that when they feel that it changes the movement direction to achieve the task better.

In both studies, *Classic* – a non-adaptive control mode inspired by the *Kinova Jaco 2* standard joystick input – relies on mode switching to access and control all DoFs one after another and was used as a baseline condition. In comparison to *Classic*, our AI-based ADMC methods significantly reduced **(1)** the task completion time, **(2)** the average number of necessary mode switches, and **(3)** the perceived workload of the user.

Users may have diverse input device preferences and capabilities. This calls for the availability of multi-modal input options or the ability to choose between different input modalities [3]. To enable such user input, our simulation environment provides a standard control approach where pressing a keyboard button moves the end effector along cardinal DoFs (x, y, z, roll, pitch, yaw, opening and closing the gripper). Using further build-in functionalities, the designated keyboard input can easily be adjusted to other input devices like gamepads, joysticks, or customized assistive input appliances.

Empirical Implications. While the goal is to keep users in control, the complexity of both the robot interaction and the DoF limitations for available input devices can easily make the system very difficult to use. Therefore, in order to find the sweet spot for shared control, we propose to start with a rather minimized set of user interaction and increase that on demand and depending on the individual capabilities. Users should not be confused by too many interaction options or overly complex movements. While optimal ways of accomplishing a goal may require complex intervention from the robot, these interventions may be difficult for users to understand, and therefore trust. In addition, keep the DoF of input devices low as this maximizes the amount of assistive devices capable of controlling the robot.

Our previous research and related work show that pick-and-place tasks are ubiquitous and necessary to perform ADLs. It is, therefore, important that shared control is first implemented for these simple tasks before more complex sequences are examined. If users struggle to understand shared controls for pick-and-place tasks, we believe it is highly likely that more complex tasks may cause further frustration.

3.3 AI Intervention

While our approaches for AI user control allow some level of intervention, since the user can decide when to accept the updated input mapping as provided by the AI, it has some limitations. If the AI algorithm – as suggested by Goldau & Frese – is not able to provide a useful mapping, the user may become stuck with little flexibility to trigger the AI to update the mapping. We are currently exploring different approaches to tackle this issue for the specific shared control mechanism. One rather straightforward approach would be to allow the user to disable the AI and go back to manual mode switching of Cartesian DoFs. This of course may decrease the acceptance and perceived usefulness of the AI.

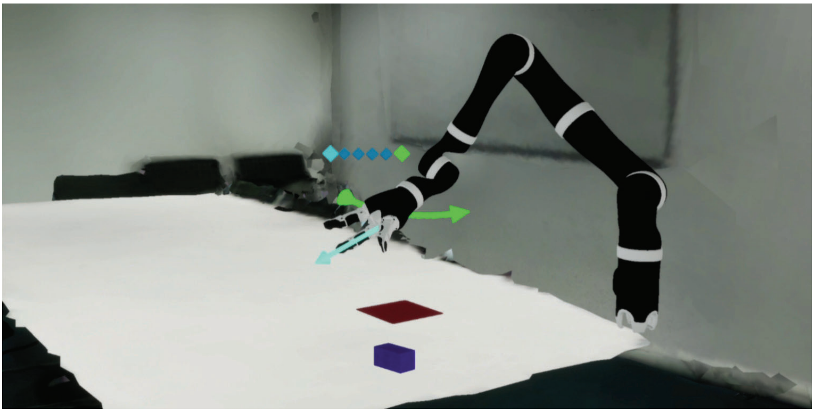


Fig. 9. The current DoF mapping (cyan arrow) does not allow to move to the blue object. By *changing the perspective* (green mode and arrow), the gripper is rotated in place to allow the CNN to suggest a new DoF mapping by an updated camera feed. (Color figure online)

A different way, which directly builds on the understanding of how the AI, in this case a CNN, operates, would be to find a way for the AI to *change perspective* – quite literally (see Fig. 9). If the robotic arm, or more specifically the gripper with the integrated or mounted camera is triggered to perform a small location repositioning, basically resulting in the robot looking around, the CNN will receive a new input which may result in a new and potentially better mapping.

Especially when the CNN is confronted with more than one choice (e.g., several objects in sight and vicinity of the gripper/camera) and has to choose one of them, based on the AI algorithm, rather than resulting in a draw. This kind of deadlock would only be resolved by a manual – Cartesian DoF – control. More suitable would be to either increasing AI's confidence of the chosen object by user input in the selected direction or decreasing by moving away, leading to a re-calculation and updated DoF mapping suggestion.

Empirical Implications. The aim is to keep the user in the loop so that they can intervene appropriately whenever the AI reaches its limits. However, this must strive for a balance that does not place sole decision dependency on the user to avoid access cognitive demand and temporal delays. Instead, establishing a four-eye principle with the AI functioning with implicit user consent until intervention is the most efficient approach to fulfilling the task's goal.

4 Conclusion

In this paper, we summarized our experiences for engineering AI-enhanced shared-control methods for assistive robotic arms. In particular, we identified three main challenges in *AI legibility*, *AI user control*, and *AI intervention*. Our work highlights the benefits and importance of sensible interaction design, which addresses these challenges and requires both a deep understanding of and interconnection with the AI technology. We also found that there is still much to be explored, in particular in the area of *AI intervention* approaches which go beyond circumventing the AI.

References

1. Abbink, D.A., et al.: A topology of shared control systems—finding common ground in diversity. IEEE Trans. Hum.-Mach. Syst. **48**(5), 509–525 (2018). https://doi.org/10.1109/thms.2018.2791570
2. Al-Halimi, R.K., Moussa, M.: Performing complex tasks by users with upper-extremity disabilities using a 6-DOF robotic arm: a study. IEEE Trans. Neural Syst. Rehabil. Eng. **25**(6), 686–693 (2017). https://doi.org/10.1109/TNSRE.2016.2603472
3. Arévalo-Arboleda, S., Pascher, M., Baumeister, A., Klein, B., Gerken, J.: Reflecting upon participatory design in human-robot collaboration for people with motor disabilities: challenges and lessons learned from three multiyear projects. In: The 14th PErvasive Technologies Related to Assistive Environments Conference - PETRA 2021. ACM (2021). https://doi.org/10.1145/3453892.3458044
4. Canal, G., Alenyà, G., Torras, C.: Personalization framework for adaptive robotic feeding assistance. In: Agah, A., Cabibihan, J.-J., Howard, A.M., Salichs, M.A., He, H. (eds.) ICSR 2016. LNCS (LNAI), vol. 9979, pp. 22–31. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-47437-3_3

5. Chung, C.S., Wang, H., Cooper, R.A.: Functional assessment and performance evaluation for assistive robotic manipulators: literature review. J. Spinal Cord Med. **36**(4), 273–289 (2013). https://doi.org/10.1179/2045772313Y.0000000132
6. Cleaver, A., Tang, D.V., Chen, V., Short, E.S., Sinapov, J.: Dynamic path visualization for human-robot collaboration. In: Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction, HRI 2021 Companion, pp. 339–343. Association for Computing Machinery, New York (2021). https://doi.org/10.1145/3434074.3447188
7. Epic Games: Unreal engine. Published online (2023). https://www.unrealengine.com. Version 4.27.2. Accessed 12 July 2024
8. Ezeh, C., Trautman, P., Devigne, L., Bureau, V., Babel, M., Carlson, T.: Probabilistic vs linear blending approaches to shared control for wheelchair driving. In: 2017 International Conference on Rehabilitation Robotics (ICORR), pp. 835–840 (2017). https://doi.org/10.1109/ICORR.2017.8009352
9. Fattal, C., Leynaert, V., Laffont, I., Baillet, A., Enjalbert, M., Leroux, C.: SAM, an assistive robotic device dedicated to helping persons with quadriplegia: usability study. Int. J. Soc. Robot. **11**(1), 89–103 (2019). https://doi.org/10.1007/s12369-018-0482-7
10. Gallenberger, D., Bhattacharjee, T., Kim, Y., Srinivasa, S.S.: Transfer depends on acquisition: analyzing manipulation strategies for robotic feeding. In: 2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 267–276 (2019). https://doi.org/10.1109/HRI.2019.8673309
11. Goldau, F.F., Frese, U.: Learning to map degrees of freedom for assistive user control: towards an adaptive DOF-mapping control for assistive robots. In: The 14th PErvasive Technologies Related to Assistive Environments Conference, PETRA 2021, pp. 132–139. Association for Computing Machinery, New York (2021). https://doi.org/10.1145/3453892.3453895
12. Herlant, L.V., Holladay, R.M., Srinivasa, S.S.: Assistive teleoperation of robot arms via automatic time-optimal mode switching. In: 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI), pp. 35–42 (2016). https://doi.org/10.1109/HRI.2016.7451731
13. Holloway, C.: Disability interaction (DIX): a manifesto. Interactions **26**(2), 44–49 (2019). https://doi.org/10.1145/3310322
14. HTC: Vive tracker 3.0. Published online (2023). https://www.vive.com/de/accessory/tracker3/. Accessed 12 July 2024
15. Jain, S., Farshchiansadegh, A., Broad, A., Abdollahi, F., Mussa-Ivaldi, F., Argall, B.: Assistive robotic manipulation through shared autonomy and a body-machine interface. In: 2015 IEEE International Conference on Rehabilitation Robotics (ICORR), pp. 526–531 (2015). https://doi.org/10.1109/ICORR.2015.7281253
16. Kim, D.J., et al.: How autonomy impacts performance and satisfaction: results from a study with spinal cord injured subjects using an assistive robot. IEEE Trans. Syst. Man Cybern. - Part A: Syst. Hum. **42**(1), 2–14 (2012). https://doi.org/10.1109/TSMCA.2011.2159589
17. Kronhardt, K., Rübner, S., Pascher, M., Goldau, F., Frese, U., Gerken, J.: Adapt or perish? Exploring the effectiveness of adaptive dof control interaction methods for assistive robot arms. Technologies **10**(1) (2022). https://doi.org/10.3390/technologies10010030
18. Lauretti, C., Cordella, F., Guglielmelli, E., Zollo, L.: Learning by demonstration for planning activities of daily living in rehabilitation and assistive robotics. IEEE Rob. Autom. Lett. **2**(3), 1375–1382 (2017). https://doi.org/10.1109/LRA.2017.2669369

19. Matsas, E., Vosniakos, G.C.: Design of a virtual reality training system for human-robot collaboration in manufacturing tasks. Int. J. Interact. Design Manuf. (IJIDeM) **11**(2), 139–153 (2015). https://doi.org/10.1007/s12008-015-0259-2
20. Meta: Meta quest 2. Published online (2023). https://www.meta.com/de/quest/products/quest-2/. Accessed 12 July 2024
21. Milgram, P., Kishino, F.: A taxonomy of mixed reality visual displays. IEICE Trans. Inf. Syst. **77**(12), 1321–1329 (1994)
22. Müller, S.L., Schröder, S., Jeschke, S., Richert, A.: Design of a robotic workmate. In: Duffy, V.G. (ed.) DHM 2017. LNCS, vol. 10286, pp. 447–456. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-58463-8_37
23. Pascher, M., Franzen, T., Kronhardt, K., Gruenefeld, U., Schneegass, S., Gerken, J.: HaptiX: vibrotactile haptic feedback for communication of 3D directional cues. In: Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems, CHI EA 2023. Association for Computing Machinery, New York (2023). https://doi.org/10.1145/3544549.3585601
24. Pascher, M., Goldau, F.F., Kronhardt, K., Frese, U., Gerken, J.: AdaptiX – a transitional XR framework for development and evaluation of shared control applications in assistive robotics. Proc. ACM Hum.-Comput. Interact. **8**(EICS) (2024). https://arxiv.org/abs/2310.15887
25. Pascher, M., Gruenefeld, U., Schneegass, S., Gerken, J.: How to communicate robot motion intent: a scoping review. In: Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems, CHI 2023. Association for Computing Machinery, New York (2023). https://doi.org/10.1145/3544548.3580857
26. Pascher, M., Kronhardt, K., Franzen, T., Gerken, J.: Adaptive DoF: concepts to visualize AI-generated movements in human-robot collaboration. In: Proceedings of the 2022 International Conference on Advanced Visual Interfaces, AVI 2022. Association for Computing Machinery, New York (2022). https://doi.org/10.1145/3531073.3534479
27. Pascher, M., Kronhardt, K., Goldau, F.F., Frese, U., Gerken, J.: In Time and space: towards usable adaptive control for assistive robotic arms. In: 2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN), pp. 2300–2307. IEEE (2023). https://doi.org/10.1109/RO-MAN57019.2023.10309381
28. Pollak, A., Paliga, M., Pulopulos, M.M., Kozusznik, B., Kozusznik, M.W.: Stress in manual and autonomous modes of collaboration with a cobot. Comput. Hum. Behav. **112**, 106469 (2020). https://doi.org/10.1016/j.chb.2020.106469
29. Quere, G., et al.: Shared control templates for assistive robotics. In: 2020 IEEE International Conference on Robotics and Automation (ICRA), pp. 1956–1962 (2020). https://doi.org/10.1109/ICRA40945.2020.9197041
30. Rakhimkul, S., Kim, A., Pazylbekov, A., Shintemirov, A.: Autonomous object detection and grasping using deep learning for design of an intelligent assistive robot manipulation system. In: 2019 IEEE International Conference on Systems, Man and Cybernetics (SMC), pp. 3962–3968 (2019). https://doi.org/10.1109/SMC.2019.8914465
31. Shafti, A., Orlov, P., Faisal, A.A.: Gaze-based, context-aware robotic system for assisted reaching and grasping. In: 2019 International Conference on Robotics and Automation (ICRA), pp. 863–869 (2019). https://doi.org/10.1109/ICRA.2019.8793804
32. Tidoni, E., et al.: Local and remote cooperation with virtual and robotic agents: a P300 BCI study in healthy and people living with spinal cord injury. IEEE

Trans. Neural Syst. Rehabil. Eng. **25**(9), 1622–1632 (2017). https://doi.org/10.1109/tnsre.2016.2626391

33. Tsui, K.M., Kim, D.J., Behal, A., Kontak, D., Yanco, H.A.: "I want that": human-in-the-loop control of a wheelchair-mounted robotic arm. Appl. Bion. Biomech. **8**(1), 127–147 (2011)
34. Varjo: Varjo XR-3 (2023). https://varjo.com/products/xr-3/. Accessed 12 July 2024
35. Wang, D., et al.: Towards assistive robotic pick and place in open world environments. In: Asfour, T., Yoshida, E., Park, J., Christensen, H., Khatib, O. (eds.) ISRR 2019. Springer Proceedings in Advanced Robotics, vol. 20, pp. 360–375. Springer, Cham (2022). https://doi.org/10.1007/978-3-030-95459-8_22

2023 32nd IEEE International Conference on Robot and
Human Interactive Communication (RO-MAN)
August 28-31, 2023. Paradise Hotel, Busan, Korea

2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN) | 979-8-3503-3670-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/RO-MAN57019.2023.10309381

In Time and Space: Towards Usable Adaptive Control for Assistive Robotic Arms

Max Pascher[1,2] and Kirill Kronhardt[1] and Felix Ferdinand Goldau[3] and Udo Frese[3] and Jens Gerken[1]

Abstract— **Robotic solutions, in particular robotic arms, are becoming more frequently deployed for close collaboration with humans, for example in manufacturing or domestic care environments. These robotic arms require the user to control several Degrees-of-Freedom (DoFs) to perform tasks, primarily involving grasping and manipulating objects. Standard input devices predominantly have two DoFs, requiring time-consuming and cognitively demanding mode switches to select individual DoFs. Contemporary Adaptive DoF Mapping Controls (ADMCs) have shown to decrease the necessary number of mode switches but were up to now not able to significantly reduce the perceived workload. Users still bear the mental workload of incorporating abstract mode switching into their workflow. We address this by providing feed-forward multimodal feedback using updated recommendations of ADMC, allowing users to visually compare the current and the suggested mapping in real-time. We contrast the effectiveness of two new approaches that a) *continuously* recommend updated DoF combinations or b) use discrete *thresholds* between current robot movements and new recommendations. Both are compared in a Virtual Reality (VR) in-person study against a *classic* control method. Significant results for lowered task completion time, fewer mode switches, and reduced perceived workload conclusively establish that in combination with feedforward, ADMC methods can indeed outperform classic mode switching. A lack of apparent quantitative differences between *Continuous* and *Threshold* reveals the importance of user-centered customization options. Including these implications in the development process will improve usability, which is essential for successfully implementing robotic technologies with high user acceptance.**

I. Introduction

While robotic devices have long been put behind fences for safety reasons, advances in the development of such (semi-) autonomous technologies have started to permeate almost all aspects of our personal and professional lives. These include increased close-quarter collaborations with robotic devices – from industry assembly lines [1] to mobility aides [2]. Assistive robotic arms are a particularly useful and versatile subset of collaborative technologies with varied applications in different fields, e.g., [3], [4].

Yet, new challenges arise when robots are tasked with (semi-) autonomous actions, resulting in additional stress for end-users if not correctly addressed during the design process [5]. Pollak et al. highlight the decreased feeling of control users experienced when using a robot's autonomous mode. Switching to manual mode allowed their study participants to regain control and decrease stress significantly. These findings are corroborated by Kim et al. whose comparative study of control methods resulted in markedly higher user satisfaction for the manual mode cohort [6].

A proposed solution from previous work [7] to these challenge are adaptive controls – referred as Adaptive DoF Mapping Controls (ADMCs) – which merge the advantages of (semi-) autonomous actions with the flexibility of manual controls. They combine multiple DoFs dynamically for a specific scenario to assist in controlling the robot. In our concept, a Convolutional Neural Network (CNN) interprets a camera's video feed of the environment and dynamically combines the most likely DoFs for a suggested movement. Building on this, we already showed that such ADMC combinations of the robot's DoFs can lead to a significantly lower number of mode switches compared to standard control methods [8]. However, our study could not show that this may also improve task completion time or reduce cognitive load. Also, challenges concerning the understanding of DoF mappings were raised during the study.

Based on these previous findings, the present study evaluates two novel ADMCs methods for an assistive robotic arm. We compare the variants *Continuous* and *Threshold*, differing in the time at which suggestions are communicated to the user, against a *classic* control method. In detail, we examine possible effects on task completion time, number of necessary mode switches, perceived workload, and subjective user experience. Our contribution is two-fold:

1) We demonstrate that both ADMC methods significantly reduce the task completion time, the average number of mode switches, and the perceived workload of the user.
2) Further, we establish that for *Continuous* and *Threshold*, each has specific advantages which some users may prefer over the other, raising the need for customizable configurations.

II. Related Work

Collaborative robotic solutions have received much attention in recent years. Previous work has generally focused on (a) different designs of robot motion intent and most recently (b) ADMCs for robots. The latter requires a critical yet seldom addressed topic in how collaborative robots can effectively communicate recommended movement directions to their user.

[1]Max Pascher, Kirill Kronhardt, and Jens Gerken are with the Westphalian University of Applied Sciences, Human-Computer Interaction, 45897 Gelsenkirchen, Germany max.pascher@w-hs.de, kirill.kronhardt@w-hs.de, jens.gerken@w-hs.de
[2]Max Pascher is also with the University of Duisburg-Essen, Human-Computer Interaction, 45127 Essen, Germany max.pascher@uni-due.de
[3]Felix Ferdinand Goldau and Udo Frese are with the University of Bremen, Mathematics & Computer Science, 28359 Bremen, Germany fgoldau@uni-bremen.de, ufrese@uni-bremen.de

A. Robot Motion Intent

Advance knowledge of the intended robot behavior and subsequent movements within the physical world are critical for effective collaboration when humans and robots occupy the same space and need to coordinate their actions [9]. In previous work, we analyzed existing techniques of communicating robot motion intent and identified different *intent types* as well as several intent properties, such as *location* and *information* or the placement of the technology [10]. Users generally prefer to have the robot's future movements represented visually [11]. To convey detailed robot motion intent, researchers often rely on Augmented Reality (AR) [12], [13], [14], as "with the help of AR, interaction can become more intuitive and natural to humans" [15].

Effective communication of robot motion intent is particularly relevant when using ADMCs for assistive robotic arms, as in such a shared or traded control environment each interaction needs to be precisely coordinated.

B. Adaptive DoF Mapping Controls

Traditionally, robot control methods include individual commands for each DoF, requiring frequent mode switches for controlling translations, rotations, and gripper functionality. Herlant et al. called into question the suitability of these standard control methods as task completion time markedly increases by using user-initiated compared to time-optimal mode switches [16].

To tackle this issue, we proposed in previous work the concept of ADMC – a dynamic combination of multiple DoFs, thus adjusted to specific scenarios or tasks [7]. This streamlining decreases the need for constant mode switching, resulting in faster and more efficient task fulfillment. In [7] we implemented a CNN as control unit to provide these dynamic DoF mappings and gave the user a triggering mechanism to request an update. In a 2D simulation study which had a 4-DoF robot control mapped to a 2-DoF input device, we found promising results.

We then extended this approach into a 3D VR simulation, thereby mapping a 7-DoF robot control to a 2-DoF input device [8]. We evaluated two ADMC methods – differing in their respective movement suggestion concept – against the baseline control method *Classic*. Simulating the effect of a CNN, our work relied on a task-specific script to provide DoF mappings based on the relative position and orientation between gripper and target. This removed the potentially confounding effect of a suboptimal CNN implementation. Results showed that the number of mode switches was significantly reduced compared to *Classic*, but task completion time was unaffected. Users reported high cognitive demand and difficulties understanding the mapping to 2 different input DoFs. In addition, the system felt difficult to predict and required trial and error [8].

III. ADAPTIVE DOF MAPPING CONTROLS

Building on our previous work [8], we created a VR simulation of a Human-Robot Interaction (HRI) experimental setup to compare different ADMC methods to a non-adaptive baseline condition *Classic*. Like in previous work [8] we applied a task-specific script to explore our ADMC methods. We tackle previous issues by 1) visualizing not only the current but also the forthcoming DoF mapping suggestion (improving predictability) and 2) reducing the input to a single DoF (reducing cognitive demand). We propose two approaches as different trade-offs between control fidelity and cognitive demand.

The VR simulation includes a virtual model of the *Kinova Jaco 2*[1] – a commercially available assistive robotic arm frequently used in HRI studies, e.g., [4], [16]. Our proposed visual feedback mimics AR, with directional cues registered in 3D space. This allows the user to understand different movement directions for the actual control and the suggested DoF combinations. To simplify understanding, we use *arrows*, a straightforward and common visualization technique to communicate motion intent [9], [17], [18].

As a control method for the ADMCs, we implemented a task-specific script. This removed any potential bias that a more generic but currently still technically limited approach such as a CNN-based control method may introduce. Of course, our approach only works in a controlled experimental setting. The task-specific script evaluates the gripper's current position, rotation, and finger position relative to a target. The DoF mapping system then suggests five different movement options (referred in the following to as *modes*) – in order of assumed usefulness – to the user.

1) *Optimal Suggestion:* Combining translation, rotation, and finger movement [opening and closing] into one suggestion, causing the gripper to move towards the target, pick it up, or release it on the intended surface.
2) An orthogonal suggestion based on (1) but excluding the finger movement. Allows the users to adjust the gripper's position while still being correctly orientated.
3) A pure translation towards the next target, disregarding any rotation.
4) A pure rotation towards the next target without moving the gripper.
5) Opening or closing of the gripper's fingers.

During movement, the ADMC system re-calculates the best DoF combinations to fulfill the specific task, which are then presented as new suggestions. Users cycle through these modes – by pressing a button on the input device – to select a suitable one or continue moving with the previous active suggestion (see Figure 1). A suggestion indicator is visible above the gripper when users are not moving the robot to distinguish between the modes. Five slanted cubes represent the possible suggestions. The cubes appear gray if no suggestion is active and turn blue to indicate that a new suggestion is selected. The cube corresponding to the selected mode increases in size. In contrast to our previous work [8] and to the dual axis system of the baseline control method (see Figure 2), only one input axis is required to control the robotic arm. Consequently, the cognitive demand

[1]Kinova Robotic arm. `https://assistive.kinovarobotics.com/product/jaco-robotic-arm`, last retrieved June 24, 2023.

on the users is reduced as they can focus on evaluating one movement rather than two simultaneous suggestions.

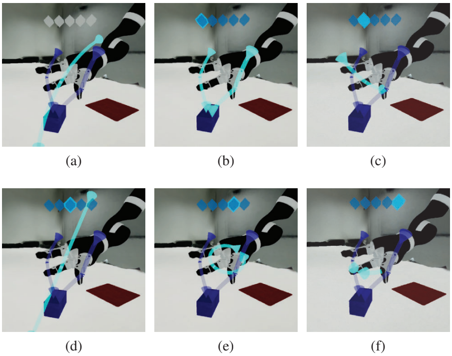


Fig. 1: Suggestions as visualized in the ADMC methods, (**a**) Continue previous movement, (**b**) Optimal Suggestion, (**c**) Adjustment Suggestion, (**d**) Pure Translation, (**e**) Pure Rotation, (**f**) Open / Close Fingers.

Continuous: This control method uses continuous feedback of robot motion intent to increase oversight of updated movement suggestions. Continuous feedback enables users to move in a direction and constantly evaluate the updated optimal suggestion by the ADMC system. If found fitting, users can switch to a new suggestion and move the robot in the updated path to fulfill the task. Here, two directional indicators are virtually attached to the robotic arm's gripper: a light blue and a dark blue arrow. The former represents the currently selected movement option (mode) mapped to the input axis. The forward movement of the input axis moves the gripper in the direction the arrow is pointing; engaging it backward moves the gripper in the arrow's reverse direction. The dark blue arrow represents the currently optimal suggestion at a given time. Users can only move the robot along the dark blue arrow if they switch to that suggestion first – which causes both arrows to overlap. While this approach increases transparency, users might be distracted by the constantly updating suggestions, potentially leading to more mode switches and perceived workload.

Threshold: In contrast to *Continuous*, *Threshold* uses time-discrete and multimodal feedback to indicate optimized movement suggestions. Again, a light blue arrow maps the selected movement option (mode) to the input axis. New suggestions are only shown to the users if the optimal mode differs – by a set degree – from the current movement. We followed Singhal et al. and used a cosine between-vector similarity measure to calculate this threshold [19], ranging from exact alignment [0%] to total opposite direction [100%]. In pretests, we determined a 20% difference between the current and optimal vector as the suggestion threshold. If exceeded, a short vibration pulse to the input device and a 1kHz sound inform the users of an updated suggestion. In addition, a dark blue arrow appears which visualizes the new suggested movement. Users can continue the active movement, switch to the new suggestion, or cycle through the other four modes before deciding on one. Unlike with *Continuous*, users can therefore entirely focus on the movement they are currently performing until explicitly notified and directed to a new suggestion. We expect *Threshold* to reduce perceived workload compared to *Continuous* as it does not require constant evaluation of the visual feedback. However, we expect task completion time to increase, as *Threshold* systematically interrupts the users' workflow. Additionally, *Threshold* might result in a perceived loss of control, potentially negatively influencing usability.

IV. Study Method and Materials

To explore the effectiveness of our ADMC methods, we conducted a supervised, controlled experiment as a VR simulation study with 24 participants. We compared our ADMC methods to *Classic*, which relies on mode switching to access and control all DoFs one after another. Approaches as *Classic* are well established (e.g., when driving a car) and are predictable and transparent for the user. Comparing ADMC methods to *Classic* allows HRI researchers to disentangle their respective advantages and disadvantages.

A. Study Design

We applied a within-participant design with *control method* as an independent variable with three conditions: (1) *Classic*, (2) *Continuous*, and (3) *Threshold*. Every participant performed eight training trials and 24 measured trials per condition, resulting in 72 measured and 24 training trials per participant and 1,728 measured trials in total. To counter learning and fatigue effects, the order of conditions was fully counter-balanced. We measured the following dependent variables:

1) **Average Task Completion Time** For each trial, we measured the time in seconds needed to pick an object and place it on the target surface.
2) **Average Number of Mode Switches** For each trial, we recorded every mode switch conducted by pressing a button on the input device.
3) **Perceived Workload** After completing each condition, we measured cognitive workload with the NASA Raw-Task Load Index (NASA Raw-TLX) questionnaire [20].
4) **Subjective Assessment** After completing each condition, we measured the five dimensions of the Questionnaire for the Evaluation of Physical Assistive Devices (QUEAD) [21]. After completing all trials, participants were further asked to rank the three conditions.

After each condition, participants were prompted with several open questions regarding their experience, their understanding of the control methods and the directional cues, plus any issue of interest they considered noteworthy. Additionally, participants were asked how they proceeded in situations when they could not solve the task at first.

Video and audio recordings of the interviews with the entire study cohort were assessed independently by two researchers. Open coding was applied to gather participants' opinions of the different control methods. We used Miro[2] – an online whiteboard [22] – to complete an affinity diagram of the open codes. Codes were then organized into themes (see Section V-F).

B. Hypotheses

Overall, we expected ADMC methods to reduce not just mode switches (as in prior work [8]) but – due to the advances in our designs – also improve on task completion time and workload.

H1: *Continuous* and *Threshold* lead to a lower task completion time compared to *Classic*. However, we expect *Continuous* to perform faster compared to *Threshold*, as the latter systematically interrupts the user during interaction.

H2: *Continuous* and *Threshold* result in fewer mode switches compared to *Classic*. We expect *Continuous* to require more mode switches than *Threshold*, as users have no clear guidance about when to switch modes. This may cause them to oversteer or accept new suggestions inefficiently.

H3: *Continuous* and *Threshold* cause lower perceived workload compared to *Classic*. However, we expect *Continuous* to cause a higher workload compared to *Threshold*, as it requires constant evaluation of the visual feedback while *Threshold* allows the user to relax until further notification.

C. Apparatus

Developing and testing new concepts for a robotic arm involves inherent challenges associated with a real robot's physical bulk and complexity. Quickly changing the experimental setup, adding feedback components, or providing information to the user further complicate testing regimes. We created a 3D testbed environment for HRI studies in VR to address these challenges. This testbed contains a simulated robotic arm (a virtual model of the *Kinova Jaco 2*) with multiple control mechanisms and a standardized pick-and-place task. Visual feedback mimics AR, with directional cues registered in 3D space. A *Meta Quest* motion controller is used as an input device to control the robotic arm.

Photogrammetry scans of an actual room were used to design the VR environment, which was created using the *Unreal Engine 4.27* and optimized for usage with a *Meta Quest* VR Head-Mounted Display (HMD) (see Figure 2). During the study, user behavior was recorded with appropriate software on a *Schenker XMG Key 17* laptop with *Windows 10 64-bit* and *Oculus Link* connected to the VR headset.

For our implementation of the baseline control method *Classic*, users cycled through four distinct modes to access all seven robot DoFs, as they are mapped on a two-DoF joystick, such as the control-stick on a *Meta Quest* motion controller:

1) X-Translation + Y-Translation
2) Z-Translation + Roll
3) Yaw + Pitch
4) Open/Close fingers

We illustrate the current mapping between the robot's DoFs and the input device through two arrows attached to the gripper. Light blue arrows indicate the robot's DoF assigned to the first, dark blue arrows to the second input axis. Looking at the joystick in VR, the same color-coded visualization is applied.

Users press a button on the input device – the A-Button of the *Meta Quest* motion controller – to switch between modes, cycling back to the first one at the end. Four blue spheres – in contrast to the slanted cubes used in our ADMC methods – above the robotic arm's gripper indicate the total number of available and the currently active mode when users are not moving the robot. The sphere representing the active mode is bigger and brighter than the spheres of inactive modes.

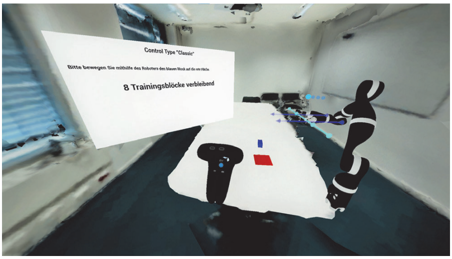


Fig. 2: Virtual environment consisting of (left to right): a virtual canvas, the motion controllers, a table with the blue object and red target, and a *Kinova JACO* with an arrow-based visualization

D. Participants

A total of 24 participants took part in our study (7 female, 17 male). The participants were aged 19 to 37, with a mean age of 26 years (SD = 4.85 years). No one declared any motor impairments that might influence reaction times. Five participants had prior experience with controlling a robotic arm. Participants were recruited from a university campus and an online appointment form.

E. Procedure

Utilizing the benefits of a standardized and portable VR simulation environment, the study was conducted in multiple comparable physical localities. Before commencing, participants were fully informed about the project objective and the various tasks they had to complete. Every participant gave their full and informed consent to partake in the study, have video and audio recordings taken, and have all the relevant data documented.

[2]Miro. https://miro.com, last retrieved June 24, 2023.

A study administrator observed the experiment on a laptop and briefed participants on using the hardware as well as the general functionalities of the study environment. Once set up, users followed command prompts embedded in the virtual simulation environment. For each of the three conditions, the following steps were performed:

1) Participants were given a written and standardized explanation of the control method used in the current condition.
2) Participants conducted eight training trials for familiarization with the respective control method.
3) Participants then conducted 24 measured trials.
4) Interview and questionnaires.

After completing all conditions, participants ranked the three control methods from most to least preferred and explained the reasoning behind their decision. The study concluded with a de-briefing. The average session lasted for 90 minutes and participants were compensated with 30 EUR.

F. Experimental Task

The experimental task is based on our previous work and resembles a common pick-and-place scenario [8]. A blue object appears on a table in front of the participant, which signals the start of a trial. The user has to control the robot from its starting position to pick the object and place it on a red target surface, also located on the table. To change the DoF mapping – for trial fulfillment – users could switch modes. Upon completion, the blue object disappears, and the robot automatically returns to the original starting position. A new blue object appears when this position is reached, and a new trial commences. For each trial, the position of the blue object is placed in one of eight possible locations spaced evenly around the red target surface. Each position occurred once during training and thrice during measured trials. However, the order of appearance was randomized. We used a neutral block shape rather than specific objects to avoid bias and ensure trial comparability.

V. Results

The study comprises 1,728 (24 participants $\times$ 3 control methods $\times$ 24 trials) measured trials. Training trials were excluded from the analysis.

We explored the distribution of the data through QQ-plots and either applied parametric Repeated Measures Analysis of Variance (RM-ANOVA) or non-parametric Friedman tests. For the latter, post-hoc pairwise comparisons using Wilcoxon signed-rank test with Bonferroni correction followed the omnibus test. Relevant effect sizes were calculated with r: $>$0.1 small, $>$0.3 medium, and $>$0.5 large effect.

A. Task Completion Time

Mean task completion time calculated per participant and control method (see Fig. 3) resulted in *Threshold* $=$ 16.54s (SD $=$ 4.09s); *Continuous* $=$ 16.61s (SD $=$ 4.77s); and *Classic* $=$ 30.96s (SD $=$ 4.89s). Outliers [N $=$ 3] with average times $\geq 2.2 * \text{IQR}$ of the mean task completion time in at least one control method were excluded [23]. The QQ-plot of the remaining 21 participants followed a normal distribution.

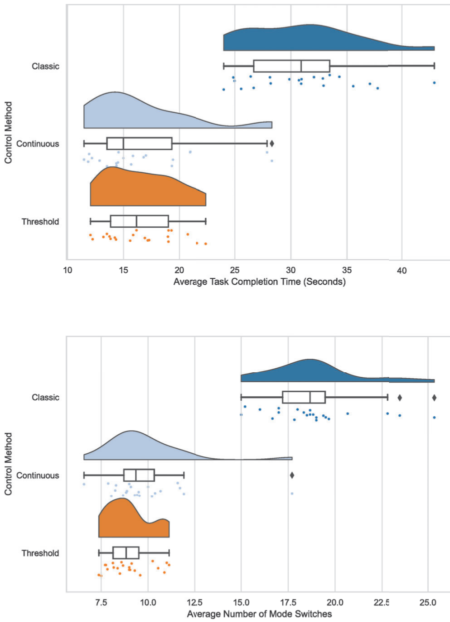


Fig. 3: Raincloud Plots for Average Task Completion Time and Mode Switches

A RM-ANOVA found a significant main effect (F(2, 36) = 130.92, p $\leq$0.001). A post-hoc pairwise comparison (Bonferroni corrected) showed a significant difference between *Continuous* and *Classic* (p $\leq$0.001) as well as between *Threshold* and *Classic* (p $\leq$0.001). No significant difference was found between *Continuous* and *Threshold* (p $\geq$0.999).

B. Mode Switches

We used a non-parametric Friedman test, as our data was not normally distributed, to determine differences between the average number of necessary mode switches between control methods. Two outliers – based on $\geq 2.2 * \text{IQR}$ of the mean value – were excluded prior to further analysis. This resulted in mean numbers of mode switches for *Threshold* $=$ 9.28 (SD $=$ 1.26); *Continuous* $=$ 9.93 (SD $=$ 1.47); and *Classic* $=$ 19.55 (SD $=$ 2.93) for N $=$ 22. We found a significant main effect ($\chi^2(2) = 33.82$, p $\leq$0.001, N $=$ 22). Post-hoc pairwise comparisons showed a significant difference between *Continuous* and *Classic* (Z $= -4.11$, p $\leq$0.001, r $=$ 0.62) as well as *Threshold* and *Classic* (Z $= -4.11$, p $\leq$0.001, r $=$ 0.62). Again, we found no significant difference between the two ADMC methods (Z $= -1.51$, p $=$ 0.131, r $=$ 0.28) (see Fig. 3).

C. Perceived Workload

NASA Raw-TLX [20] scores [scale from 1 to 100] for all participants resulted in mean task load values of *Threshold* = 22.67 (SD = 13.86); *Continuous* = 23.23 (SD = 13.26); and *Classic* = 34.24 (SD = 14.65). We applied a Friedman test which revealed a significant main effect for perceived task load: ($\chi^2(2) = 9.87$, $p = 0.007$, $N = 24$). Post-hoc pairwise comparisons show significant differences between *Continuous* and *Classic* ($Z = -3.03$, $p = 0.002$, $r = 0.44$), *Threshold* and *Classic* ($Z = -2.76$, $p = 0.006$, $r = 0.40$), but not between *Continuous* and *Threshold* ($Z = -0.21$, $p = 0.830$, $r = 0.03$).

D. Evaluation of Physical Assistive Devices

The QUEAD encompasses five individual scales (3 to 9 items each, 7-point Likert). Friedman tests for individual dimensions revealed significant main effects for *Perceived Usefulness (PU)*, *Perceived Ease of Use (PEU)*, *Emotions (E)*, and *Comfort (C)*, but not for *Attitude (A)*. Post-hoc pairwise comparisons indicate significant differences between *Continuous* and *Classic* for *PU*, *PEU*, and *C* as well as between *Threshold* and *Classic* for *PU* and *PEU* (refer to Table I for detailed scores).

TABLE I: Statistics for individual QUEAD dimensions: Perceived Usefulness (PU), Perceived Ease of Use (PEU), Emotions (E), Attitude (A), and Comfort (C).

	PU	**PEU**	**E**	**A**	**C**
Descriptive Statistics					
$M_{Classic}$	4.98	4.87	5.00	4.81	5.65
$SD_{Classic}$	1.39	1.20	1.71	1.75	1.71
$M_{Continuous}$	5.68	5.80	5.90	5.42	6.44
$SD_{Continuous}$	1.05	1.04	1.25	1.48	0.78
$M_{Threshold}$	5.77	5.90	5.68	5.44	6.13
$SD_{Threshold}$	1.02	0.97	1.43	1.58	1.14
Friedman Tests					
$\chi^2(2)$	7.49	15.22	7.20	1.76	6.39
p	0.022	$\leq$0.001	0.026	0.422	0.040
N	24	24	24	24	24
Pairwise Comparisons					
Classic vs. Continuous					
$\|Z\|$	2.32	2.47	1.85	—	2.29
p	0.021	0.014	0.064	—	0.022
r	0.33	0.36	0.27	—	0.33
Classic vs. Threshold					
$\|Z\|$	2.68	2.90	1.28	—	1.23
p	0.007	0.003	0.202	—	0.220
r	0.39	0.43	0.18	—	0.18
Continuous vs. Threshold					
$\|Z\|$	0.62	0.38	1.03	—	1.70
p	0.538	0.706	0.302	—	0.089
r	0.09	0.05	0.15	—	0.25

E. Individual Ranking

Participants ranked the control methods in order of preference from 1 = *favorite* to 3 = *least favorite*. Mean values in ascending order are *Continuous* = 1.67; *Threshold* = 2.04; and *Classic* = 2.29. A Friedman test revealed no significant main effect ($\chi^2(2) = 4.75$, $p = 0.100$, $N = 24$).

F. Qualitative Insights

Overall, the open coding process led to the identification of five main themes, as discussed below.

1) Familiarization: While all three control methods included a training phase, comments suggest that in particular the ADMC methods required familiarization. Here, participants felt the controls were sometimes "inverted" (P3) and wanted to "move the stick in the direction the arrow was pointing at" (P6). They also reported that "it takes a while to get used to" (P24), but "routine set in fast" (P18).

2) Handling Adaptive DoF Mapping Suggestions: The study cohort showed a relatively uniform response to the two ADMC methods with clear distinctions between *Threshold* and *Continuous*. In *Threshold*, many participants "trusted the system" (P23) and switched to the new suggestion as soon as they perceived the multimodal indicator. They "did not have to think a lot" (P4) and "relied on what the suggestion says" (P7). This dependence on the system caused some to "draw a blank when something went wrong because [they] forgot they had other options" (P8). One participant even tried using the *Threshold* control method with eyes closed, which "worked surprisingly well" (P7).

In contrast, participants evaluated the suggestions in *Continuous* more thoroughly, as they had to decide when to switch without the help of threshold-based indicators. Some participants waited for suggestions with relatively simple direction cues, such as "straight arrows" (P6, P16) as an indication to switch modes, while others trusted their "gut feeling" (P23). Uncertainties of "How do I approach this?" (P23) were more frequent in this control method than *Threshold*. Participants dealt with problems in both ADMC conditions in one of two ways to find alternative suggestions that better align with their needs. They cycled through the further offered suggestions for an alternative option or reversed their current movement direction until a different suggestion was offered.

3) Visualization: Overall, participants understood the different visualizations. Yet, difficulties arose in all three conditions relating to depth perception and understanding if the gripper is positioned correctly to pick or place the object. Some participants suggested a "laser pointer" (P16) to indicate the gripper's position above the table for improved depth perception. This is a known problem for robot teleoperation. In the past, researchers have suggested and explored AR *Visual Cues* to counter that, which include similar approaches as the ones mentioned by our participants [24], [25].

Interestingly, some participants "manipulated" the second mode of *Classic* (X- and Y-Translation) to mimic this effect, as that mode shows straight up- and downward pointing arrows as directional cues along the y-axis.

4) Multimodal Feedback: As described above, most participants used *Threshold* as intended, switching to the next suggestion when they received the multimodal feedback.

However, some participants experienced the haptic and audio indicators as "irritating" (P20) or "weird and horrible" (P17). The poignant statement "If I had to do this for five more minutes, it would be too annoying." (P7) reveals some participants' strong reactions to this control method. As a possible mitigation, one participant suggested implementing multiple thresholds of varying intensity instead of a singular one that "instantly beeps loudly at me and says 'Do this now!'" (P24).

5) Control vs. Comfort: Participants reported substantial differences in the level of control and comfort between *Classic*, *Continuous*, and *Threshold*. By nature, *Classic* offers the highest control level but requires participants to decide individually on every task step. In contrast, *Threshold* allowed participants to perform tasks "entirely brainlessly" (P16) and only press "forward, then A, then forward, then A" (P17). Many participants expressed that they "felt too directed by [*Threshold*]" (P8), attesting *Continuous* a higher level of comfort or "freedom to experiment" (P24). Overall, participants described *Continuous* as a reasonable compromise or "the golden middle" (P14) between the comfortable execution in *Threshold* and the high level of control in *Classic*.

VI. Discussion

Adaptive DoF mapping controls have already been indicated to have benefits over classic methods [7], [8]. Yet, research is still limited, and analysis of *time-based dimensions* of directional cues is lacking. In this paper, we examined to what extent the two ADMC methods, *Continuous* and *Threshold*, differ from the *Classic* baseline – and each other – in terms of task completion time, necessary mode switches, perceived workload, and subjective assessment.

Significant results for all four metrics partially support our initial hypotheses. Most strikingly, ADMC methods reduced task completion time (*H1*) and mode switches (*H2*) by 50% respectively compared to *Classic*. As previously suggested by Kim et al., this establishes that ADMC methods lead to faster and less involved execution of pick-and-place tasks [6]. These findings are in line with previous work [7], underlining the benefits of ADMCs compared to *Classic* controls.

In contrast to previous results [8], our novel ADMC methods were able to significantly lower task completion time and perceived workload compared to the *Classic* method. The latter finding also partially supports *H3*. This highlights that ADMCs which communicate the suggested recommendation to the user – irrespective of timing – were able to increase usability. Notably, the decreased workload of ADMCs is particularly meaningful as the end goal should be the smooth integration of robotic devices into people's lives and workflows, not to add stress.

Turning to the second part of our analysis – contrasting different time-based communication of feed-forward recommendations – we found no significant differences in the four metrics between *Continuous* and *Threshold*. The lack of measurable differences between *Continuous* and *Threshold* implies that both discrete and continuous communication of movement suggestions allows users to use ADMC methods efficiently. Insights gained by the results of the QUEAD and our qualitative interviews corroborate these findings, while the latter also helped to provide a more distinguished analysis.

Overall, participants expressed a positive stance regarding the ADMC methods. However, individual preferences vary greatly between *Continuous* and *Threshold*. While some participants preferred the higher level of control *Continuous* allowed, others favored the comfortable execution possible with *Threshold*. Consequently, future development of ADMC methods should – in accordance with Burkolter et al. – include individualization options to increase comfort and end-user acceptance [26]. Customizations would be particularly beneficial for *Threshold*-based controls as participants repeatedly criticized the multimodal feedback. Allowing users to adjust the modalities, the signal intensity, and even the threshold itself may improve usability while still offering the advantages of ADMC.

In contrast to expectations derived from our initial hypotheses, qualitative insights revealed that the *Classic* control method could still be a valuable addition in specific situations. Participants felt an apparent lack of control when the ADMC suggestions did not match their expectations. To improve usability, ADMC methods could incorporate static suggestions for certain situations. A potential way to address this could be combining ADMC and static suggestions using only the most common input-DoFs.

However, further experimental studies are needed to disentangle exactly which factors shape personal preferences and how customizations or crossover methods can deliver the best results.

A. Limitations

We explored the proposed ADMC methods in a VR simulation environment. While the usage of virtual simulations in industrial settings has been successfully established [27], [28], [29], future work should confirm if our promising findings can be replicated in the real world with a physical robot.

VII. Conclusions

Our ADMC methods *Continuous* and *Threshold* are promising approaches to communicate proposed directional cues effectively. We extend our previous work [8] by demonstrating that ADMCs significantly reduce task completion time (1), the average number of necessary mode switches (2), and the perceived workload of the user (3). Further, we establish that *Continuous* and *Threshold* perform equally well in quantitative measures while qualitative insights reveal individual preferences.

The observations of this study provide valuable implications for any HRI researcher involved in designing novel ADMC methods for human-robot collaborative settings. Future work should focus on disentangling quantitative and qualitative feedback of focus groups to develop optimal robot motion control methods, thus increasing usability, safety and – ultimately – end-user acceptance.

ACKNOWLEDGMENT

This research is supported by the *German Federal Ministry of Education and Research* (BMBF, FKZ: 16SV8563, 16SV8565). Our study is approved by the Ethics Committee of the *Faculty of Business Administration and Economics of the University of Duisburg-Essen* with the ID: 2202007.

REFERENCES

[1] S. Bragança, E. Costa, I. Castellucci, and P. M. Arezes, *A Brief Overview of the Use of Collaborative Robots in Industry 4.0: Human Role and Safety*. Cham: Springer International Publishing, 2019, pp. 641–650. [Online]. Available: https://doi.org/10.1007/978-3-030-14730-3_68

[2] C. Fattal, V. Leynaert, I. Laffont, A. Baillet, M. Enjalbert, and C. Leroux, "SAM, an Assistive Robotic Device Dedicated to Helping Persons with Quadriplegia: Usability Study," *International Journal of Social Robotics*, vol. 11, no. 1, pp. 89–103, 2019.

[3] T. B. Pulikottil, S. Pellegrinelli, and N. Pedrocchi, "A software tool for human-robot shared-workspace collaboration with task precedence constraints," *Robotics and Computer-Integrated Manufacturing*, vol. 67, p. 102051, Feb. 2021. [Online]. Available: https://doi.org/10.1016/j.rcim.2020.102051

[4] M. Beaudoin, J. Lettre, F. Routhier, P. S. Archambault, M. Lemay, and I. Gélinas, "Impacts of robotic arm use on individuals with upper extremity disabilities: A scoping review," *Canadian Journal of Occupational Therapy*, vol. 85, no. 5, pp. 397–407, 2018, pMID: 30866682.

[5] A. Pollak, M. Paliga, M. M. Pulopulos, B. Kozusznik, and M. W. Kozusznik, "Stress in manual and autonomous modes of collaboration with a cobot," *Computers in Human Behavior*, vol. 112, p. 106469, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0747563220302211

[6] D.-J. Kim, R. Hazlett-Knudsen, H. Culver-Godfrey, G. Rucks, T. Cunningham, D. Portee, J. Bricout, Z. Wang, and A. Behal, "How Autonomy Impacts Performance and Satisfaction: Results From a Study With Spinal Cord Injured Subjects Using an Assistive Robot," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 42, no. 1, pp. 2–14, 2012.

[7] F. F. Goldau and U. Frese, "Learning to map degrees of freedom for assistive user control: Towards an adaptive dof-mapping control for assistive robots," in *The 14th PErvasive Technologies Related to Assistive Environments Conference*, ser. PETRA 2021. New York, NY, USA: Association for Computing Machinery, 2021, p. 132–139. [Online]. Available: https://doi.org/10.1145/3453892.3453895

[8] K. Kronhardt, S. Rübner, M. Pascher, F. Goldau, U. Frese, and J. Gerken, "Adapt or perish? exploring the effectiveness of adaptive dof control interaction methods for assistive robot arms," *Technologies*, vol. 10, no. 1, 2022.

[9] M. Walker, H. Hedayati, J. Lee, and D. Szafir, "Communicating robot motion intent with augmented reality," in *Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '18. New York, NY, USA: Association for Computing Machinery, 2018, p. 316–324. [Online]. Available: https://doi.org/10.1145/3171221.3171253

[10] M. Pascher, U. Gruenefeld, S. Schneegass, and J. Gerken, "How to communicate robot motion intent: A scoping review," in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems - CHI '23*, 2023. [Online]. Available: https://doi.org/10.1145/3544548.3580857

[11] A. Cleaver, D. V. Tang, V. Chen, E. S. Short, and J. Sinapov, "Dynamic path visualization for human-robot collaboration," in *Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '21 Companion. New York, NY, USA: Association for Computing Machinery, 2021, p. 339–343. [Online]. Available: https://doi.org/10.1145/3434074.3447188

[12] N. J. Hetherington, E. A. Croft, and H. M. V. der Loos, "Hey robot, which way are you going? Nonverbal motion legibility cues for human-robot spatial interaction," *IEEE Robotics and Automation Letters*, vol. 6, no. 3, pp. 5010–5015, jul 2021. [Online]. Available: https://doi.org/10.1109%2Flra.2021.3068708

[13] R. T. Chadalavada, H. Andreasson, M. Schindler, R. Palm, and A. J. Lilienthal, "Bi-directional navigation intent communication using spatial augmented reality and eye-tracking glasses for improved safety in human–robot interaction," *Robotics and Computer-Integrated Manufacturing*, vol. 61, p. 101830, 2020.

[14] M. Pascher, K. Kronhardt, T. Franzen, U. Gruenefeld, S. Schneegass, and J. Gerken, "My caregiver the cobot: Comparing visualization techniques to effectively communicate cobot perception to people with physical impairments," *Sensors*, vol. 22, no. 3, p. 755, 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/3/755

[15] Z. Makhataeva and H. A. Varol, "Augmented reality for robotics: A review," *Robotics*, vol. 9, no. 2, 2020. [Online]. Available: https://www.mdpi.com/2218-6581/9/2/21

[16] L. V. Herlant, R. M. Holladay, and S. S. Srinivasa, "Assistive teleoperation of robot arms via automatic time-optimal mode switching," in *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*, ser. HRI '16. IEEE Press, 2016, p. 35–42.

[17] M. C. Shrestha, A. Kobayashi, T. Onishi, E. Uno, H. Yanagawa, Y. Yokoyama, M. Kamezaki, A. Schmitz, and S. Sugano, "Intent communication in navigation through the use of light and screen indicators," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2016, pp. 523–524.

[18] I. Shindev, Y. Sun, M. Coovert, J. Pavlova, and T. Lee, "Exploration of intention expression for robots," in *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '12. New York, NY, USA: Association for Computing Machinery, 2012, p. 247–248. [Online]. Available: https://doi.org/10.1145/2157689.2157778

[19] A. Singhal, "Modern information retrieval: A brief overview," *IEEE Data Eng. Bull.*, vol. 24, no. 4, pp. 35–43, 2001.

[20] S. G. Hart, "Nasa-Task Load Index (NASA-TLX); 20 Years Later," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 50, no. 9, pp. 904–908, 2006.

[21] J. Schmidtler, K. Bengler, F. Dimeas, and A. Campeau-Lecours, "A questionnaire for the evaluation of physical assistive devices (quead): Testing usability and acceptance in physical human-robot interaction," in *2017 ieee international conference on systems, man, and cybernetics (smc)*. IEEE, 2017, pp. 876–881.

[22] G. Harboe and E. M. Huang, "Real-world affinity diagramming practices: Bridging the paper-digital gap," in *Proc. 33rd Annual ACM Conf. Human Factors in Computing Systems*. ACM, 2015, pp. 95–104.

[23] D. C. Hoaglin and B. Iglewicz, "Fine-tuning some resistant rules for outlier labeling," *Journal of the American Statistical Association*, vol. 82, no. 400, pp. 1147–1149, 1987.

[24] S. Arévalo-Arboleda, F. Ruecker, T. Dierks, and J. Gerken, "Assisting manipulation and grasping in robot teleoperation with augmented reality visual cues," in *CHI Conference on Human Factors in Computing Systems (CHI '21)*. ACM, 2021. [Online]. Available: https://hci.w-hs.de/pub_VisualCuesCHI_compressed/

[25] S. Arévalo Arboleda, T. Dierks, F. Rücker, and J. Gerken, "Exploring the visual space to improve depth perception in robot teleoperation using augmented reality: The role of distance and target's pose in time, success, and certainty," in *Human-Computer Interaction – INTERACT 2021*, C. Ardito, R. Lanzilotti, A. Malizia, H. Petrie, A. Piccinno, G. Desolda, and K. Inkpen, Eds. Cham: Springer International Publishing, 2021, pp. 522–543.

[26] D. Burkolter, B. Weyers, A. Kluge, and W. Luther, "Customization of user interfaces to reduce errors and enhance user acceptance," *Applied Ergonomics*, vol. 45, no. 2, Part B, pp. 346–353, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0003687013000902

[27] E. Matsas and G.-C. Vosniakos, "Design of a virtual reality training system for human–robot collaboration in manufacturing tasks," *International Journal on Interactive Design and Manufacturing (IJIDeM)*, vol. 11, no. 2, pp. 139–153, Feb. 2015.

[28] S. L. Müller, S. Schröder, S. Jeschke, and A. Richert, "Design of a robotic workmate," in *Lecture Notes in Computer Science*. Springer International Publishing, 2017, pp. 447–456.

[29] E. Tidoni, M. Abu-Alqumsan, D. Leonardis, C. Kapeller, G. Fusco, C. Guger, C. Hintermuller, A. Peer, A. Frisoli, F. Tecchia, M. Bergamasco, and S. M. Aglioti, "Local and remote cooperation with virtual and robotic agents: A p300 BCI study in healthy and people living with spinal cord injury," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 9, pp. 1622–1632, Sept. 2017.

Exploring of Discrete and Continuous Input Control for AI-enhanced Assistive Robotic Arms

Max Pascher
max.pascher@udo.edu
TU Dortmund University
Dortmund, Germany
University of Duisburg-Essen
Essen, Germany

Kevin Zinta
kevin.zinta@studmail.w-hs.de
Westphalian University of Applied Sciences
Gelsenkirchen, Germany

Jens Gerken
jens.gerken@udo.edu
TU Dortmund University
Dortmund, Germany

ABSTRACT

Robotic arms, integral in domestic care for individuals with motor impairments, enable them to perform Activities of Daily Living (ADLs) independently, reducing dependence on human caregivers. These collaborative robots require users to manage multiple Degrees-of-Freedom (DoFs) for tasks like grasping and manipulating objects. Conventional input devices, typically limited to two DoFs, necessitate frequent and complex mode switches to control individual DoFs. Modern adaptive controls with feed-forward multi-modal feedback reduce the overall task completion time, number of mode switches, and cognitive load. Despite the variety of input devices available, their effectiveness in adaptive settings with assistive robotics has yet to be thoroughly assessed. This study explores three different input devices by integrating them into an established XR framework for assistive robotics, evaluating them and providing empirical insights through a preliminary study for future developments.

CCS CONCEPTS

• **Computer systems organization** → **Robotic control**; • **Human-centered computing** → *Visualization techniques*; *Virtual reality*.

KEYWORDS

assistive robotics, human-robot interaction (HRI), shared user control, virtual reality, visual cues

ACM Reference Format:
Max Pascher, Kevin Zinta, and Jens Gerken. 2024. Exploring of Discrete and Continuous Input Control for AI-enhanced Assistive Robotic Arms. In *Companion of the 2024 ACM/IEEE International Conference on Human-Robot Interaction (HRI '24 Companion), March 11–14, 2024, Boulder, CO, USA.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3610978.3640626

1 INTRODUCTION

The progress in the development of (semi-)autonomous technologies compelled their incorporation into numerous sectors, reshaping how we live and work. This integration includes scenarios of close collaboration with robotic devices, ranging from industrial assembly lines [4] to personal mobility aids [7]. Among these collaborative technologies, assistive robotic arms emerge as a particularly valuable and versatile subset, finding applications across various domains (e.g., [3, 26]).

Assistive robotic arms can enhance the independence of individuals with restricted mobility [15, 21]. These technologies – particularly when integrated with Artificial Intelligence (AI) – empower individuals to perform Activities of Daily Living (ADLs), which often entail tasks like gripping and manipulating objects in their surroundings, without reliance on human assistance [24]. However, current Human-Robot Interaction (HRI) research underscores a notable challenge faced by developers: optimizing the autonomy level of assistive robots [16]. Striking a balance is crucial, as purely autonomous systems may diminish user interaction and trust, while manual controls could prove impractical for users with specific impairments [12, 25, 31]. Shared control – combining manual input with algorithmic assistance – emerges as such a balanced approach and a promising research direction.

In this work, we explore three different input devices for controlling an assistive robotic arms in shared control applications:

- *Joy-Con*: A motion controller with continuous data input, suited for one-handed operation.
- *Head*: User control input by head-based movements, using continuous data.
- *Button*: A set of assistive buttons to control the robot in an accessible manner with discrete input data.

2 RELATED WORK

Standard control devices with a high Degree-of-Freedom (DoF), like gaming joysticks and keyboards, often pose challenges for users with severe motor impairments. Addressing these issues requires alternative solutions, such as specialized training or different interfaces [8, 28]. An approach proposed by Herlant et al. addresses these challenges by reducing the number of DoFs through mode switches. In their successful implementation, a joystick was used to control a *Kinova Jaco* assistive robotic arm [10].

Alternatively, Arévalo-Arboleda et al. introduced a hands-free multi-modal interaction by combining head movements, using a head-gaze based cursor to point, and speech commands to execute specific actions for tele-operating a robotic arm [2]. However, while speech commands provide enhanced accessibility, challenges like environmental noise or speech impairments encounter, impacting their effectiveness [18].

The control of assistive robotic arms involves a wide array of possible input devices, each targeted to suit the preferences and

HRI '24 Companion, March 11–14, 2024, Boulder, CO, USA

ACM ISBN 979-8-4007-0323-2/24/03.
https://doi.org/10.1145/3610978.3640626

capabilities of the respective user [1]. Despite this diversity, there remains a gap in the evaluation of these input devices within the context of AI-enhanced shared control applications for assistive robots.

In previous research, we introduced the *AdaptiX* framework, an open-source XR tool designed for Design and Development (D&D) operations [22]. *AdaptiX* consists of a Virtual Reality (VR) simulation environment to prepare and test study settings as well as a Robot Operating System (ROS) interface to control a physical robotic arm. The framework also includes a general input adapter, facilitating the development and evaluation of different input technologies and devices. Leveraging these capabilities, *AdaptiX* is used as the basis for this research project.

Through an algorithmic approach, the robotic arm's DoFs are configured to enable precise control with a low-DoF input device. This adaptive DoF mapping, denoted as *Adaptive DoF Mapping Control (ADMC)*, aims to present the user with a set of DoF mappings, organized based on their effectiveness in executing the pick-and-place task employed in the experiment (*optimal suggestion*, *adjusted/orthogonal suggestion*, *translation-only*, *rotation-only*, and *gripper*). The underlying concept of "usefulness" posits that optimizing the cardinal DoFs of the robot aligned with an input DoF while advancing towards the next goal represents the most advantageous approach.

3 DISCRETE AND CONTINUOUS CONTROL METHODS

Owning to *AdaptiX*'s integration of ADMC, users control the robotic arm forwards or backwards along a defined path based on the DoF mapping. Consequently, only a single-DoF input device is necessary for the movement. To choose from the different DoF mapping suggestions of the system, an additional one-dimensional input is required to perform a *mode switch* action, providing flexible and efficient control of the robotic arm.

Expanding upon the functionalities of *AdaptiX*, this study focuses on discrete and continuous control methods serving as assistive input devices for the ADMC shared control application. The framework's general input adapter provides a *float* value (-1.000 – 1.000) for the *Adaptive Axis* and a Boolean trigger for *Switch Mode*.

3.1 Motion Controller

Prior studies [14, 23] used a *Meta Quest* motion controller to interact with the *AdaptiX* framework. To add to this, we integrated a *Nintendo Joy-Con* [20], which is well suited for one-handed operation. For the integration, we used *UE4-JoyConDriver* [6] – a plugin for *Unreal Engine 4.27/5.2*. The plugin creates a connection between *Unreal* and *Nintendo Joy-Con* and provides sensor data such as accelerometer, gyroscope and Inertial Measurement Units (IMUs).

The left controller was selected for its balanced layout, accommodating both left- and right-handed users. The thumbstick – providing continuous data – was tilted up or down to move the robotic arm forward or backward. The mode switch is performed by pressing the *Up*-button of the controller right beneath the thumbstick. This design ensures single-handed control of the robot while preventing simultaneous movement and mode switching for enhanced usability.

3.2 Head-based Control

This control method eliminates the need for extra, specialized input devices as it utilizes orientation data from a device the user is already using – the Head-Mounted Display (HMD) [29]. Furthermore, it offers an accessible approach by allowing users with impaired hand motor function to operate the robotic arm.

The HMD's internal sensor technology, specifically the IMUs, facilitates the measurement of head rotations along three axes (*roll*, *pitch*, and *yaw*). This coordinate system is anchored to the object, positioned at the center of the user's head. Positive and negative rotations are possible around each axis, facilitating the mapping of six distinct actions to the corresponding axis rotations.

When the user tilts their head in a positive manner (*pitch*; rotating the head upwards), the robotic arm is advanced along the DoF mapped trajectory. Conversely, tilting the head in the opposite direction causes the arm to move backward along that path. Rolling the head to the right triggers the mode switch action, selecting the next ADMC suggestion.

Along each head rotation axis, a 20° resting zone has been set to prevent unintentional controlling of the robot. In this application, the user's head serves as a continuous data source for controlling the robot, akin to a joystick or the *Joy-Con's* thumbstick.

3.3 Assistive Buttons

Integrating the *Microsoft Xbox Adaptive Controller* [19], emphasizing flexibility and accessibility, enables the use of assistive buttons (e.g., *Logitech Adaptive Gaming Kit* [17]). These can be quickly and flexibly arranged to ensure comfortable operation by the user.

Similar to a gamepad control for discrete input data, the elementary actions for moving forward and backwards are mapped onto the adaptive buttons. The buttons marked *Arrow up* and *Arrow down* are mapped for moving the robotic arm, while a button with an *A*-marking was assigned to the mode switch.

4 STUDY

This preliminary study gathered initial user experiences with different modalities and operating modes for AI-enhanced assistive robotic arms. Through a controlled Mixed Reality (MR) user study involving 14 participants (6 female, 8 male), we systematically compare the advantages and disadvantages of the selected input methods. Four participants had prior experience with robotic arms.

4.1 Study Design

We employed a within-participant experimental design, with the *control method* as the independent variable, comprising three conditions: (1) *Joy-Con*, (2) *Head*, and (3) *Button*. Each participant underwent eight trials per condition. To mitigate the potential impacts of learning and fatigue, the condition order was fully counterbalanced.

4.2 Apparatus

Our study used the *AdaptiX* [22] framework to integrate and assess the selected control methods. We operated the framework in its MR mode, employing the *Varjo XR-3* [29] HMD and a *Kinova Jaco 2* [13] assistive robotic arm, as shown in Figure 1. We connected the *Varjo XR-3* and all input devices to a *Schenker Media Station* computer to facilitate this setup. Furthermore, we established connections between the *Schenker Media Station*, the ROS server, and the *Kinova Jaco 2* through a wired Local Area Network (LAN).

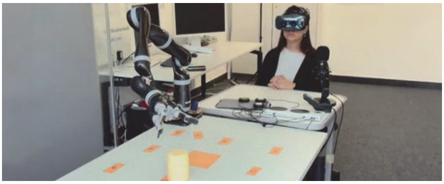


Figure 1: Overview of the study setup. The participant is wearing a *Varjo XR-3* HMD and controls the *Kinova Jaco 2* via head movements. The goal is to grasp the light-colored rounded block and place it on the large orange square in the middle of the table. The small orange markings are potential starting points for the rounded block.

4.3 Procedure

Before starting, participants received a detailed explanation of the project's objectives and the tasks involved. Each participant provided informed consent for their participation, including recording video, audio, and any other relevant data. A study administrator, overseeing the experiment on a laptop, provided instructions on using the hardware and the study environment. Once set up, participants followed command prompts within the MR environment. For each of the three conditions, the following steps were performed:

(1) Participants were given a written and standardized explanation of the control method used in the current condition.
(2) Participants conducted eight trials, grasping the object and placing it on the target surface.
(3) Interview and questionnaires.

After completing all conditions, participants ranked the three control methods from *most* to *least preferred* and explained their decision. The study concluded with a de-briefing.

4.4 Experimental Task

The experimental procedure builds on prior research that employed the *AdaptiX* framework (refer to [23]). The present study expands the configuration to a real-world environment, replicating a typical pick-and-place scenario.

To commence each trial, the study administrator positioned an object on a table. The participant aimed to navigate the robot from its initial location to grasp the object and deposit it onto a designated target area on the same table. For each trial, the object's starting position varied among eight possible predetermined locations. These positions were randomized in their sequence. We employed uniform rounded block shapes as objects to ensure impartiality and trial comparability, eliminating bias and allowing for consistent trial comparisons. Users could adjust the robot's DoF mapping by toggling between modes to fulfill the task. Following a successful execution, the object was removed, and the robot returned to its initial position. The object was then placed in a new starting position for a subsequent trial to begin. Upon completing each condition, we assessed workload using the NASA Raw-Task Load Index (Raw-TLX) questionnaire [9] and measured the five dimensions of the Questionnaire for the Evaluation of Physical Assistive Devices (QUEAD) [27]. The task completion time was recorded from the moment the participant initiated the movement of the robotic arm until the block was successfully placed.

5 RESULTS

This research focused on collecting subjective feedback from participants to improve the future development and integration of control input methods for shared control applications. The presented study encompasses a total of 336 (14 participants × 3 control methods × 8 trials) measured trials.

5.1 Perceived workload

Raw-TLX [9] scores [scale from 1 to 100] for all participants resulted in mean task load values of *Button* = 35.90 (SD = 12.98), *Joy-Con* = 41.17 (SD = 18.66), and *Head* = 59.65 (SD = 19.64). We applied a Friedman test which revealed a significant main effect for perceived task load ($\chi^2(2) = 18.00$, $p \leq 0.001$ ***, N = 14). The post-hoc pairwise comparisons (Bonferroni corrected) using Wilcoxon signed-rank tests revealed significant differences between *Head* and *Button* (Z = −3.27, $p \leq 0.001$ ***, r = 0.67), *Head* and *Joy-Con* (Z = −3.02, p = 0.002 **, r = 0.62), but not between *Button* and *Joy-Con* (Z = −1.44, p = 0.487, r = 0.29). The resulting task load scores per individual dimension of the Raw-TLX are presented in Figure 2.

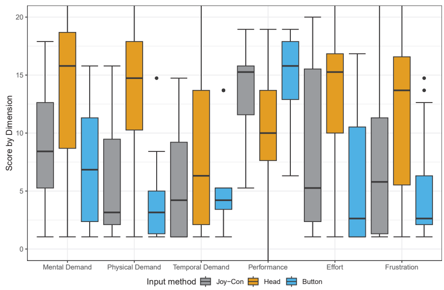


Figure 2: Comparison of the task load dimensions for the three different control methods: *Joy-Con, Head,* and *Button*

5.2 Evaluation of Physical Assistive Devices

The QUEAD included five individual scales (7-point Likert). Friedman tests for individual dimensions revealed significant main effects for all dimensions. Post-hoc pairwise comparisons indicate significant differences between *Head* and *Button* for all five dimensions as well as between *Head* and *Joy-Con* for *PU*, *PEU*, *E*, and *C*. For *Joy-Con* and *Button* only *PU* and *PEU* show significant differences (refer to Table 1 for detailed scores).

Table 1: Statistics for individual QUEAD dimensions: Perceived Usefulness (PU), Perceived Ease of Use (PEU), Emotions (E), Attitude (A), and Comfort (C).

	PU	PEU	E	A	C
Descriptive Statistics: mean value (standard deviation)					
Joy-Con	4.84 (1.18)	5.11 (1.55)	5.19 (1.41)	4.46 (2.16)	5.71 (1.87)
Head	2.93 (1.28)	3.37 (1.54)	2.86 (1.67)	2.82 (2.24)	3.57 (1.96)
Button	5.63 (1.44)	6.06 (1.15)	5.90 (1.22)	5.61 (1.84)	5.79 (1.46)
Friedman Tests					
$\chi^2(2)$	20.48	13.00	16.57	9.69	11.78
p	≤0.001 ***	0.002 **	≤0.001 ***	0.008 **	0.003 **
N	14	14	14	14	14
Pairwise Comparisons					
Joy-Con vs. *Head*					
$\lvert Z\rvert$	2.99	2.42	2.58	1.70	2.88
p	0.004 **	0.040 *	0.023 *	0.282	0.006 **
r	0.56	0.46	0.49	0.32	0.54
Head vs. *Button*					
$\lvert Z\rvert$	3.27	3.17	3.11	3.02	2.97
p	≤0.001 ***	0.001 ***	0.002 **	0.003 **	0.006 **
r	0.62	0.60	0.59	0.57	0.56
Joy-Con vs. *Button*					
$\lvert Z\rvert$	2.58	2.39	1.79	1.54	0.16
p	0.023 *	0.044 *	0.227	0.382	≥0.999
r	0.49	0.45	0.34	0.29	0.03

5.3 Individual Ranking

All participants – except one – ranked conditions from 1 = *favorite* to 3 = *least favorite*. Mean values in ascending order are *Button* = 1.46 (SD = 0.52); *Joy-Con* = 1.69 (SD = 0.63); and *Head* = 2.85 (SD = 0.55). A Friedman test revealed a significant main effect ($\chi^2(2) = 14.31$, p =≤0.001 ***, N = 13). The post-hoc pairwise comparisons indicate significant differences between *Head & Button* (Z = 3.02, p = 0.005 **, r = 0.59) and *Head & Joy-Con* (Z = 2.52, p = 0.026 *, r = 0.49), but not between *Button & Joy-Con* (Z = −0.83, p ≥0.999, r = 0.16).

5.4 Subjective Feedback

Participants noted an increased mental workload during the *Head*-based interaction. P01 highlighted that the movement execution for "forward felt opposite to the suggested arrow direction". Additionally, P01 got quickly distracted by a conversation with the experimenter, and P02 required substantial assistance due to difficulties in perceiving the arrows and mapping them to the head-movement direction. Participants P01 – P04 suggested introducing an additional mode switch to display the previous suggestion rather than presenting the next one. Participants P04, P11, and P12 preferred a non-continuous control by moving the head (i.e., only stop and go) to "prevent unintentional robot control when returning their head to the zero position" (P11).

Similar to the *Head*-based interactions, participants P01 – P04 mentioned a discrepancy between the suggested arrows by the system and the control input. In certain situations, the system suggests movements in the user's direction. To move the robot along this trajectory (*forwards*), the thumbstick of the *Joy-Con* or *Arrow Up* assistive button had to be pressed, which felt "discrepant". Participant P04 suggested using the thumbstick of the *Joy-Con* instead of the selected button for mode switching, for example, by tilting it sidewards.

Additionally, it was observed that specific initial placements of the object were perceived as disadvantageous compared to others, as the robot is fixed in place and has to perform – for the novice users – un-legible movements to reach the target.

6 DISCUSSION

All participants were able to control the robotic arm with each input device to fulfill the project task. Yet, the study's findings indicate that the effectiveness of the *Head*-based interaction method for controlling the robotic arm is relatively low compared to both hand-operated input methods. A notable insight derived from these results is the potential issue of the *Varjo XR-3* HMD being too bulky and heavy for sustained and precise *Head*-based control. To address these concerns, a more lightweight and comfortable solution, such as utilizing external IMUs for *Head*-based interaction [11, 30], could be considered.

Nevertheless, the HMD remains essential for visualizing directional cues, even with the integration of IMUs. Looking forward, advancements in technology are expected to yield significantly more compact and lighter devices, thereby enhancing user comfort and immersion.

Further, participants pointed out a discrepancy between the robot's movement direction and the mapping of user inputs. This could lead to an unclear mental model, particularly since the robot is controlled in a *first-person view*. To counteract this issue, a more extensive familiarization phase might be beneficial.

7 CONCLUSION

The input methods *Joy-Con* and *Button* represent promising approaches for controlling a robotic arm in a shared control application. Notably, both hand-operated input methods – irrespective of whether they provide discrete or continuous input data – (1) reduced perceived user workload and (2) improve *Perceived Usefulness*, *Perceived Ease of Use*, *Emotions*, and *Comfort*. These findings hold valuable implications for HRI researchers involved in the design of input technologies for assistive robotic arms. Future research efforts should prioritize the nuanced analysis of both quantitative and qualitative feedback obtained from focus groups. This comprehensive approach aims to refine and develop optimal methods for robot motion control, with the overarching goal of improving usability, safety, and end-user acceptance of these technologies.

Still, given the diverse likes and dislikes of the participants, future development of adaptive input control methods should – in line with Burkolter et al. – include individualization options to increase comfort and end-user acceptance [5].

ACKNOWLEDGMENTS

This research is supported by the *German Federal Ministry of Education and Research* (BMBF, FKZ: 16SV8565). Our study is approved by the Ethics Committee of the *Faculty of Business Administration and Economics of the University of Duisburg-Essen.*

REFERENCES

[1] Stephanie Arévalo-Arboleda, Max Pascher, Annalies Baumeister, Barbara Klein, and Jens Gerken. 2021. Reflecting upon Participatory Design in Human-Robot Collaboration for People with Motor Disabilities: Challenges and Lessons Learned from Three Multiyear Projects. In *The 14th PErvasive Technologies Related to Assistive Environments Conference - PETRA 2021* (2021-06-29). ACM. https://doi.org/10.1145/3453892.3458044

[2] Stephanie Arévalo-Arboleda, Franziska Ruecker, Tim Dierks, and Jens Gerken. 2021. Assisting Manipulation and Grasping in Robot Teleoperation with Augmented Reality Visual Cues. In *CHI Conference on Human Factors in Computing Systems (CHI '21)* (2021-05-03). ACM. https://doi.org/10.1145/3411764.3445398

[3] Maude Beaudoin, Josiane Lettre, François Routhier, Philippe S. Archambault, Martin Lemay, and Isabelle Gélinas. 2018. Impacts of robotic arm use on individuals with upper extremity disabilities: A scoping review. *Canadian Journal of Occupational Therapy* 85, 5 (2018), 397–407. https://doi.org/10.1177/0008417418820878 PMID: 30866682.

[4] Sara Bragança, Eric Costa, Ignacio Castellucci, and Pedro M. Arezes. 2019. *A Brief Overview of the Use of Collaborative Robots in Industry 4.0: Human Role and Safety.* Springer International Publishing, Cham, 641–650. https://doi.org/10.1007/978-3-030-14730-3_68

[5] Dina Burkolter, Benjamin Weyers, Annette Kluge, and Wolfram Luther. 2014. Customization of user interfaces to reduce errors and enhance user acceptance. *Applied Ergonomics* 45, 2, Part B (2014), 346–353. https://doi.org/10.1016/j.apergo.2013.04.017

[6] João Pedro Dante. 2024. UE4-JoyConDriver. Published online (last visited on January 16, 2024). https://github.com/jpdante/UE4-JoyConDriver

[7] Charles Fattal, Violaine Leynaert, Isabelle Laffont, Axelle Baillet, Michel Enjalbert, and Christophe Leroux. 2019. SAM, an Assistive Robotic Device Dedicated to Helping Persons with Quadriplegia: Usability Study. *International Journal of Social Robotics* 11, 1 (2019), 89–103. https://doi.org/10.1007/s12369-018-0482-7

[8] Peralta H., Monacelli E., Riman C., Baklouti M., Ben Ouezdou F., Mougharbel I., Laffont I., and Bouteille J. 2007. PLEIA: A Reconfigurable Platform for Evaluation of HCI Acting. In *Rehabilitation Robotics*, Sashi S Kommu (Ed.). IntechOpen, Rijeka, Chapter 23. https://doi.org/10.5772/5173

[9] Sandra G. Hart. 2006. Nasa-Task Load Index (NASA-TLX); 20 Years Later. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 50, 9 (2006), 904–908. https://doi.org/10.1177/154193120605000909

[10] Laura V. Herlant, Rachel M. Holladay, and Siddhartha S. Srinivasa. 2016. Assistive Teleoperation of Robot Arms via Automatic Time-Optimal Mode Switching. In *The Eleventh ACM/IEEE International Conference on Human Robot Interaction* (Christchurch, New Zealand) *(HRI '16)*. IEEE Press, 35–42.

[11] Anja Jackowski, Marion Gebhard, and Roland Thietje. 2018. Head Motion and Head Gesture-Based Robot Control: A Usability Study. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* 26, 1 (2018), 161–170. https://doi.org/10.1109/TNSRE.2017.2765362

[12] Dae-Jin Kim, Rebekah Hazlett-Knudsen, Heather Culver-Godfrey, Greta Rucks, Tara Cunningham, David Portee, John Bricout, Zhao Wang, and Aman Behal. 2012. How Autonomy Impacts Performance and Satisfaction: Results From a Study With Spinal Cord Injured Subjects Using an Assistive Robot. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans* 42, 1 (2012), 2–14. https://doi.org/10.1109/TSMCA.2011.2159589

[13] Kinova inc. 2021. Jaco assistive robot - User guide. Published online (last visited on January 16, 2024). https://assistive.kinovarobotics.com/uploads/EN-UG-007-Jaco-user-guide-R05.pdf EN-UG-007-r05-202111.

[14] Kirill Kronhardt, Stephan Rübner, Max Pascher, Felix Goldau, Udo Frese, and Jens Gerken. 2022. Adapt or Perish? Exploring the Effectiveness of Adaptive DoF Control Interaction Methods for Assistive Robot Arms. *Technologies* 10, 1 (2022), 24 pages. https://doi.org/10.3390/technologies10010030

[15] Maria Kyrarini, Fotios Lygerakis, Akilesh Rajavenkatanarayanan, Christos Sevastopoulos, Harish Ram Nambiappan, Kodur Krishna Chaitanya, Ashwin Ramesh Babu, Joanne Mathew, and Fillia Makedon. 2021. A survey of robots in healthcare. *Technologies* 9, 1 (2021), 8.

[16] Audrey Lebrasseur, Josiane Lettre, François Routhier, Philippe S. Archambault, and Alexandre Campeau-Lecours. 2019. Assistive robotic arm: Evaluation of the performance of intelligent algorithms. *Assistive Technology* 33, 2 (May 2019), 95–104. https://doi.org/10.1080/10400435.2019.1601649

[17] Logitech. 2024. Adaptive Gaming Kit. Published online (last visited on January 16, 2024). https://www.logitechg.com/en-us/products/gamepads/adaptive-gaming-kit-accessories.943-000318.html

[18] Thombare Mayuri, Dhawale Poonam, Amale Rohini, and L. Aishwarya. 2020. EMIVR Based Smart Computer. *International Journal of Scientific Research in Science and Technology* 7 (2020), 397–400. https://doi.org/10.32628/IJSRST207380

[19] Microsoft. 2024. Xbox Adaptive Controller. Published online (last visited on January 16, 2024). https://www.microsoft.com/en-us/d/xbox-adaptive-controller/8nsdbhz1n3d8

[20] Nintendo. 2024. Nintendo Joy-Con. Published online (last visited on January 16, 2024). https://www.nintendo.de/Hardware/Zubehor/Zubehor-1183144.html

[21] Max Pascher, Annalies Baumeister, Stefan Schneegass, Barbara Klein, and Jens Gerken. 2021. Recommendations for the Development of a Robotic Drinking and Eating Aid - An Ethnographic Study. In *Human-Computer Interaction – INTERACT 2021* (2021-09-01), Carmelo Ardito, Rosa Lanzilotti, Alessio Malizia, Helen Petrie, Antonio Piccinno, Giuseppe Desolda, and Kori Inkpen (Eds.). Springer, Cham. https://doi.org/10.1007/978-3-030-85623-6_21

[22] Max Pascher, Felix Ferdinand Goldau, Kirill Kronhardt, Udo Frese, and Jens Gerken. 2024. AdaptiX – A Transitional XR Framework for Development and Evaluation of Shared Control Applications in Assistive Robotics. *Proc. ACM Hum.-Comput. Interact.* 8, EICS (2024). Pripront on arXiv: https://arxiv.org/abs/2310.15887.

[23] Max Pascher, Kirill Kronhardt, Felix Ferdinand Goldau, Udo Frese, and Jens Gerken. 2023. In Time and Space: Towards Usable Adaptive Control for Assistive Robotic Arms. In *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 2300–2307. https://doi.org/10.1109/RO-MAN57019.2023.10309381

[24] Laura Petrich, Jun Jin, Masood Dehghan, and Martin Jagersand. 2022. A Quantitative Analysis of Activities of Daily Living: Insights into Improving Functional Independence with Assistive Robotics. In *2022 International Conference on Robotics and Automation (ICRA)*. 6999–7006. https://doi.org/10.1109/ICRA46639.2022.9811960

[25] Anita Pollak, Mateusz Paliga, Matias M. Pulopulos, Barbara Kozusznik, and Malgorzata W. Kozusznik. 2020. Stress in manual and autonomous modes of collaboration with a cobot. *Computers in Human Behavior* 112 (2020), 106469. https://doi.org/10.1016/j.chb.2020.106469

[26] Terrin Babu Pulikottil, Stefania Pellegrinelli, and Nicola Pedrocchi. 2021. A software tool for human-robot shared-workspace collaboration with task precedence constraints. *Robotics and Computer-Integrated Manufacturing* 67 (Feb. 2021), 102051. https://doi.org/10.1016/j.rcim.2020.102051

[27] Jonas Schmidtler, Klaus Bengler, Fotios Dimeas, and Alexandre Campeau-Lecours. 2017. A questionnaire for the evaluation of physical assistive devices (quead): Testing usability and acceptance in physical human-robot interaction. In *2017 ieee international conference on systems, man, and cybernetics (smc)*. IEEE, 876–881. https://doi.org/10.1109/SMC.2017.8122720

[28] Atieh Taheri, Ziv Weissman, and Misha Sra. 2021. Design and Evaluation of a Hands-Free Video Game Controller for Individuals With Motor Impairments. 3 (2021). https://doi.org/10.3389/fcomp.2021.751455

[29] Varjo. 2024. Varjo XR-3. Published online (last visited on January 16, 2024). https://varjo.com/products/xr-3/

[30] Lukas Wöhle, Stanislaw Miller, Jens Gerken, and Marion Gebhard. 2018. A Robust Interface for Head Motion based Control of a Robot Arm using MARG and Visual Sensors. In *2018 IEEE International Symposium on Medical Measurements and Applications (MeMeA)*. 1–6. https://doi.org/10.1109/MeMeA.2018.8438699

[31] Jakub Złotowski, Kumar Yogeeswaran, and Christoph Bartneck. 2017. Can we control it? Autonomous robots threaten human identity, uniqueness, safety, and resources. *International Journal of Human-Computer Studies* 100 (2017), 48–54.

UDE

People with motor impairments experience reduced mobility, social exclusion, and dependence on caregivers. Although assistive technologies have the potential to enhance independence and well-being, their development often overlooks user involvement. This oversight -- coupled with designs that limit user autonomy -- leads to unmet needs and increased stress for end users.

This thesis addresses current limitations by focusing on user integration and improving shared control approaches for assistive robotics enhanced by Artificial Intelligence (AI). Key contributions include identifying user needs, exploring robot motion intent communication, introducing the innovative Adaptive DoF Mapping Control (ADMC) shared control approach, and presenting the AdaptiX framework for developing and evaluating multi-modal interaction designs.

The effectiveness of ADMC and AdaptiX are demonstrated in real and simulated scenarios, emphasising user-centred design, AI-enhanced applications, and in-silico testing. This thesis also outlines future research opportunities to advance AI-enhanced assistive robotics, aiming for the full inclusion of people with physical impairments in social and professional spheres.

ISBN: 978-3-00079-648-7
9 783000 796487