



UMEÅ UNIVERSITY

Being connected to the world through a robot

Patrik Björnfot

Department of Informatics
Umeå 2022

This work is protected by the Swedish Copyright Legislation (Act 1960:729)
Dissertation for PhD
ISBN: 978-91-7855-821-6
ISBN: 978-91-7855-822-3
ISSN: 1401-4572 RR-22.01
Cover Art and Design by Hanna Nordin
Electronic version available at: <http://umu.diva-portal.org/>
Printed by: Xxxx Xxxxxxx
Umeå, Sweden 2022

”Jos met freistaama,

pruuaama, freistaama

kyllä se häätyy kannattaa!”

Torvald Pääjärvi - *”Ei se kannatte”*

Table of Contents

Abstract	iii
Acknowledgement	iv
Preface	vi
Prologue	vii
1 Introduction	1
1.1 Exploring the relationship between humans and the world through robotic telepresence robots	4
1.2 Structure of the Thesis	7
2 Introducing Robotic Telepresence Technology	9
2.1 Robotic Telepresence: The object of study	9
2.2 Overview of Robotic Telepresence systems	11
3 HCI and Robotic Telepresence	20
3.1 The Pilot's experience	20
3.2 The Local User's experience	25
3.3 Social Interaction	27
3.4 Summary	29
4 Theoretical Foundations	31
4.1 Activity Theory	32
4.2 Phenomenology and Postphenomenology	42
4.1 Activity Theory and Phenomenology	49
4.1 Summary of Theoretical Foundations	50
5 Research Design and Method	52
5.1 Methodological Considerations	52
5.2 Study Design and Methodology	54
5.1 Study Methods	56
6 Summary of Studies	62
6.1 Study 1	63
6.2 Study 2	65
6.3 Study 3	67
6.4 Study 4	68
6.5 Study 5	70
7 Remote Embodiment and Robotic Telepresence	72
7.1 Remote Embodiment	72
7.2 Operating Robotic Telepresence Technology	74
7.3 Being remotely embodied as a robot in a local environment	76
7.4 Social Remote Embodiment	80
7.5 The humans' connection to the world through remote embodiment	86
8 Possible Futures of Robotic Telepresence	89
8.1 Robotic telepresence is a multipurpose technology	89
8.2 Design for Remote Embodiment	96
8.3 Theory and Robotic telepresence	99
9 Conclusion	101
10 References	103

Abstract

Robotic telepresence systems enable humans to be present physically and socially in a distant environment. Robotic telepresence technology is the latest in the line of communication technology development. The unique feature of such technology is that its users can act in a distant environment and interact with other people through these systems. The robot is the user's physical avatar through which they act. This thesis aims to understand how people connect to the world through robotic telepresence. The aim includes addressing how humans operate the robotic telepresence system, how the robotic telepresence supports performing actions in a distant location and supports social interaction, and how a human experience being in a robotic body.

The thesis is based on five studies, reported in five papers, that explore different aspects of robotic telepresence. The theoretical foundations consist of activity theory and phenomenology, two traditions that are arguably compatible and complementary. The concept of remote embodiment is proposed to describe the relationship between the human and robotic telepresence systems. Remote embodiment is a phenomenon, design concept, and feature that enables robotic telepresence to be used in a wide variety of activities. Furthermore, I use the concept of remote embodiment to outline possible futures of robotic telepresence.

Acknowledgement

“No man is an island entire of itself; every man is a piece of the continent, a part of the main;” - Devotions upon Emergent Occasions, John Donne

It is truly a privilege to have devoted time to explore a subject to the extent that I have had the opportunity to do. During all this time, I have been surrounded by people who have, in different ways, helped me through this journey.

I want to express my gratitude to my main supervisor Victor Kaptelinin; thanks for all the discussions and collaborations; it has been inspiring to work with you.

I would also like to thank my other main supervisor Karin Danielsson; a special thanks for supporting me in finishing this thesis.

A big thanks to Rikard Harr, my co-supervisor and my office neighbor: you have been a great bollplank and listener to all my rants.

I would also like to direct a special thanks to Mikael Wiberg, who has been involved in the project that this Ph.D. has been a part of and also for serving as a panelist in my pre-seminar.

The feedback given by the panel during the pre-seminar was of great value to me. The panel consisted of Mikael (again), Teresa Almeida, Ott Velsberg, and John Waterworth; many thanks!

My process has been close to Robyn Schimmer's work; I am grateful that we followed each other through the years and had many great fikas, lunches, and beers.

This thesis was done in the Department of Informatics; I want to thank all of you for this time, and I look forward to future collaborations! In particular, I would like to thank all the Ph.D. students that I have had the pleasure of meeting; our community has been valuable to me.

Furthermore, I wish to thank the 86 participants that have been involved in my studies and pre-studies.

Jag vill också ta tillfället i akt att tacka min familj och mina vänner som jag fått äran att dela dessa år med, år som både varit fyllda av glädje och sorg, hopp och förtvivlan.

Jag vill tacka min älskade Therese Erixon för allt!

Citatet i början av denna avhandling är i mångt och mycket en dedikation till mina rötter i Tornedalen. Således är det också en dedikation till de tre i min familj som var med mig vid arbetets början men tyvärr inte kunde vara med till slutet: min mor Ruth Björnfot, min farfar Yrjö Klingestål och min farmor Alma Klingestål. Samtliga tre illustrerar på sitt egna sätt att om man försöker så lönar det sig i slutändan.

/Patrik

Preface

This dissertation is based on five studies reported in five appended papers.

Paper 1: Björnfot, P., & Kaptelinin, V. (2017, March). Probing the design space of a telepresence robot gesture arm with low fidelity prototypes. In *2017 12th ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 352-360). IEEE.

Paper 2: Björnfot, P., Bergqvist, J., & Kaptelinin, V. (2018). Non-technical users' first encounters with a robotic telepresence technology: an empirical study of office workers. *Paladyn, Journal of Behavioral Robotics*, 9(1), 307-322.

Paper 3: Björnfot, P. (2021, April). Evaluating Input Devices for Robotic Telepresence. In *European Conference on Cognitive Ergonomics 2021* (pp. 1-8).

Paper 4: Kaptelinin, V., Björnfot, P., Danielsson, K., & Wiberg, M. (2020, October). Performance, Power, and Place: User Experience of Contactless Object Manipulation in Robotic Telepresence. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society* (pp. 1-12).

Paper 5: Kaptelinin, V., Björnfot, P., & Danielsson, K. (2021, April). Exploring the Relationship Between Physical Presence, User Experience, and Task Parameters in Robotic Telepresence. In *European Conference on Cognitive Ergonomics 2021* (pp. 1-5).

Prologue

In general, digital technology has become increasingly intertwined with humans. The smartphone is so integrated into our everyday life that it has become a part of us. We use clothes to shield ourselves from the cold weather, and we use smartphones to connect to the world. When I was a child, *I was “on” the computer*, and *I was “on” the internet*: these were *states*. Being online, it is no longer a “state” but a part of my lifeworld. In the same way, as I never have “utilized electricity” as an “act,” I no longer go “online”, instead, I am accessible through this medium in the same sense that my attention is accessible if a person is in the same room as me.

I have been writing this thesis during the COVID-19 pandemic. Many of us have worked from home during this time, while others have carried out their work activities while maintaining a safe distance. The majority of the earth’s population has lived under restrictions to reduce the virus’s spread. The primary purpose of these restrictions is to achieve “social distancing.” However, we are not “socially distant” due to communications technology. Instead, we are “physically distant.” This thesis is about novel communication technology, robotic telepresence, and it is also about the social and physical aspects of digitally mediated interaction. Robotic telepresence technology is one of the latest technologies to reduce the social distance between physically distanced humans.

Robotic telepresence is also a solution for when people want to be at a particular location but cannot be there physically. In 2017, there was a ban on entering the USA for people with a passport from certain countries. This travel ban also affected people within the USA because they could not re-enter if they left the country. A consequence of this was that many people could not visit conferences. To mitigate this, some conferences, including two conferences that I attended (HRI’2017 and CHI’2017), offered attendance through robotic telepresence technology (“Special Statement,” .d.).

Due to the restrictions, academic conferences moved to a virtual world instead. It thus eliminated the need for a particular physical location. Working from home means that much of the previously location-dependent work has become location-independent. One wonders if this will ever return to normal, as people start to see clear benefits of not being tied to a particular place. Instead, we can choose our physical location based on aspects other than work.

Robotic telepresence beautifully combines being virtual and physical. Therefore, I think that insights into this technology might be one of the keys to understanding our future relationship with digital technology.

1 Introduction

In recent years, digital technology has become interwoven with our lives. We have an online presence that makes us constantly reachable by friends and acquaintances. However, the separation between the virtual and physical worlds remains, since our virtual actions only transfer to the physical world through human mediators. However, various remotely controlled devices, such as robotic telepresence systems, blur the virtual and physical division by enabling the users to engage in real-time with the physical world.



Figure 1: BEAM+ Telepresence Robot from SutableTech. The system used in studies.

The topic of this thesis is robotic telepresence systems (see Figure 1) that consist of remotely controlled robots that users can use to interact with people in a distant physical location. Thus, a remote user can log in to a robot in a distant place and interact with other people and the local environment. Robotic telepresence combines videoconference functionality and a physical mobile body.

Telepresence was firstly introduced by Minsky (1980), when he described the following scenario:

"You put on a comfortable jacket lined with sensors and muscle-like motors. Each motion of your arm, hand, and fingers is reproduced at another place by mobile, mechanical hands. Light, dexterous, and strong, these hands have their own sensors through which you see and feel what is happening. Using this instrument, you can "work" in another room, in another city, in another country, or on another planet. Your remote presence possesses the strength of a giant or the delicacy of a surgeon. Heat or pain is translated into informative but tolerable sensation. Your dangerous job becomes safe and pleasant." (Minsky, 1980).

Forty years later, many affordable commercial robotic telepresence systems are available. As with a human body, a typical robotic telepresence system (see Figure 1) has feet (wheeled base), a body (a pole), and a head (computer with a screen). The head consists of a camera, microphone and speaker, and the display shows the user's face via a webcam.

These robots have not yet reached the technological level of Minsky's vision. Instead, their design is similar to one of the first Robotic Telepresence systems, PROPs (Personal Roving Presences), developed in the 1990s (Paulos and Canny, 1998a). Paulus and Canny never intended for the PROp to be the type of telepresence system envisioned by Minsky. Instead, PROp resulted from an effort to support a limited set of human skills in a remote environment based on the technology at hand.

"Instead, PROPs attempt to achieve certain fundamental human skills without a human-like form. More importantly, our research is driven by the study and understanding of the social and psychological aspects of extended human-human interactions rather than the rush to implement current technological advances and attempt to re-create exact face-to-face remote human experiences." (Paulos and Canny, 1998b)

As this quote highlights, the replication of human skills was the focus of PROp. This approach was logical since Paulos and Canny had previously developed a telepresence system, built as a blimp, that floated around but lacked the general "fundamental human skills" (Paulos and Canny, 1997). Arguably, the approach and design of PROp were ultimately successful because investigations of current robotic telepresence systems for various activities yielded positive results. Such activities include office work (Lee and Takayama, 2011), medicine (O'Neill et al., 2001), academic conferences (Rae and Neustaedter, 2017), visiting museums (Claudio et al., 2017), education (Edwards et al., 2016), elderly care (Cesta et al., 2016), and even long-distance relationships (Yang et al., 2017). Additionally,

experimental uses such as shopping together (Yang et al., 2018) and geocaching (Heshmat et al., 2018) indicate further exciting potential. In other words, robotic telepresence has already been successfully appropriated into activities of various kinds.

Robotic telepresence is the first communication technology where the user controls an avatar in the physical world for social interaction. While there are various remotely controlled devices, such as flying planes and drones, underwater rovers, and remotely controlled cars, these do not afford social interaction. A robotic telepresence robot is both a "drone" and a representation of the user. The user is engaged with the world through a robotic body, excellently described in the article title "Now I have a Body" (Lee and Takayama, 2011). Compared to regular videoconference systems, the mobility provided by robotic telepresence systems increases the level of engagement; the user has freedom in exploring the remote location and can start a conversation with whoever happens to be around.

Robotic telepresence is also a novel experience for the people in the robots' environment (called local users) since most have never interacted with a human physically represented by a robot body that can move around independently.

Thus, as I write this, the technology is in its middle ground, that is, it is mature enough to be used but still so novel that we haven't yet see the real-world consequences of their use. Arguably, the technology's current state makes it a suitable subject for Human-Computer Interaction (HCI) and Interaction Design (IxD) research.

The field of HCI has a long history of understanding human and digital technology relationships. Stemming from computer science, HCI focuses on developing systems *for* humans (Carroll, 2013). The original focus of HCI researchers was the investigation of actual interactions and the human factors that limit interaction potential (Bannon, 1995). Gradually the focus of HCI shifted towards the use of digital technology for meaningful activities (Bannon, 1995) and user experience (Bødker, 2006).

Thus, the initial object of study in HCI was the relationship between humans and digital technology. As digital technology became increasingly integrated into our lives, the object of study expanded from human low-level interaction with digital technology to include aspects of how digital technologies affect our lifeworld. Thus, the development of HCI as a research field has co-evolved with digital technology. In other words, as digital technology evolves, new human-technology relationships emerge, resulting in the need for further theoretical development.

Furthermore, Informatics in general, and HCI in particular, have a long history of affecting technological development, including methods and processes for designing and evaluating designs. The relationship between research and design is an ongoing discussion, but it is evident that it has had a significant influence on HCI and related fields (Fallman, 2003; Hevner and Chatterjee, 2010; Hevner et al., 2004; Redström, 2017; Zimmerman et al., 2007).

This thesis acknowledges and employs the evolution of HCI as a research field and its relationship to design. Robotic telepresence is, in a way, "yet another technology" of interest for HCI. Furthermore, we can understand large parts of the robotic telepresence-human relationship through studies and theoretical frameworks developed for previously developed digital technologies.

However, understanding robotic telepresence also poses a significant challenge for HCI theory. The human-technology relationship has become increasingly entangled due to digital technologies such as virtual reality (Frauenberger, 2019). The robot mediates the users' engagement in the world, and the body becomes a *physical* and *social* representation of the user. Thus, the human-digital technology relationship includes both the direct *in situ* interaction and how digital technology mediates our activities and alters our experience.

1.1 Exploring the relationship between humans and the world through robotic telepresence robots

The relatively recent human ability to physically *act in remote places in real-time*, enabled by robotic telepresence and similar technologies, poses unique challenges to HCI research and requires a systematic empirical and conceptual analysis. Such an investigation aims to create an understanding that forms a basis for exploring the use and design space of robotic telepresence.

This thesis explores the complex and entangled relationship between humans and the world through robotic telepresence systems. The broad perspective on the human-technology relationship is motivated by the realization that the transformative power of digital technology is indisputable and technology changes our perception of the world; thus, it changes our lifeworld (Ihde, 1990).

HCI can utilize the philosophy of technology to find out what "good" technology is (Fallman, 2011), an inquiry that upon a particular existing technology should be presided by insights from the current human-technology relations. Robotic telepresence alters the users' lifeworld, and if the previously outlined visions of robotic telepresence become realized, this could happen for many thousands of people. The novel functionality of robotic telepresence makes it difficult to understand how the users' lifeworld would be affected.

Robotic telepresence's novel capabilities render it impossible to understand fully its relationship to humans by only extrapolating findings from previous human-technology relationships. Robotic telepresence is a mix of other well-researched technologies, such as drones and videoconferencing systems, whose history and insights are worthy of consideration. However, the similarities between these technologies are mainly limited to isolated details and not the higher-level features. Robotic telepresence enables humans to be both social and physical actors at a distant location. The dual nature of being both physical and social is comparable to being in an environment in person. But, it has an inherent difference in that some actors are present in their bodies while others are in their robots.

There has been a wide range of studies of robotic telepresence in HCI and related fields. The studies have produced a considerable body of empirical evidence that has yielded several important insights. However, there is a notable limitation to existing research. The research has mainly focused on particular issues and phenomena related to robotic telepresence technologies. In contrast, the understanding of the consequences of using the technology for human activity and experiences, in general, has been somewhat limited.

An underlying assumption of the study, reported in this thesis, is that various aspects of the design and use of robotic telepresence technology *may affect how people relate to and experience their engagement with the world*. This thesis is a step towards a systematic analysis of such effects as it presents and discusses a series of empirical studies exploring the impact of robotic telepresence technology design and its effect on users' activities and experiences. The overall aim is thus *to understand the role and impact of robotic telepresence technology in relation to humans*. Thus, at a high level of abstraction, the overarching research question in this thesis is:

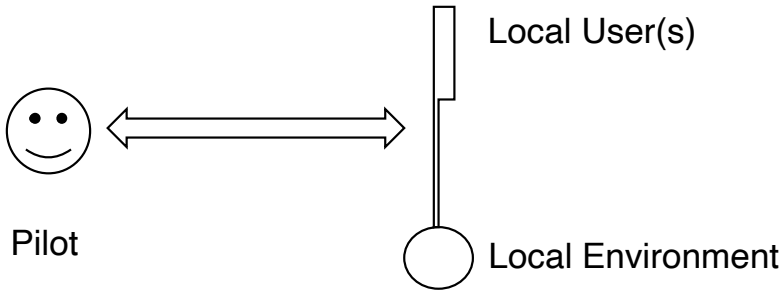
How does robotic telepresence influence humans' connection to the world?

The term “connection” is chosen to capture the holistic aim of the thesis. Arguably a fundament of the thesis is to understand various aspects of this “connection.”

The research question contains four sub-questions, based on the three main types of relationships: pilot and robot, pilot and local environment, and pilot and local users (Figure 2).

1) Operating Robot

4) Social Interaction



3) Being in the robots body

2) Actions in the Local environment

Figure 2: Illustration of the four sub-questions

1) How does the user operate the robot?

Operating the robot is a prerequisite for using the system in meaningful activities. How users operate digital technology is one of the first issues addressed by HCI (Bannon, 1995; Carroll, 2013). While insights from previous technologies are included in the designs of robotic telepresence systems, and are relevant when addressing the operation of a robotic telepresence system, it is essential to acknowledge the novelty of robotic telepresence. This novelty is the context surrounding robotic telepresence. In other words, it is possible to extrapolate lower-level interactions from earlier technologies but, due to the novel context of robotic telepresence, it is hard to estimate the consequences of real-world usage.

2) How does robotic telepresence support performing actions in a distant physical location?

One of the core functionalities of robotic telepresence systems is support the carrying out of actions in a distant location. Acting in a distant location is a relatively novel capability, especially for a general user group.

3) How does the pilot experience being in the robotic body?

Another unique aspect of robotic telepresence systems is that the user connects to the world through a physical robotic body. As a result, this robotic body is the representation of its user in a distant place. When using the system, the user *being* in the distant location is mediated by the robotic body, that is, the user is

in this body. Therefore, a pilot's experience of being in the robotic body is crucial in understanding how humans connect to the world through robotic telepresence.

4) *How does robotic telepresence support social interaction?*

One primary feature of robotic telepresence is mediating social interaction between the pilot and local user(s). How humans interact with each other is a vital part of their "connection to the world." The social interaction mediated through use of a robot is similar to other previous communication technologies. Still, the physical representation of a robot in combination with the acting capabilities of the pilot is novel and might affect social interaction.

This thesis utilizes two main perspectives on the relationship between human and robotic telepresence systems: activity theory (Kaptelinin and Nardi, 2006; Leontiev, 1977) and phenomenology (Dourish, 2001; Ihde, 1990). The activity theory perspective focuses on how technology helps humans reach their motives, while the phenomenological perspective focuses on an individual's experience of technology.

By addressing the research questions, I aim to contribute to the knowledge about the relationships between humans and telepresence robots as well as the future design and usage of these systems. In other words, this thesis aims at contributing to both research and design relating to robotic telepresence and HCI.

1.2 Structure of the Thesis

The rest of the thesis is structured as follows. The following chapter, "Introducing Robotic Telepresence," provides an overview of robotic telepresence technology from a commercial and research point of view.

The chapter is followed by "HCI and Robotic Telepresence," where I address related research on robotic telepresence from an HCI perspective.

In the chapter Theoretical Foundations, I outline theoretical concepts used in this thesis. In short, I use *activity theory* to understand the structure and dynamics of activities mediated by the robotic telepresence system and *phenomenology* to understand users' experiences.

The Research Design and Method chapter outlines the overarching methodological aspects considered for addressing the research question and sub-questions. The chapter also outlines how the five studies, each presented in individual papers, relate to the main research question and their study design and methodological considerations.

The Summary of Studies chapter describes the five studies that this thesis consists of:

Study 1. “Probing the design space of a telepresence robot gesture arm with low fidelity prototypes.”

Study 2. *“Non-technical users’ first encounters with a robotic telepresence technology: An empirical study of office workers.”*

Study 3. *“Evaluating Input Devices for Robotic Telepresence.”*

Study 4. *“Performance, Power, and Place: User Experience of Contactless Object Manipulation in Robotic Telepresence”*

Study 5. *“Exploring the Relationship Between Physical Presence, User Experience, and Task Parameters in Robotic Telepresence”*

The Research Design and Method chapter addresses the four sub-questions based on the five studies. The chapter Remote Embodiment and Robotic Telepresence presents a discussion of the main research question using the theoretical foundations. In chapter Possible Futures of Robotic Telepresence I use activity theory to describe the complexity of robotic telepresence as a factor of its multi-purpose nature and phenomenology to describe the concept of remote embodiment. The cover paper ends with Conclusion.

2 Introducing Robotic Telepresence Technology

This chapter introduces robotic telepresence technology by discussing its general functionality and outlining the current design and use space.

2.1 Robotic Telepresence: The object of study

Robotic telepresence systems have two main user groups, *pilot* and *local users*¹, whose interactions with each other are mediated by a *robot* (see Figure 3). The *pilot* controls the robot from a *remote environment*, and the *local user* is physically present in the *local environment*.

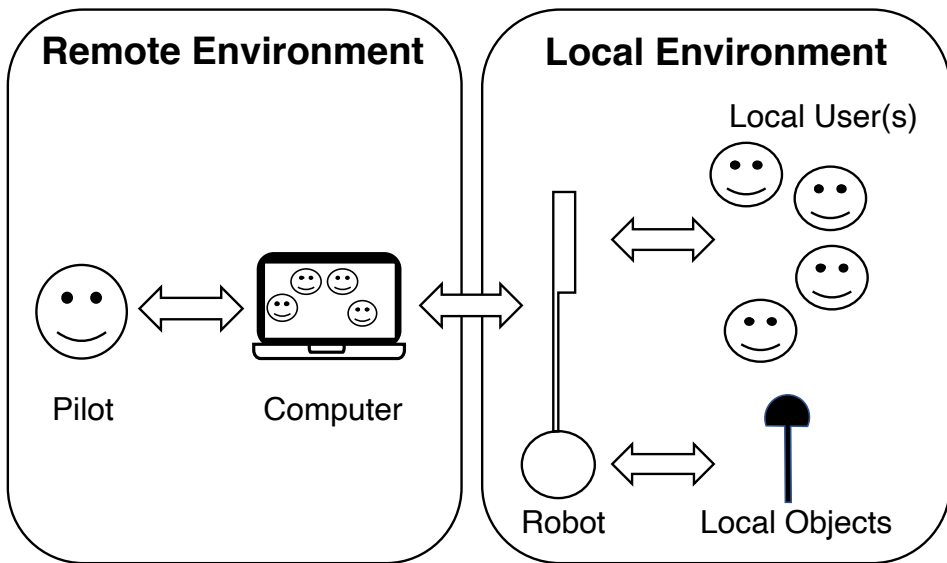


Figure 3: Robotic Telepresence and humans

The *remote environment* can be anywhere the pilot can set up the necessary equipment and internet connection. The *local environment* is an environment that offers connectivity and accessibility for the robot. In general, telepresence robots are limited to navigation in indoor environments that lack stairs and other floor-level obstacles.

The pilot controls the robot using either a computer or a handheld device (pad or smartphone). The interaction is both directed to and through the robot.

¹ Terms based on Kristoffersson et al. (2013)

Interactions directed towards the robot are mainly concerned with changing the settings, such as altering volume and speed.

The interaction that is focused on the local environment is limited to moving and looking around; the pilot can also move some light objects by bumping into them, but this is not encouraged by the manufacturers.

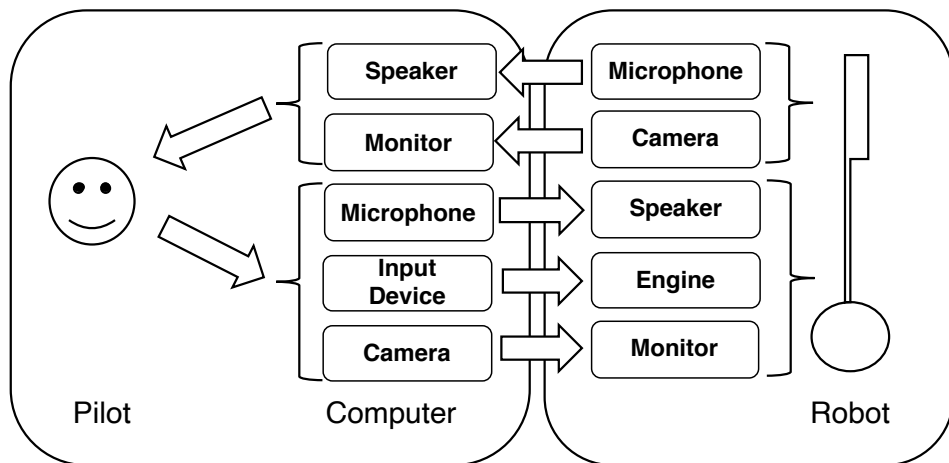


Figure 4: Input and output sub-devices

Interactions between the pilot and the local user occur through speakers, microphones, cameras and screens, so the same types of interaction as experienced during a normal videoconference call (see Figure 4). The pilot can produce a limited set of body language movements, such as rotating the robot and independently moving around in the local environment.

The local user has some control over the robot, for example, they can choose to disable it and can intervene physically. The degree of interaction possibilities that the local users have over the robot varies between systems. In some, the local user needs to answer calls to the robot from the pilot for a connection to be established² and charge it while, in other systems, no intervention from the local user is necessary. Some systems allow the local user to change settings such as sound level.

Scoping the object of study

The robotic telepresence systems discussed in this thesis support social interaction between a pilot and a local user, provide independence of movements

² Similar procedure as establishing a regular phone call.

to the pilot, and resemble a human in size, as with the types of robots shown in Figure 1.

This scope excludes multiple types of telepresence systems, including table-based robots such as KUBI (“KUBI Telepresence Robot | Welcome,” n.d.), flying drone-like systems, small robotic telepresence systems such as MeBot, and humanoid robots that often have telepresence qualities used for research purposes (Thellman et al., 2017).

2.2 Overview of Robotic Telepresence systems

This section outlines the research and design of robotic telepresence systems. The data sources are both from commercial communications and academic research. The purpose of both datasets is to provide an overview of the current design and use space of robotic telepresence systems. Using both commercial and research data for this kind of review is in line with the applied nature of HCI and has previously been used for reviewing robotic telepresence (see Kristoffersson et al., 2013).

Table 1: Summary of commercial Robotic Telepresence systems

Robot Model	Reference	Height (cm)	Weight (kg)
Double 3	(“Double Robotics,” n.d.)	119–150*	7.3
Ohmni Supercam	(“Learn How to Control Ohmni,” n.d.)	142	9
Beam Pro	(“Beam Quickstart Guide,” n.d.)	158	41
Boteyes Pro	(“Telepresence Robot BotEyes,” n.d.)	110	11
Vgo	(“VGo Solution User Guide,” n.d.)	121	8.6
Giraff	(“Giraff,” n.d.)	170	-
Ava	(“Ava Robotics,” n.d.)	166*	60
Teleme2	(“MantaroBot - MantaroBot TeleMe 2 -- Datasheet,” n.d.)		
Endurance	(“Telepresence robot / system for video conferencing - Endurance,” n.d.)	170*	12
Vita	(“Telehealth Devices & Equipment - Vita InTouch Health,” n.d.)	168	79.8
Webot	(“The Webot telepresence robot .,” n.d.)	140	-
Cobalt Robot	(“Cobalt Robotics,” n.d.)	156	68

* Adjustable height

- No information found

The list of commercial robotic telepresence systems (see Table 1) was compiled by searching the internet and particular retail sites. The inclusion criteria for the commercial systems were that 1) it could be purchased, 2) had accessible

documentation, and 3) was usable “out-of-the-box,” i.e., only a minor setup is needed for the system to be fully functional.

Applying these criteria resulted in a selection of a total of 12 systems. The data about these systems are from their respective companies’ official websites. It is worth highlighting that many research studies have used commercial systems.

In studies 2-5, the Beam Pro was used. In a pre-study for study 4 (see Kaptelinin et al., 2017), a Double 2 was used (predecessor to Double 3).

The related research items are from a dataset³ consisting of search results from ACM Digital Library, Web of Science, and Scopus. The search query was: "Robotic telepresence," "mobile remote presence," and "mobile telepresence." The search was also enhanced by including papers that contained "presence" and "robo*." The popular abbreviation "MRP" was not used due to copious generation of false positives.

General Designs of Robotic Telepresence systems

As previously mentioned, one of the first Robotic Telepresence systems, developed in the 1990s, was PROPs (Paulos and Canny, 1998a). By this time, the main functionality seen in current commercial systems had already been implemented. These included connecting a computer to the robot, interacting using videoconferencing technology, and moving around in the local environment.

Overall, all the 12 robots (see Table 1) have similar capabilities and have all the functionality summarized in Figure 4. All robots have the same movement capabilities: they can go forwards, backwards, and turn. The turn commands rotate the robot. Combining forwards/backwards commands with direction commands makes the robot turn while moving forwards/backwards.

The majority of the robots have four wheels, two driving wheels and two stabilizer wheels. The only exception is the two-wheeled Double 3, which uses a gyro system, similar to Segway, to balance automatically.

The robots weigh between 7.3kg (Double 3) and 79.8kg (Vita), meaning it is possible to lift and carry the lighter robots (see Table 1). Omni Supercam is

³ While the dataset could be used for a structured literature review, this section does not intend to be that rigid, instead focusing on presenting to the reader the overall design of robotic telepresence. This dataset was compiled in 2021.

foldable to allow it to “*be carried with one hand or put in a car*” (“Learn How to Control Ohmni,” n.d.).

The shortest of the robots is BotEyes Pro (110 cm), and the tallest is Giraff (170 cm). Three robots have adjustable heights: Ava, Double 3, and Endurance. Adjusting the height of Ava involves sliding the screen down the body. For Double 3 and Endurance, the entire head moves and thus the total height changes. The dynamic length of the robot is determined by the pilot wanting to be at approximately eye level with the local users. ScalableBody builds on this idea by supporting dynamic scaling to adapt to the local users (Matsuda and Rekimoto, 2016). The ScalableBody supports eye-to-eye conversations with local users of various heights and when the local user is sitting or standing. The idea of eye-to-eye contact is also a motivator for the use of face-tracking, which tilts the screen towards the local user (Chua et al., 2012).

Capabilities and Functionality

In general, the commercial systems have similar capabilities, with the main exceptions being the three specialized robots, Cobalt (security), VITA (medicine), and Giraff (elderly care). However, much research is carried out where there is no immediate commercial application.

All robots have one or more forward-facing cameras that capture events occurring in front of the robot. The main difference between the systems’ forward-facing cameras is the type of lens used (e.g. regular, fish-eye, or wide-angle) and whether images from multiple cameras are stitched together into one image (see Double 3).

A challenge for a pilot is understanding where the robot is in the local environment, including knowing the location in “the world” and situational awareness of objects in the robot’s proximity (Chen et al., 2007; Yanco and Drury, 2004).

The robot’s camera is an obvious candidate for improvement since it is the pilot’s primary method of understanding the local environment. Such efforts have focused on various forms of wide-angle and panoramic solutions (Bazzano et al., 2019; Johnson et al., 2015; Lazewatsky and Smart, 2011; Vaughan et al., 2016). The conclusion is that a wider camera angle increases situational awareness but can be distracting (Bazzano et al., 2019; Johnson et al., 2015; Vaughan et al., 2016).

Many systems, such as the Beam Pro, have a secondary downwards-facing camera that makes it easier for the pilot to navigate in tight spaces by seeing the

area around the robot's base. Cobalt provides surveillance functions by having 360°, night vision, thermal and depth cameras.

In research, there have been efforts to improve the pilot's awareness of the robot's immediate surroundings by combining a head-mounted display (HMD) and 360° camera. Investigations show promising results when using the system outdoor for geocaching (Heshmat et al., 2018) and during a collaborative assembly task (Kratz and Ferriera, 2016). A benefit of using HMD with a 360° camera is that the pilot can move their head to look around.

An alternative way of increasing awareness by providing the ability to look around is to support head movements independent of the robot's body (Bamoallem et al., 2016; Bazzano et al., 2019; Clotet et al., 2016). The main advantage of head movement seems to be that the user can inspect smaller details (Bazzano et al., 2017; Clotet et al., 2016). An additional factor is that the pilot can use head movements to create gestures during social interaction (Bamoallem et al., 2016; Nakanishi et al., 2008).

Six commercial robots can tilt their head, two of them only tilt up and down, and four can also tilt sideways. Tilting the head down gives the same function as a downwards-facing camera.

Various mapping solutions are looking promising in providing a better holistic understanding for pilots, especially when combined with a video stream (Nielsen and Goodrich, 2006).

The screens used by commercial robots mainly differ in size and resolution. However, in research, various alternative screen solutions have been examined, most notably, the use of monitors (Jouppi et al., 2004) or holograms (Tokuda et al., 2013). In particular, a system consisting of multiple screens also provides multiple screens for the pilot, thus making it possible for them to look around in a similar fashion to the previously mentioned HMD solutions (Tokuda et al., 2013).

The challenge of designing microphones for robotic telepresence systems is that various environments present different challenges. In noisy environments, the pilot needs to be able to focus on the sounds of interest. A noisy environment is also a problem because the noisier it is, the louder the speaker has to be, and vice versa (Hayamizu et al., 2014). The same type of issue also exists regarding the level of the robot's speakers. In this case, the pilot's biggest problem might be knowing how loud they speak, and especially whether they are speaking too loudly (Neustaedter et al., 2016).

The difference in the microphones used by commercial systems is mainly the number used, from simple solutions found in everyday handheld devices to multiple microphones (Double 3).

Studies have investigated utilizing audio for improved immersion and presence (Kiselev et al., 2015b; Liu et al., 2015), as well as focusing in on specific sounds (Izumi et al., 2014). The idea of improving immersion and presence is also a question of awareness. The design should provide necessary information about where the sound is coming from, basically mimicking the human’s auditory sense.

Operating Telepresence Robots

Commercial systems rely on standard input devices, such as a keyboard, mouse, touchpad, and game controller, to control the robot (see Table 2)⁴. In research, alternative input devices have been developed to target users unable to use their hands to control the robot.

All the robots move between 1.5km/h and 5.5km/h, except Cobalt, which can reach 9.7km/h when in “emergency mode.” Thus, the typical speed of telepresence robots is close to that of a human walking indoors.

Table 2: Selection of commercial systems and their supported input devices

Robot model	Keyboard	Mouse	Game controller*	Touch screen
Double 3	+	+	-	-
Ohmni supercam	+	+	+	-
Beam pro	+	+	+	-
Boteyes pro	-	+	-	-
Vgo	+	+	-	-
Giraff	+	+	-	-
Ava	+	+	-	-
Teleme2	+	+	-	+
Endurance	+	+	-	-
Vita	-	+	+	-
Webot	+	+	-	-
Cobalt robot	+	+	-	-
Total	10	12	3	1

* xbox or playstation controllers

⁴ Study 3’s focus is on input devices, thus it includes a more comprehensive overview and discussion on the topic than in this section.

The most common input devices are keyboard and mouse. The users control the robot with a keyboard using the arrow or WSDA keys, a standard design for moving virtual avatars in various games.

There are three different variations of interaction with the mouse: 1) on-screen buttons (see Ohmni), 2) press-and-move (see Beam Pro), and 3) click-on-goals⁵ (see Double 2).

The *on-screen buttons* function as a virtual keyboard where the users press the button corresponding to the desired action. In the *press-and-move* design, the user holds down a mouse button over a specific UI area, and the robot moves in response to the mouse movements. The UI area in Beam Pro is the video feed from the downwards-facing camera. In the *click-on-goal* design, the user clicks on the robot's desired position, and the robot moves there autonomously. The user can either select the desired position on a map (Cobalt) or in the video feed (Double 3).

The two systems that support use of a game controller (Beam Pro and Double 3) utilize the game controller's "thumbstick" to move the robot. It is notable that only a discontinued robot, InTouch RP-7, supports a traditional joystick (Mendez et al., 2013).

TeleMe 2 is the only system that uses a touchscreen as an input device for the computer. In this case, the implementation uses a combination of an on-screen keyboard and gyros, so tilting the device turns the robot.

In research, there have been experiments with alternative input devices such as body tracking (Berri et al., 2014; Kang et al., 2018), brain robot interface (BRI) (Leeb et al., 2015), and eye-tracking (Zhang et al., 2019). The body-tracking can either mirror the pilot's movements (Kang et al., 2018) or control the robot through particular gestures (Berri et al., 2014).

The development of BRI (Leeb et al., 2015) and eye-tracking (Zhang et al., 2019) has been motivated by the desire to support people with "severe motor disabilities," such as those who are unable to use hand-based input devices. However, one major weakness of BRI and eye-tracking is that both can produce involuntary actions.

⁵ Also known as point-and-click (Bazzano et al., 2017) and through-the-screen (Vaughan et al., 2016). The choice of using "click-on-goal" is based on the action that is occurring, and not how the interaction necessarily looks. I also believe "click-on-goal" to be a better description of the interaction.

Possible uses for Robotic Telepresence

This section highlights possible uses for robotic telepresence by combining commercial marketing and research. While this illustrates how researchers and developers anticipate robotic telepresence usage, it is worth acknowledging that underlying agendas affect the highlighted uses. Marketing is driven by the desire to find customers, whereas researchers are interested in the field of research and what types of projects can receive funding.

Table 3: Uses of commercially available systems

	Office work	Home visits	Culture and art	Retail	Education	Medicine	Care	Security	Manufacturing	Open environments	Sell property
Double 3	+	-	-	+	+	+	-	+	-	-	-
Ohmni supercam	+	+	-	-	+	-	+	-	-	-	-
Beam pro	+	+	+	+	+	-	+	-	-	-	-
Boteyes pro	+	-	-	-	-	-	-	-	+	-	-
Vgo	+	+	-	-	+	+	+	-	-	-	-
Giraff	+	-	-	-	-	-	+	-	-	-	-
Ava	+	-	-	-	-	-	-	-	+	+	-
Teleme2	+	+	-	-	-	-	-	-	-	+	-
Endurance	+	-	-	-	-	+	+	-	-	+	+
Vita	-	-	-	-	-	+	-	-	-	-	-
Webot	+	-	+	+	-	-	-	-	+	-	-
Cobalt robot	-	-	-	-	-	-	-	+	-	-	-
Total	10	4	2	3	4	4	5	2	3	3	1

The most common use of commercial systems is *office work* (see Table 3). Office work includes hiring professional workers, showing customers around the company, and collaboration between multiple sites. The target user groups include engineers, supervisors, project managers, and executives.

Care is the second most common use of commercial systems and has gained the most research attention (33 research items in total). The research into *elderly care* focuses on using robotic telepresence technology to enable relatives and caregivers to visit elderly adults in their homes (Orlandini et al., 2016). Privacy concerns regarding the use of robotic telepresence in the home environment have resulted in Giraff implementing a solution where the caregiver calls the Giraff, and the care recipient has to answer, something which the caregiver can override in emergencies (Cesta et al., 2016).

Aside from increasing the efficiency of elderly care, a robotic telepresence system can reduce loneliness in the elderly (Cesta et al., 2016; Moyle et al., 2014). Designs acknowledging this include enabling the care recipient to initiate a call using a "call me button" (Aaltonen et al., 2017).

While the previously mentioned uses mainly focus on older adults as local users, some research has been carried out into how older adults might want to use robotic telepresence systems themselves (Beer and Takayama, 2011; Rae et al., 2015). Applications include visiting *museums* and sharing meals.

Interestingly, *museums* are among the earliest (Agah and Tanie, 1999) and most researched area of application for robotic telepresence technology. The research has investigated using the robot as a guide (Burgard et al., 1999; Pang et al., 2017; Roussou et al., 2001) or as a method to allow remote visits (Agah and Tanie, 1999; Brayda et al., 2009; Burgard et al., 2003; Claudio et al., 2017; Ng et al., 2015; Pang et al., 2017). The development of the uses of robotic telepresence systems is motivated by a desire to help those with mobility issues (Ng et al., 2015), visit inaccessible parts of a museum (Claudio, 2017), to help those with visual impairment (Park et al., 2015), or as a part of the experience (Agah and Tanie, 1999; All and Nourbakhsh, 2001).

The use of such systems in *medicine* is similar to that in care. The standard scenario is that a doctor visits a patient in a hospital, often accompanied by a physically present medical worker (i.e. doctor or nurse). The Vita system is specialized for medical applications and has various medical-related add-ons, including a stethoscope, to assist the doctor.

Robotic telepresence has been used in various *medical contexts*. Many of the studies into medical use have used medical robots from InTouch. These examples include mentoring (Agarwal et al., 2007), use in rural environments (Mendez et al., 2013) and use in intensive care units (McNelis et al., 2012).

There are two particular examples that illustrate the most common reasons behind the use of robotic telepresence in the medical context. First, it can be used to help patients in rural areas where access to medical professionals is generally limited (Goodridge and Marciniuk, 2016; Mendez et al., 2013; Selic, 2014). The idea is simply that the medical doctor can be present and interact with the patients in these rural areas through robotic telepresence.

The second example is connecting patients to experts, where the expertise is somewhat scarce. An example is the emergency stroke network found in the USA (Morales-Vidal and Ruland, 2013), which hospitals can contact when receiving a

patient with a stroke. The critical element of these examples is that rapid delivery of the correct medicine significantly improves the situation.

Three companies mention various forms of open indoor spaces, referred to as "*open environment*." AVA Robotics uses the word "hospitality" and this includes helping guests at a hotel and visiting conferences. Similarly, there have been studies on using telepresence systems at conferences (Neustaedter et al., 2018, 2016; Rae and Neustaedter, 2017). While an *open environment* shares similarities with *office work*, the main difference is the size of the environment.

Security and surveillance are an example where the main selling point is that the security officer can be present in a surveillance room while patrolling the site. Cobolt explicitly focuses on security and is equipped with a vast range of automation, various types of surveillance and can, if needed, be set to move faster than any other robot. Cobolt also has an automatic surveillance mode that allows a security officer to override the robot when it detects something suspicious.

Additional areas of use include properties, retail stores, laboratory environments, and manufacturing processes. The AVA robot is designed for use in sterile laboratory environments. TeleMe 2's manufacturer, MantaRobot, also has a robot version called TeleTrak that expands TeleMe 2's capabilities by having continuous tracks suitable for uneven terrain. Thus, TeleTrak enables outdoor use in places such as industrial building projects.

There are examples where robotic telepresence systems can be used in education: helping students who could not otherwise attend classes due to illness, allowing university students to attend classes remotely, and allowing teachers to take classes remotely.

Last, Ava highlights that the robotic telepresence system has the potential to enable people with various forms of disabilities to be present in locations and events that their disability would have restricted them from attending. Such an argument would also be applicable to people who were immunocompromised, such as those mentioned in education.

In conclusion, while the target user group for such systems is professionals who could use the robot in their daily work, this may be a result of commercial strategies. Thus, as the technology becomes cheaper and more used, examples of use from outside the professional world might become increasingly common.

3 HCI and Robotic Telepresence

This chapter outlines the research into robotic telepresence in the field of HCI. Robotic telepresence has become a research topic in Human-Computer Interaction (Beer and Takayama, 2011; Herring et al., 2016; Rae et al., 2014; Tsui and Yanco, 2013) as its technological maturity now enables investigation of real-world uses.

It is evident from several user studies that robotic telepresence offers another dimension of connection to the world compared to traditional videoconference systems (for example, Lister, 2020; Rae and Neustaedter, 2017; Takayama and Go, 2012).. Robotic telepresence presents similar possibilities and problems as related technologies do i.e. controlling the robot is like controlling a remote car, drone, and playing various video games, while interactions are closely related to those using videoconference systems. However, robotic telepresence also offers novelty by combining said technologies, and through this combination creates additional areas of usage.

Robotic telepresence systems increase the pilot's social presence while the pilot experiences an increased spatial presence and independence in the remote environment (Takayama and Go, 2012; Venolia et al., 2010). This mobility and independence enable ad hoc conversations and the freedom to move around and look at objects.

It is worth noting that the pilot and local users often compare the experience of robotic telepresence to either videoconferencing calls or being physically present. Since these are two extremes in a continuum where robotic telepresence is somewhere in the middle, it causes a sense of ambivalence as robotic telepresence provides additional value over other technologies while still being limiting compared to actual physical presence. When the frame of reference is being physically present, then the system becomes a filter between the pilot and the world.

This chapter consists of three sections: the first two focus on the pilot and local user experience, and the third section focuses on the social interaction between pilots and local users. The social interaction between the pilot and the local user is, in many cases, an aggregation of the two user groups' experience of the system.

3.1 The Pilot's experience

This section will outline the *general usability and the pilot's initial experience* to form an understanding of the lower-level usages of robotic telepresence

technology. On this basis, two particular experiences of being connected to the world through robotic telepresence will be focused on, *being in the local environment* and the sense of *being in a robotic body*. Last, there is a particular focus on how the pilot can engage with the world, and especially consideration of the pilot's independence in the local environment.

General Usability and Initial Experiences

It is evident from many user studies that robotic telepresence systems are relatively easy to use, even for new users (Jones et al., 2020b). The literature also indicates that previous experience in playing video games positively affects the pilot's performance (Takayama et al., 2011).

While studies show that users successfully operate the robot in a variety of challenging conditions, there are clear indications that users experience a relatively high cognitive load (Kiselev & Loutfi, 2012). Thus, it seems that the act of using the robotic telepresence system for a long time is exhausting, partially because it requires a higher degree of attention to avoid objects and people (Neustaedter et al., 2018).

Therefore, it is unsurprising that there have been attempts to simplify operations using semi-automation to address common issues, including automatic obstacle avoidance (Takayama et al., 2011), follow-me system (Cheng et al., 2019; Cosgun et al., 2013), automatic navigation to pre-determined locations (Kiselev et al., 2015a), and being able to select the location on a map where the robot then automatically moves to (Coltin et al., 2012).

Being in the local environment

Robotic telepresence pilots experience "being in the local environment," a spatial presence, a feeling of being in the robot's location (Lister, 2020) with an increased degree of immersion in the situation (Neustaedter et al., 2016). This sense of spatial presence is arguably a critical benefit, since it enhances the feeling of being at a location the pilot wants to be at.

The experience of being in the local environment consists of three main components: spatial presence, awareness of surroundings, and macro-level understanding. While robotic telepresence reportedly supports spatial presence (Lister, 2020; Neustaedter et al., 2016), awareness of surrounding and macro-level understanding is a challenge for users.

A lack of awareness of the surroundings includes not understanding who else is present in that environment and not appreciating the physical dimensions of the robot body and objects in its proximity. Such problems are particularly acute

when there are large numbers of people in the local environment, for example, at conferences (Rae and Neustaedter, 2017).

As mentioned in the previous section, successful efforts to increase the spatial presence include using HMD devices combined with a 360° camera (Heshmat et al., 2018; Jones et al., 2021). The main benefit of using an HMD is that it can utilize head movements, so instead of moving the robot around to look at things, the user can move their head. Consequently, this enables quick and effortless assimilation of the robot's surroundings. However, major drawbacks include the recognized HMD issue of dizziness (Jones et al., 2021) and the fact that the HMD hides the upper part of the pilot's face.

Methods of enhancing the awareness of the surroundings include the previously mentioned obstacle avoidance solutions and haptic feedback systems, such as “FeetBack,” a system for providing feedback to a pilot's feet (Jones et al., 2020a).

While pilots often seem effective in navigating in a room, they have difficulty understanding and navigating large areas. There are two problems with navigating larger areas. *First*, it is hard for the pilot to form a macro-level understanding, that is, know where the robot is in the building. *Second*, the resolution of the camera is not sufficient to give the pilot adequate navigational clues, such as reading signs (Neustaedter et al., 2018, 2016; Rae and Neustaedter, 2017). The most obvious solution for navigation in larger areas is the use of maps, sometimes combined with autonomous driving (Coltin et al., 2012).

However, maps only provide an overarching picture of where the pilot is, and their primary function is to simplify navigation. The macro-level understanding of being at a particular place includes other factors, such as how the surrounding environment looks and feels. The remote conference participant who wanted to look through a window to see the surroundings of the conference provides an excellent example of the need for macro-level understanding (Neustaedter et al., 2016).

Research into presence

Research into the psychological sensation of presence (Riva et al., 2014) is related to the feeling of spatial presence when using robotic telepresence systems. The work on presence has focused on the psychological sensation of feeling presence at another location. Although presence and robotic telepresence share the same roots from Minsky (1980), there has been little overlap. Instead, the primary technology for investigating presence has been virtual reality.

Meanwhile, users sense presence when using robotic telepresence systems (Rae et al., 2014), and both phenomenology (Dolezal, 2009) and activity theory (Riva

et al., 2011) have been used to analyze that sense; however, this has limited impact on this work. In this thesis, the usage of presence will primarily be one of the multiple experiences mediated through robotic telepresence technology. It is evident that this is a narrow use of a much broader concept, but it is out of scope for this thesis to combine robotic telepresence and HCI with the work on presence.

Furthermore, this thesis is not isolated to understanding the presence or any particular physical sensation of telepresence. Instead, the thesis is about understanding human connections to the world through robotic telepresence. Setting a larger picture of the relationship between human and robotic telepresence affords a top-down approach when formulating a theoretical framework. Arguably, a bottom-up approach, such as focusing on presence, could have resulted in the same outcome. Nevertheless, such an approach would always risk unintentionally omitting productive perspectives.

Experiencing presence in a remote environment is a positive aspect of robotic telepresence systems. Further understanding of what affects the sense of spatial presence, how to improve, and benefit from this experience, could be essential for further development of robotic telepresence technology.

Being in the body

A robotic body is a "proxy" for the pilot's own body, "an extension of themselves" (Neustaedter et al., 2018). However, the pilot's relationship to the body is multifaceted, and includes bodily ownership, bodily awareness, and understanding of the robotic impact in the context.

The pilot sometimes even embody the technology to such a degree that it becomes an extension of themselves, that is, they feel that the body is "theirs" (Lee and Takayama, 2011; Rae and Neustaedter, 2017). This body ownership is evident when pilots feel negative emotions after local users have violated the pilot's personal space (Takayama and Go, 2012). Likewise, the pilot's "personal space" around the robot highlights a sense of ownership.

The pilot lacks a third-person view of themselves. They have difficulty understanding a local user's perception of their appearance (Neustaedter et al., 2016) and how their actions are perceived.

The standard telepresence robots are rather sterile in their design. Efforts to enable personalization of the robot have produced mixed results. While personalization supports a pilot's self-expression, it remains hard to understand how others perceive them (Khojasteh et al., 2019; Neustaedter et al., 2018, 2016).

There are additional issues with the pilot's lack of external self-awareness, including cases where the pilot has broken social norms by talking too loud (Neustaedter et al., 2016), making too much noise when moving (Neustaedter et al., 2018; Takayama and Go, 2012), and physically blocking local users (Lee and Takayama, 2011; Neustaedter et al., 2016). The following quote from a telepresence conference attendant illustrates the issue: *“However, it was more difficult for the remote attendees [pilots] to ask the local attendees [local people] if they were blocking their view. They lacked the ability to easily turn around and whisper or ask them in a low voice”* (Neustaedter et al., 2016).

While blocking a local user's field of view could be solved by increasing a pilot's spatial awareness, it is harder for a pilot to understand how loud they sound. There are two principal solutions for this: giving feedback to the pilot about the volume level through use of sidetone (Paepcke et al., 2011) or automatically adjusting the volume dependent on the ambient noise (Hayamizu et al., 2014). While these solutions have a positive effect, it is still worth acknowledging that the pilot's limited understanding of the consequences their robot-mediated acts have on the environment is problematic.

While the sense of being in a remote body has been a reported phenomenon, there have been limited theoretical discussions regarding this in the general robotic telepresence literature. In early developments of PRoP, the term “tele-embodiment” was introduced (Paulos and Canny, 1998b). However, tele-embodiment has never been grounded in a larger theoretical framework. As will later be addressed more thoroughly, grounding the phenomenon into a theoretical framework provides benefits of connecting it into a larger body of knowledge and provides a perspective through which the phenomenon can be understood.

Pilot's independence in the local environment

Unlike videoconferencing systems, robotic telepresence offers users the freedom to move in the local environment. The pilot's mobility increases their independence from local users. Robotic telepresence provides a physical mobile representation of the pilot in the local environment, increasing autonomy, social engagement, and agency in learning (Lister, 2020) in comparison to videoconferencing systems.

However, there are cases where the pilot becomes *explicitly* dependent on the local users. While the pilot has a physical presence through the robot, they are extremely limited when undertaking basic physical actions such as opening a door or turning on lights (Tsui et al., 2011; Yang et al., 2017; Yang and Neustaedter, 2018). Typically, the only physical actions possible relate to driving

the robot into things. As a result, a common metaphor for Robotic Telepresence systems is "person with disabilities" (Neustaedter et al., 2016; Takayama and Go, 2012).

There are numerous case studies where the limited actions available have resulted in negative experiences for the pilot. Such cases include missing meetings because of the pilot's inability to open a door (Lee and Takayama, 2011), pilots who were unable to attend a conference session because they could not use the elevator (Rae and Neustaedter, 2017; Takayama and Go, 2012), and remote partners in long-distance relationships feeling inadequate about not being more helpful with household chores (Yang et al., 2017).

The lack of possible actions makes the pilot dependent on local users. Eliciting help from local users is not always experienced negatively (Yang et al., 2017) but is still a concern (Neustaedter et al., 2016; Takayama and Go, 2012; Tsui et al., 2012; Yang and Neustaedter, 2018). Furthermore, recruiting help from a local user affects the social context (Boudouraki et al., 2021).

There are mainly two types of designs aimed at increasing the physical actions available to pilots, namely adding physical manipulation capabilities, such as a grip arm (Koceska et al., 2019), and using existing smart home technologies such as Nest Home Hub (Yang and Neustaedter, 2020).

The potential of telemanipulation technology, such as robotic arms, is that the pilot can use these for a diverse set of tasks. However, there are concerns concerning safety and how to control these efficiently (Kristoffersson et al., 2013)

While using voice-activated home assistance tools is safer, these also have limitations. Standard home assistance/automation tools do not explicitly target the requirements of robotic telepresence users.

3.2 The Local User's experience

It is clear that the local user's experience of robotic telepresence systems is vastly different from that of the pilot. The difference is easy to understand since the local user is in the local environment with their own body. The local user often does not need to interact directly with the robot. Instead, they interact with the pilot through the robot.

The local users' view of the robot

The robot is perceived both as a machine and a person (Lee and Takayama, 2011). While this machine-person dualism is present for the pilots, it is most evident to the local user.

Efforts to alter the perception for the local user include various degrees of personalization (Neustaedter et al., 2016) and a rather interesting Augmented Reality (AR) solution (called VROOM), where an avatar projection of the pilot is visible for the local user using AR glasses (Jones et al., 2020b). The VROOM system presents a virtual avatar based on the pilot. Interestingly, the VROOM system highlights that the uncanny valley problem⁶ (Wang et al., 2015), common in human-like robotics, could potentially be a critical problem for telepresence (Jones et al., 2020b).

Studies in the wild indicate that the novelty effect affects the behavior of local users more profoundly than the pilot experiences (Neustaedter et al., 2016; Niemelä et al., 2019; Rae and Neustaedter, 2017). For example, the local user might take selfies and investigate the robot's functions (Rae and Neustaedter, 2017), reflecting behavior that seems to relate to viewing the robot as a machine. In a field study of the use of robotic telepresence at an HCI conference, there were even cases of repeated bullying and bad behavior (Neustaedter et al., 2018).

Security and privacy

The presence of local users also raises privacy and security issues. The two main issues are that the pilot can log in to the robot without any explicit permission from the local user (Cesta et al., 2016), and that local users can perceive the robot as a surveillance device (Heshmat et al., 2018; Niemelä et al., 2017).

The ability of the pilot to log in to the robot independently is a necessary requirement in many applications while, for example, in a home care setting, such a feature could infringe privacy (Cesta et al., 2016). In this case, the design solution is that the local user needs to grant permission each time a pilot logs in⁷ (Cesta et al., 2016).

Local users can also perceive the robot as a device for surveillance because they do not know how the video data is used (Heshmat et al., 2018). A problem related to this is that local users cannot tell if the device is active or not (Rae and Neustaedter, 2017). The personalization factor is also an issue for the local users, since they cannot identify the pilot without seeing the face on the screen (Neustaedter et al., 2016).

⁶ The uncanny valley problem is a negative feeling towards an object that closely resembles a human but lacks vital human traits (Mori et al., 2012; Wang et al., 2015)

⁷ In many ways, this is similar to a telephone call where the person being called has to pick up the receiver in order for the call to take place.

3.3 Social Interaction

Supporting social interaction is a fundamental capability of robotic telepresence systems. Robotic telepresence enables a type of informal social ad hoc meeting that was previously reserved for in-person interactions. These include actions such as meeting in a hallway, visiting someone's office, and talking in the break room (Lee and Takayama, 2011).

What is evident is that the interaction between the pilot and the local user is, in many ways, a collection of the problems outlined in previous sections. Social interaction adds a level of complexity with unique problems and opportunities.

One major underlying cause of many issues in social interaction is the asymmetry between the local user and the pilot. The pilot and the local user have a different experience of the situation, and they find it difficult to understand each other's perspective. Furthermore, it is also a question about robot-human dualism, where the robotic telepresence offers capabilities previously reserved for physically present actors.

Asymmetric relationship

There is an asymmetric relationship between the pilot and the local user caused by the technological setup (Boudouraki et al., 2021). The root of this asymmetry is that the pilot is present through a robot, and the local user is present with their body. Thus, the two user groups do not have the same experience and capabilities.

There are multiple factors to this asymmetry, including the pilot being in two locations simultaneously (Yang and Neustaedter, 2018), the pilot and the local user having a different view of what the robot is, and the problem of understanding each other's experience.

As Yang and Neustaedter pointed out (2018), while the local user is only present in the local environment, the pilot is present at both the local and the remote environment. As a result, the pilot's actions and decision-making are less transparent than those of a local user (Neustaedter et al., 2016).

As discussed, there are many metaphors for robotic telepresence (Takayama and Go, 2012), but one, in particular, stands out, namely that of being an *object*. That local users adopt the object metaphor is problematic, since the pilot views the robot as a proxy of themselves, it is "their" body, and will thus experience behavior that treats the robot body as an object negatively (Neustaedter et al., 2018; Takayama and Go, 2012). This behavior includes moving the robot (Neustaedter et al., 2018), turning off the robot (Takayama and Go, 2012), and,

arguably, the reported “novelty effect” when people approach the pilots for selfies or through curiosity (Rae and Neustaedter, 2017)

It is thus encouraged to move beyond the metaphor of seeing the robot as an object, towards metaphors that “humanize” the pilot’s avatars without building expectations of the pilot’s capabilities too high (Takayama and Go, 2012).

This asymmetry is also evident in the pilot and local users’ understandings of each other. The pilot has a better understanding of the experience of being a local user than vice versa. Perhaps the most telling examples are when the local user asks the pilot questions such as “can you see this?” (Boudouraki et al., 2021). Meanwhile, the pilot often finds it difficult understanding how the local user perceives them, both from the simple view of how the robot looks, to more socially complex situations, including being too loud (Neustaedter et al., 2016) or unintentionally blocking a local user’s field of view (Neustaedter et al., 2016).

The asymmetry between the local user and the pilot is rooted in the technological setup and, while it is possible to mitigate this through design and the development of social norms, it will not be solved completely. Consequently, it is crucial to consider both local users and pilots when designing and understanding robotic telepresence. Since the difference is so large between these user groups, it is not feasible to extrapolate the results from one group to another.

Helping and power structure

The pilot’s lack of action capabilities in the local environment makes them dependent on local users to solve problems (Boudouraki et al., 2021; Heshmat et al., 2018; Jones et al., 2020b; Yang et al., 2018). Such dependence can unbalance the social relationship between the pilot and the local user (Boudouraki et al., 2021).

Compared to videoconferencing systems, the added independence creates a social dilemma. Videoconferencing affords a higher degree of assistance from the local user and has a lower degree of expected independence.

Consequently, local users expect attendees using a videoconferencing system to need assistance in all activities involving moving and looking around, while robotic telepresence pilots only need assistance for particular situations.

In a way, describing the telepresence user as a “person with disabilities” (Takayama and Go, 2012) captures this perspective. On the one hand, the pilot has a similar degree of freedom as the local users but, on the other hand, they lack the same degree of capabilities.

Social Body

While it is evident that the robot body provides an additional degree of social presence, it is also the case that the robot's capabilities are limited when expressing gestures and social cues. As Tsui and Yanco wrote:

"As social interaction is the primary goal of social telepresence robots, failure to design for eye gaze, facial expressions, and nonverbal gestures will result in systems that hinder the ability to achieve telepresence for the user and/or the interactant." (Tsui and Yanco, 2013)

Gestures are fundamental in human-human communication and might even be the first form of cooperative communication (Tomasello, 2010). In light of this, it is reasonable that researchers have emphasized a need for robotic telepresence systems to support communicational gestures for telepresence robots (Cohen et al., 2011; Kaptelinin, 2016; Stahl et al., 2018).

Gesture communication is limited to gesturing through the video feed. It is notable that early robotic telepresence systems, including PRoP (Paulos and Canny, 1998a), PEBBLES (Kristoffersson et al., 2013) and GestureMan (Kuzuoka et al., 2000), were equipped with a rudimentary arm/hand for conveying simple gestures.

One form of gesture is to direct attention to some object in the room through pointing. Both QB and MantaroBot (Kristoffersson et al., 2013) addressed the lack of pointing capabilities by attaching laser pointers to the robot. Some more recent solutions have included a pointing stick (Fitter et al., 2019) and augmented reality solutions that project a body onto the robot (Jones et al., 2021).

It is worth emphasizing that body language involves far more than just hand gestures. Body language is also about how the robotic body is positioned and moves (Stahl et al., 2018).

Furthermore, studies have also reported that the pilots of the robotic telepresence systems might struggle to see the body language of the local users (Khojasteh et al., 2019; Neustaedter et al., 2018).

3.4 Summary

While current research in robotic telepresence is extensive, some unaddressed issues are relevant when addressing the research question by understanding how robotic telepresence shapes users' connections to the world.

First, there should be investigation into how robotic telepresence users can use body gestures for communication. Such investigation could uncover the social dimensions that limit the design space of social interaction.

Second, while several studies have focused on the use of robotic telepresence systems, few studies have explored potential future users' expected appropriation. Thus, to understand in what ways robotic telepresence can affect people's lives, there needs to be an understanding as to what appropriation a general user group might anticipate. To understand the appropriation potential is also to understand the limits and possibilities of robotic telepresence.

Third, while there have been controlled experiments on how users operated the robot, there has been a general lack of comparison between input modalities and their effect on performance. Thus, this issue focuses on how the pilots operate the robot.

Fourth, while research acknowledges that the pilot's action space in the local environment is limited, there has been little investigation into methods for expanding the action space and what effects such expansions might have.

Fifth, studies of robotic telepresence have highlighted how the pilots feel a sense of spatial presence. However, while there is an extensive body of studies on the underlying factors of spatial presence in virtual reality, few have addressed robotic telepresence.

Furthermore, it is crucial to acknowledge the complex relationship between pilots and local users. Various forms of changes to the system can affect this relationship. Due to the asymmetric experience between the user groups, it might even be hard to understand such an effect. Consequently, it is essential to consider both the pilot and the local users' experience, especially when introducing capabilities that alter the range of potential actions.

4 Theoretical Foundations

The role of theory in this thesis is to act as a lens through which it is possible to position the research questions and the empirical findings within a larger knowledge construct (Halverson, 2002; Kaptelinin and Nardi, 2006). Thus, the theoretical foundations act somewhat as an ideology⁸ that one can see the world through while at the same time providing an overarching structure for analyzing the empirical findings.

I apply two different perspectives to address the research question. The first perspective focuses on the particular usage of the tool as a mediator, and the other focuses on the human experience. The theoretical foundations of this thesis need to address both perspectives while also addressing the complex and entangled relationships between people and the robotic telepresence without oversimplifying them.

The conceptual foundation of this thesis consists of activity theory (Kaptelinin and Nardi, 2006) and phenomenology (Dourish, 2001; Ihde, 1990). Activity theory and phenomenology are consistent at some of their most basic conceptual foundations (Kaptelinin and Nardi, 2012, 2006; Kosaka, 2013; Macdonald, 2000) while having significantly different starting perspectives. Combining activity theory with phenomenology is not unique to this work (see, for example, Hasse, 2013).

Activity theory and phenomenology provide two different focuses on the main research question and the sub-questions. The activity-centered perspective focuses on appropriate activities for appropriating the robotic telepresence and how this appropriation reshapes the activity. The experience-centered perspective of phenomenology emphasizes the user's uniquely personal *experience* of engaging with the world through robotic telepresence technology. Throughout the thesis, it will be evident that these two perspectives both contrast and complement each other.

Activity theory and phenomenology are comprehensive and have a long history of development, refinement, and application. However, the usage of activity theory and phenomenology in this thesis is focused solely on HCI and IxD, and this is thus rather narrow compared to their respective traditions.

Through activity theory it is possible to understand robotic telepresence as a tool used in meaningful activities, i.e., robotic telepresence framed as a mediator

⁸ See Eagleton (2007) for further elaboration of the relationship between ideology and theory.

technology between humans and their ideal objects. This theoretical lens makes it possible to comprehend further appropriation possibilities based on users' needs.

The phenomenology perspective complements activity theory by emphasizing human experience. In phenomenology, the tool is a mediator that affects our being in the world. It is thus possible to analyze how robotic telepresence enables a new embodiment form by applying a phenomenological perspective on the human–robotic telepresence relationship.

My usage of phenomenology focuses mainly on the role of tools and technology, emphasizing in particular the concept of embodiment. As a result, the main theoretical influences come on the whole from the HCI usage of phenomenology (Dourish, 2001) and the notion of postphenomenology⁹ (Ihde, 2010, 1990) that builds on important foundational works, such as (Heidegger, 2010; Merleau-Ponty, 2013).

4.1 Activity Theory

The core motive for using activity theory here is its inherent activity focus. Activity theory provides a framework that makes it possible to analyze acts and connect them to the user's overarching motive. Furthermore, mediational artifacts, such as a robotic telepresence, are vital in activity theory.

Activity theory found its way into HCI during the post-cognitivist wave. Since then, it has been a popular choice of theory in HCI (Clemmensen et al., 2016). In this thesis, the use of activity theory is mainly limited to previous HCI adoptions (Bødker, 1991; Kaptelinin and Nardi, 2006).

As the name implies, activity theory is a framework that focuses on activity. Activity is a subject's (e.g., a human being's) interaction with the object (i.e., the world). The subject has needs that the activity should fulfill.

The needs of a human include the most fundamental needs – such as food and shelter – as well as cultural and social needs. Thus, the object of an activity originates both from biological and societal conditions:

“But the main thing is ignored, that in society man finds not only his external conditions to which he must adapt his activity, but also that these very social

⁹ Postphenomenology was developed by Don Ihde (1998) in the field of Science and Technology Studies. While building upon phenomenology, it is also differentiated from it to such an extent that the prefix “post” is used to signify the distinction. Differences between phenomenology and postphenomenology will be addressed in section 1.2 Phenomenology and Postphenomenology.

conditions carry in themselves the motives and aims of his activity, the ways and means of its realization; in a word, that society produces human activity.” (Leontiev, 1977)

Thus, the social dimension is inherent within the activity. The activity has a double transformative characteristic that renders it in a state of constant change. The activity will change the world, and this changes the activity. Changing the world affects how the subject can perform the activity and the subject’s motives.

In activity theory, tools mediate between humans and the world. The definition of a tool includes technological tools, e.g., a hammer, and psychological tools, e.g., maps, blueprints, and algebraic notation (Kaptelinin and Nardi, 2006). Psychological tools can be external or internal; humans can internalize both. Consequently, an activity will be modified when appropriating a tool into it. Thus, introducing a robotic telepresence system modifies the activity and transforms it into a new activity.

Tools are also an expression of culture: *“Tools reflect the previous experience of other people, which experience is accumulated in the structural properties of tools, such as their shape or material, as well as in the knowledge of how the tool should be used”* (“Activity Theory,” n.d.). Thus, the contemporary robotic telepresence has resulted from previous technological advances and their usage. In many ways, robotic telepresence is only the latest artifact in a technological evolution that began with the telegraph.

While previous technologies have been developed over time, and changes have taken place based on users’ experience of the technology, robotic telepresence is currently in such an early phase that many uses and designs are yet to be explored.

The following sub-sections outline the main concepts of activity theory used in this thesis, i.e., the hierarchical structure of activity, mediation, and the role of artifacts in activities. The hierarchical structure of activity allows for analyzing small acts as parts of a larger whole, enabling an understanding of the sub-question relationships.

The section on mediation aims to outline the complexity of digital tools as mediators in activities, which is especially relevant in robotic telepresence systems because they consist of multiple artifacts.

Lastly, the section on the role of artifacts in activities focuses on how the introduction of tools changes the activity structures, especially as robotic telepresence is a novel technology for many potential users.

The hierarchical structure of activity

An activity consists of three leveled “acts”: activity, action, and operation (see Figure 5). The object of an activity is a motive that corresponds to a user's need. The goal is the object of an action, and the condition is the object of an operation (Kaptelinin and Nardi, 2006). An activity consists of actions that consist of operations. These three levels correspond to the three analysis questions "why?" (Activity), "what?" (Action), and "how?" (Operation) (Bødker and Klokmoose, 2011). An activity is an act that humans undertake to fulfill a motive. Actions are conscious goal-oriented acts that are the focus of the human agent. Thus, the object of an action is the goal. Operations are low-level acts that a human manages to do with a low cognitive effort – they are basically acts that are in some sense “automated.”

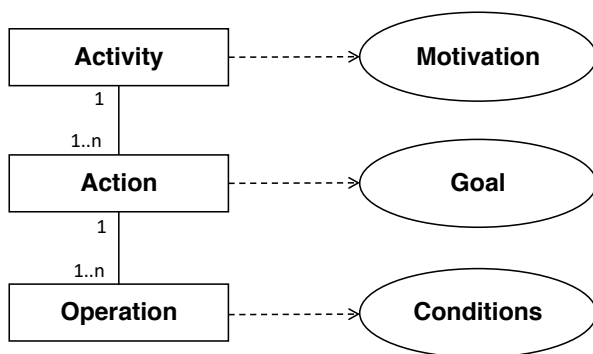


Figure 5: Hierarchical structure of activity

When conducting an activity, the human focus is on the action. While the human might be aware of the activity, it is not on this level that conscious acts are taking place. In the same way, humans do not need to focus on the operations since they are acts that do not need conscious attention.

Consequently, activity theory postulates that a person can only do a single action at any given time. However, evidently humans can conduct multiple operations at the same time: it is possible to “multi-task,” i.e., do multiple operations from multiple actions. “Multi-tasking” is also possible by rapidly switching between actions.

To exemplify this point: it is possible to drive a car and interact with the passengers. The driver can handle the vehicle on an operational level while focusing on social interaction. However, if the action that involves driving the car needs special attention, the driver needs to shift focus and, hopefully, pause the social interaction that is going on.

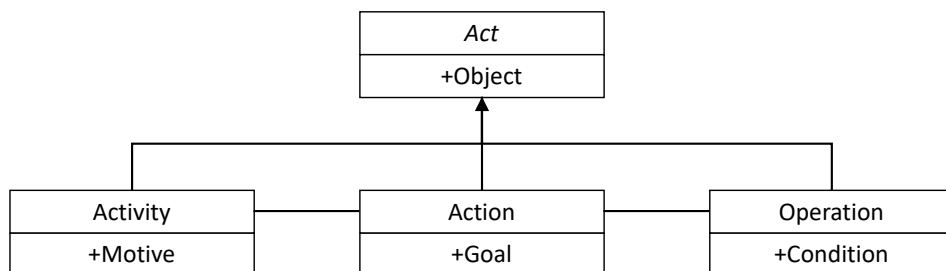


Figure 6: Acts and the relationships where all concrete acts, i.e., activity, action, and operation, inherit from the abstract class Act.

Figure 6 further illustrates the relationships between activity, action, and operation by using a UML-based class diagram. This diagram highlights that the classes Activity, Action, and Operation have a common abstract ancestor named “Act.” They have a common ancestor because an activity can become an action, an action can become an activity or an operation, and an operation can become an action. In other words, an act can take the shape of all three (see Figure 7). Note, however, that when an act changes, a ripple effect results in more extensive changes in the activity system than is conveyed in Figure 7.

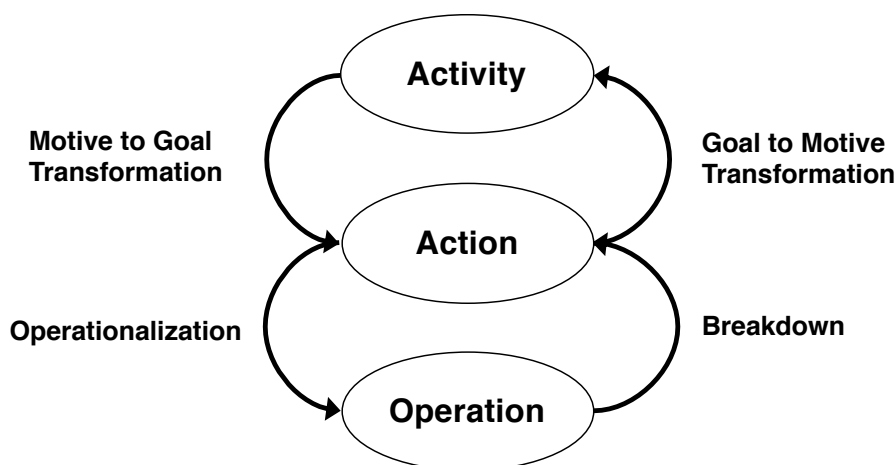


Figure 7: State changes of acts

An action can become an activity when the person's goal becomes a motive; for example, if the activity is learning, one's action goal is to get good grades. Getting good grades can be so influential that the person's sole motive is to get good grades. Another example is multiple people in a work-related meeting, with one person present via a robotic telepresence system. Instead of focusing on the work activity, the people involved only focus on the robot due to its novelty.

An action becomes an operation when the user no longer needs to have a specific focus on the action. This transformation occurs either by learning or via a re-design that enables operationalization.

An example of the learning aspect is a pilot's first interaction with a telepresence robot's controls. In the initial phase, the pilot needs to form actions to figure out how to handle it, i.e., they need to think, "to go forward, I hit this button." After some practice, the pilot has learned to use the controls efficiently; the action has become an operation. When sufficiently learned, the pilot does not need to focus on controlling the robot; instead, the focus can be on actions mediated by the robot.

Enabling an action to become an operation can be achieved through re-design by fully automating it, by automating parts of it so that the human can operationalize the rest or enable learning. When an action becomes an operation, it becomes a part of a new or modified action. The goal of the operationalized action becomes interwoven into the new or altered action.

It is also possible for an operation to become an action by a breakdown. In the previous example, the user had learned to control the robot to the extent that it became an operation. However, if the robot suddenly stops responding to the user's input, the user's focus would be on the input device; thus, the operation breaks down and becomes an action.

The structure of activities can be used both for descriptive and prescriptive analyses. It is possible to analyze activities by describing them from the lowest to the highest level; the analysis scope dictates the limit. The model is also feasible for prescriptive analyses since acts need to be of a specific class for the activity to unfold in the desired way.

This hierarchical structure highlights a number of essential factors. *Firstly*, it is evident that the activity's motive comes inherently from the user and is not designable per se. *Secondly*, actions are the level where the most outspoken acts occur. The actions' goals should fulfill the activity, which means that tools should aim to accomplish goals aligning with the activity's motives. *Thirdly*, for the user to successfully use a tool in an activity, the tool needs to be designed so that lower-level acts are operations and thus become withdrawn from the user's focus.

The conditions of the operation are designable. Thus, one can state that it is possible to design operations. However, it is impossible to design particular activities and actions; it is only possible to design *for* activities and actions. But there is one notable thing: since activities and actions *consist* of operations, the operation design will propagate to the higher levels.

The activity structure offers a coherent model to analyze an activity from the user's motives to the operation's particular conditions. As a consequence, every activity's structure is overwhelmingly complex. Therefore, it is crucial to limit the scope of the analysis to only include what is relevant for the purpose.

The hierarchical structure is used to understand how robotic telepresence mediates activities. The activity structures also form a foundation for understanding how robotic telepresence technology shapes the activities. The critical aspect is modeling essential parts of the activity mediated via robotic telepresence and the relations between the lower and higher levels via this structure.

This activity structure makes it possible to combine the sub-questions posed into one coherent whole: because, as indicated by the relation between the overarching research question and sub-questions, they are interdependent, and understanding them is to understand their relation to each other. For example, to understand the entirety of the social interaction between a pilot and a local user, it is necessary to understand how the pilot operates and experiences the robot and local environment, because these interrelate.

Mediation in activity theory

Mediation is one of the cornerstones when applying activity theory to HCI. The fundamental idea of mediation is that the subjects act through a mediational artifact toward the world. This is perhaps best captured by the title "Through the interface" (Bødker, 1991).

This crucial concept relates to the previously mentioned hierarchical level and the demand for lower-level acts classified as operations: because to focus *through* the tool, the focus cannot be *on* the tool. For example, if a user is piloting a robotic telepresence robot, the focus should be on the local environment and not on the input device.

The issue with the traditional concept of mediation is that it focuses on one particular artifact. However, as digital artifacts become more complex and interwoven into our everyday lives, we use multiple artifacts in various combinations in our activities.

Complex Mediation (Bødker and Andersen, 2005) and the Human-Artifact Model (Bødker and Klokmoose, 2011) discuss this issue by expanding the

traditional model using semiotics. These relationships can be described as related, chained, and leveled.¹⁰

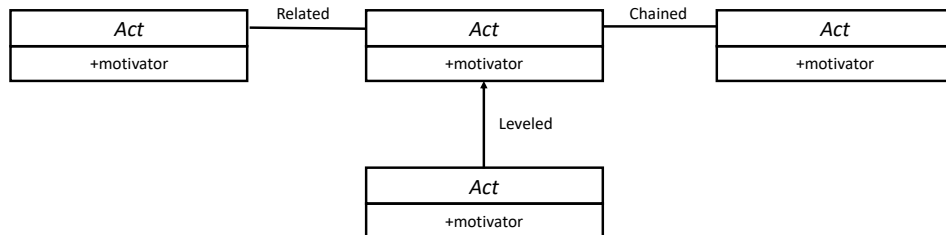


Figure 8: Relations between acts

Previous debates have elaborated on the “fluent” nature of the act, and the relationships that (Bødker and Andersen, 2005) outline are arguably relationships that all of the acts can have to each other; Figure 8 outlines these relationships as acts.

However, when looking at particular appropriations, the focus is on the activity. The description of an act *includes* the relationships to other acts within the activity. With this activity-centric approach, we view the various interactions and types of mediations that create a web of mediation that is both leveled and chained (Bødker and Andersen, 2005).

Leveled mediation is simply an activity that has sub-activities. The chain mediation is circular and has an "interconnectedness of activities through materials and mediators," i.e., the material and outcome become the mediator in a structured process.

Acknowledging this meta meditation is crucial for two aspects. Firstly, using the computer in such a way is not familiar to typical users and is error-prone. Secondly, the computer is a mediator for multiple activities that might interrupt the current ongoing activity.

Social dimension

As has been mentioned, activities are inherently social because humans are social beings. Also, the tools in an activity are manifestations of human knowledge (Leontiev, 1977). Furthermore, tools are influenced by the activity they were designed to mediate.

¹⁰ Leontiev acknowledged that multiple activities are related, but he never addressed “how” (Kaptelinin and Nardi, 2006).

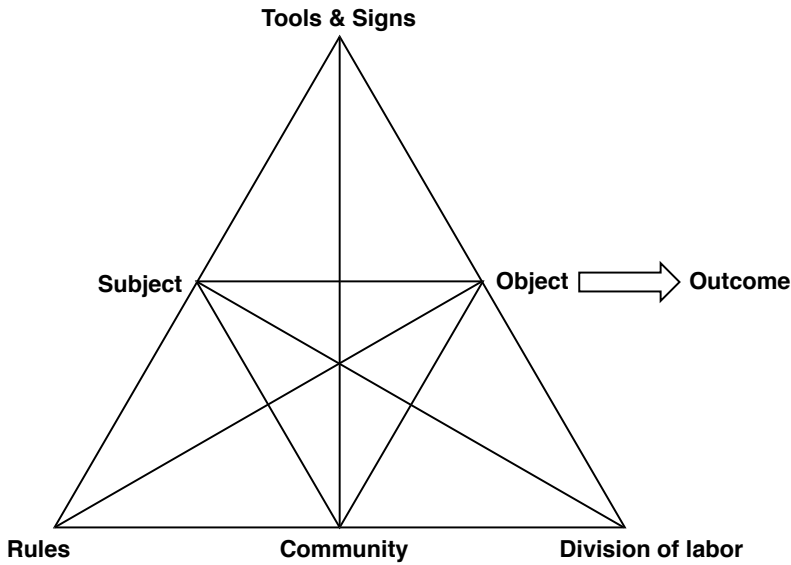


Figure 9: Leontiev's activity system modeled by Engeström (1987)

As seen in Figure 9, community was added to the original model, creating a three-way interaction between subject, object, and community, mediated by three different means. On the top, the relationship between subject, tools and signs, and object remains as previously described. At the base of the triangle are rules, community, and division of labor.

The community consists of people who “share the same object” (Engeström, 1987). Rules include laws, policies, and norms. The rules result from the context and govern the relationships between persons in the community. Thus, the rules are social conventions of the community that the subject needs to obey to fit into the community. The division of labor is how the community relates to the “shared” object; i.e., how can we reach the desired outcome as a community?

Another aspect of the activity system is that the object and outcome is separated. The object is what is shared and transformed into an outcome. The outcome of one activity system can be combined with the outcomes from other activity systems into a larger whole (for an overview of this, see Engeström and Sannino, 2021).

Digital technology and activity theory

Recent developments in activity theory have addressed the fact that people utilize an artifact ecology consisting of multiple artifacts (Bødker and Klokmoose, 2012; Kaptelinin and Bannon, 2012), and also the appropriation of tools into

instruments (Tchounikine, 2017). Furthermore, there is a strand of contemporary discourse that argues that current challenges are motivating the creation of a new generation of activity theory, including a fluent set of stakeholders (Spinuzzi, 2020), increased complexity of the object (Engeström and Sannino, 2021), and the flexibility of digital tools (Karanasios et al., 2021).

As our usage of digital artifacts has increased, we have created a dynamic ecology of artifacts (Bødker and Klokmoose, 2012). An artifact in such an ecological system can belong to multiple activities, creating a web of activity. Bødker and Klokmoose (2012) identified three states of artifact ecologies: unsatisfactory, exiting, and stable. In the unsatisfactory stage, the artifact ecology does not meet the requirements of the activity. In the exiting stage, the changes in the artifact ecology trigger an exploratory stage where the user explores new affordances of the technology. Lastly, in the stable state, the artifact ecology and activity have reached equilibrium.

To understand appropriation and how to design for it, Tchounikine uses activity theory and instrumental genesis to form a theoretical foundation (2017). The core idea is that technological introduction into an activity changes the activity and the instrument. In other words, the model acknowledges a form of a dialectical process where the resulting synthesis is a new/transformed activity. However, the idea of doing the activity only remains stable if no changes occur for a long time. Thus the designer needs to acknowledge that the tool will change the activity and that the activity will change the tool (Rabardel, 2001; Rabardel and Bourmaud, 2003; Tchounikine, 2017). Precisely how a human appropriates is an empirical question, and as we can understand from the definition of activities, the appropriation takes place in unique contexts.

The appropriation perspective adds a crucial component to the understanding of robotic telepresence usage by acknowledging that within the process of appropriation the robotic telepresence system goes from being a tool to being an instrument. During this appropriation process, both the activity and the tool will change. A tool becomes fluent through the appropriation perspective, and the user becomes an actor who does the final parts of the design.

The critical point for robotic telepresence appropriation into activity is that it addresses users' needs and requirements. The fundamental need that robotic telepresence technology can address is for users to be in a location other than the one occupied by their body. The challenge is that such a need is not limited to a small set of similar activities.

Appropriation occurs in collective activities, meaning that the effect of appropriation occurs for more than the primary user of robotic telepresence (i.e.,

pilots). When forming an understanding of how robotic telepresence technology can be used, it is necessary to acknowledge that the structures of activities change as the technology is appropriated.

Activity theory consists of three generations (Karanasios et al., 2021; Spinuzzi, 2020; Engeström and Sannino, 2021) who have expanded the subject of study.

The third generation of activity theory has adopted an interventionist stance by using co-design principles. However, as activity systems are increasingly unstable and now come with a fluent set of stakeholders, interventions based on solving tensions in activity systems based on the stakeholders' needs become infeasible (Spinuzzi, 2020). Consequently, Spinuzzi argues that activity theory requires a qualitative transformation to address these challenges.

Engeström and Sannino (2021) agree on the need for this particular focus. They still argue, however, that the main issue is the nature of the object that has caused the problem, where the object is not something that a single technological solution can solve.

Karanasios et al. (2021) complement the discourse by highlighting that today's digital tools are radically different from the tools analyzed by Vygotsky and Leontyev. The earlier tools were developed for particular activities, while digital tools are increasingly flexible and can have a more interwoven relationship with their users. One example is how integrated smartphones are into people's everyday life. The smartphone is a platform where the end-user can add apps for various activities.

Robotic telepresence directly relates to the fourth generation of activity theory challenges. As mentioned earlier, robotic telepresence is a tool designed by using knowledge about how a potential user might benefit from appropriating it in an activity. In a sense, the designer has an idea of usage while not knowing exactly how it might happen. The user might better understand how the tool can be a mediator in their activity. Additionally, the user might know how a tool fits into their activity; however, even while understanding their needs, it is still hard to predict how appropriation will occur.

Arguably, the robotic telepresence is a tool that mediates between the subject and the object, in a way that is similar to traditional technology. Additionally, robotic telepresence is also a technology that can affect the social dimension, i.e., the bottom part of Engeström's activity system (see Figure 9). It is evident that norms differ when using robotic telepresence compared to being there in person (see Lee & Takayama, 2011), and robotic telepresence pilots have other action possibilities than local users, resulting in another form of division of labor.

4.2 Phenomenology and Postphenomenology

“Phenomenology studies conscious experience as experienced from the subjective or first-person point of view.” (Smith, 2018)

Phenomenology focuses on how humans experience. Thus, the focus is not on the “objective” world. Instead, it is about our personal lifeworld (i.e., the experienced world). Consequently, phenomenology is a field with great diversity, and by and large it could be considered as a method (Farina, 2014).

Employing phenomenology on robotic telepresence focuses on robotic telepresence as a “phenomena,” i.e., how robotic telepresence is experienced and affects the “user’s” lifeworld.

Various forms of phenomenology have a long history in HCI and related fields, and on the whole such works are mainly dependent on a smaller set of works related to technology, such as selections of Heidegger's writings (Heidegger, 2010, 1977) and Merleau-Ponty’s (Merleau-Ponty, 2013).

Perhaps the most influential HCI works have been Dourish’s work on embodied interaction (Dourish, 2001). He uses phenomenology to form a philosophical foundation for HCI called “embodied interaction.” By analyzing tangible and social computing through embodied interaction, he highlights the embodied nature of computer-mediated actions.

In the field of science and technology, Don Ihde introduced postphenomenology as a philosophy to better understand the relationship between humans and technology (Ihde, 1990). *“The postphenomenological approach combines an empirical orientation with philosophical analysis,”* where the “post” in postphenomenology *“emphasizes that it distances itself from the romanticism of classical phenomenology”* (Rosenberger and Verbeek, 2015).

Postphenomenology aims at understanding the relationship between humans and technology by combining phenomenology and pragmatism. Postphenomenology is part of the “empirical turn” in technology philosophy which emerged during the 1980s and 1990s, the core of which was to have a pragmatic and descriptive stance toward technology. This stance was radically different from the classical philosophy of dystopian and deterministic technologies (Brey, 2010). Furthermore, postphenomenology was a way to update traditional phenomenology to include contemporary technologies (Ihde, 2009). In this work, the main focus is on the concept of embodiment and mediation. Consequently, related concepts such as intersubjectivity, intentionality, and how technology changes the lifeworld will also be utilized.

Embodiment

Embodiment is a concept which originates from phenomenology, with its roots in Husserl's late writings. Multiple philosophers have expanded the concept, including Heidegger (2010) and Merleau-Ponty (2013). Instead of “searching for a truth,” phenomenological philosophy focuses on how humans experience the world. To make sense of this experience of being in the world, it is not reasonable to have a model-based Cartesian dualism because the separation of mind and body is impossible when dealing with perception. Thus, Merleau-Ponty argues that the subject is inseparable “from this body and this world” (Merleau-Ponty, 2013).

“Embodiment” is also a term used for the process, and outcome, of learning how to use a tool with such efficiency that it becomes virtually transparent. Heidegger used the term embodiment when analyzing the relationship between a human and tools by using a hammer as an example. He argued that the hammer is “presence-at-hand for an inexperienced hammer-user.” When the user gains experience, s/he can use it with a low cognitive effort, and the hammer becomes ready-at-hand (Heidegger, 2010).

Merleau-Ponty extended this notion by adding the aspect of perception. In his two examples, the blind man's cane¹¹ and the feather on the hat, he illustrates that objects can become more than merely transparent; they can become an extension of the senses (Merleau-Ponty, 2013). The blind man's cane is suddenly an organ of perception; in the feather example, the wearer senses where the feather is, and thus it is included in the body schema. These embodiments are similar to how a driver embodies a car (Ihde, 1990). By using multi-sensory sensing, the car becomes both an extension of one's body, as in the feather example, and something that provides information about the world, such as the blind man's cane. All the while it is still a tool that the person acts through with the world, as with Heidegger's hammer. These examples show one crucial idea of Merleau-Ponty, namely that perception is more than the senses; human perception is embodied.

Merleau-Ponty uses the concept of body schema to outline how external objects can be “embodied” (2013). Body schema is the system that enables us to control our bodies efficiently without much of a focus. The embodiment of external objects is how these objects become part of our body schema. By being a part of the body schema, this perception can thus be extended to other artifacts, i.e., embodied perception, and arguably also into the digital realm (Svanæs, 2013).

¹¹ It is worth noting that Bateson (2000), independently of Merleau-Ponty, used the blind man's cane/stick to ask the question “Where does the person stop, and the human begin?” Bateson's example might be more famous in HCI since Hutchens used it to illustrate distributed cognition (1995).

Embodiment has affected technology research in general, and HCI in particular, through embodied interaction (Dourish, 2001). It has had a vital role in postphenomenology from the beginning (Ihde, 1990), and development has continued in this tradition (Ihde, 2010, 2002).

Embodied interaction (Dourish, 2001) is an approach for design and analysis where the embodied nature of our being is central. Dourish argues for using the fundamental phenomenological idea of embodiment as a center for understanding HCI. He does this by defining embodied interaction thus:

“Embodied interaction is the creation, manipulation, and sharing of meaning through an engaged interaction with artifacts.” (Dourish, 2001)

The focus on the fundamental idea of “embodiment” has meant that embodied interaction considers both the embodiment of tools and social embodiment. Here, the focus is on social embodiment in an upcoming section.

In postphenomenology, embodied technology forms a unit with the subject, notated as: (I-Technology)-World. Ihde emphasizes that the interaction with the technology is transparent so that the user experiences their actions to be toward the world. However, Ihde argues further that total transparency is impossible since the mediating technology alters humans’ relationships with the world. For example, a magnification is also a reduction in vision, or a hammer enables hitting, but at the same time occupies the hand (for further analysis about mediation see (Kiran, 2015)). Embodiment is when the tools become semi-transparent and become an integrated part of the body, thus changing the human’s action possibilities and perception of the world (Ihde, 1990).

While understanding embodiment from the simplest of examples, i.e., a person with a hammer, it is worth noting some additional levels of embodiment proven in psychology. It is essential to acknowledge that our senses collaborate to create “a whole.” Meanwhile, one of the most crucial senses for embodiment is proprioception – the sense of understanding the location of the body.

Due to proprioception, without looking we know the location of the hand and since the hammer has a fixed position in and extending from the hand, we know the location of the hammer as well. People who lack proprioception cannot locate their bodies and consequently have a hard time doing everyday tasks, such as walking without looking at their legs (Sacks, 1994).

The so-called Rubber Hand Illusion illustrates that vision can “replace” proprioception for people with fully functioning proprioception. The Rubber Hand Illusion is an experiment where the experimenter manages to get a

participant to sense that a rubber hand belongs to them, and through this hand the participants can “feel” touch (Botvinick and Cohen, 1998). In the original experiment, the participants felt that the rubber hand belonged to their “body image” but did not sense that they could do any actions (called body schema). However, by introducing control over the hand, in an experiment where one physical hand controls two fake hands (one right and one left), both of the hands are incorporated into the body schema (Newport et al., 2010).

Embodiment as a concept is thus strongly related to various forms of technology use. However, traditional embodiment analyses have focused on tools/instruments that we incorporate into our body schema. When applying the traditional notions of embodiment to robotic telepresence, we can see a form of embodiment between the input/output devices and the human. However, as we look at Figure 3, it is evident that there is another level of mediation. On the remote location, the pilot has input devices such as a microphone, keyboard, and mouse, used together with output devices such as screens and speakers. The pilot is thus interacting with the robot through equipment seen in regular personal computers, creating the following structure of embodiment: (I-Personal Computer)-Robot-World. In addition, there is also a relation between the person and the robot body.

Re-embodiment

The term re-embodiment suggests that an “embodiment” happens again. Thus, our first embodiment is our carnal body. We exist in this world as embodied beings. Re-embodiment is thus a state of embodiment other than the original. There are two types of re-embodiment in the literature: a re-embodiment of self (Aas and Wasserman, 2016; De Preester, 2011; Ihde, 2012; Papadimitriou, 2008; Standal, 2011) and a re-embodiment of another body (Besmer, 2015; Dolezal, 2009).¹²

The re-embodiment of “self” is when the person learns a new embodiment which is different from the original. Re-embodiment of “self” has its base in studies of rehabilitation where the patient has lost a bodily function and needs to appropriate technology to regain this capability, using things such as prosthetics and wheelchairs (Aas and Wasserman, 2016; De Preester, 2011; Ihde, 2012; Papadimitriou, 2008; Standal, 2011). For example, Papadimitriou (2008) looked at how patients with spinal cord injuries learn how to use a wheelchair to become “en-wheeled.” The re-embodiment of a wheelchair includes conducting certain

¹² The definition of re-embodiment and the categorization is compatible with all of the relevant articles, indexed by Google Scholar, that address “re-embodiment” in their title, or are referred to by these, i.e. (Aas and Wasserman, 2016; Besmer, 2015; De Preester, 2011; Dolezal, 2009; Ihde, 2012; Papadimitriou, 2008; Standal, 2011).

maneuvers and “adapting to a new self.” These maneuvers include the ability to do things such as a wheelie, i.e., standing only on the two back wheels.

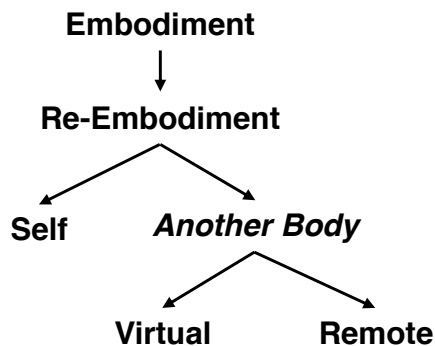


Figure 10: Overview of the concept of re-embodiment

The term “re-embodiment” highlights that this is a process of embodying a new embodied self. In the case of a person with a wheelchair, it is first a transition from being able to walk to being unable to walk, followed by learning to use the wheelchair. At first glance, the embodiment of a wheelchair might look similar to that of a hammer. However, there are two crucial differences. Firstly, the starting phase of this re-embodiment is not in the “original” embodiment; instead, it is in a body with altered capabilities. Secondly, the wheelchair aims to provide capabilities that existed in the original embodiment but cannot do so in the same way as the old body. Thus, the wheelchair allows the user to move, but not in the same way as walking.

Re-embodiment in another body comes from studies of virtual reality (VR) and robotic telepresence (Besmer, 2015; Dolezal, 2009). The experience is defined as “*given a sense for the user to be an active, perceiving agent, both ‘here’ and ‘there’*” (Besmer, 2015). Similarly, Dolezal defines it thus: “*In effect, these technologies hope to create a sense of re-embodiment, displacing the motor-intentional behavior of the body without rupturing the phenomenological coincidence of agency and ownership*” (Dolezal, 2009). The re-embodiment of another body differs from the re-embodiment of self in that the original embodiment exists in parallel to the re-embodiment.

The body re-embodiment can be *virtual* or *remote* (Besmer, 2015; Dolezal, 2009).¹³¹⁴ The main distinction between the remote and the virtual embodiment is that the virtual embodiment is an artificial, digital world created together with the virtual body, whereas the remote embodiment is in the same world as the original embodiment. This distinction is crucial because the physics in the virtual world are not the same as in the real world. The virtual world is inaccessible without mediating technology, and actions in the virtual world do not directly affect the real world.¹⁵

While the virtual embodiment differs from the remote embodiment, it is still essential to notice that humans experience VR similar to the original embodiment. For example, VR has successfully been used in exposure therapy to treat patients who suffer from anxiety disorders (Allom et al., 2016), particularly public speaking anxiety (Anderson et al., 2005).

There are crucial differences between re-embodiment of another body and self. Based on the empirical cases of extensional re-embodiment, Standal (2011) argues that re-embodiment is both a change in body image – “*the reconstitution of self-identity in relation to the person’s new bodily state*” (Seymour, 2012) – as well as a change in the body schema: “*rearrangements and renewals of body schema*” (Merleau-Ponty, 2013). It is highly questionable if the new body is incorporated into the body schema (Besmer, 2015; Dolezal, 2009) due to its distance from the carnal body.

Lastly, it is vital to address the significant differences that these various re-embodiments have on the lifeworld of the human. Extensional re-embodiment focuses on people regaining lost capabilities; thus, the artifact will become a key element in their lifeworld in these scenarios. Nevertheless, there might be future scenarios of re-embodiment that have less impact, such as wearing exoskeletons for conducting manual labor.

The empirical cases of re-embodiment of another body are only temporary compared to extensional re-embodiment. While the user feels an intense

¹³ Besmer (2015) uses the terms “Virtual” and “Robotic”, and Dolezal (2009) uses the terms “Virtual” and “Telepresent.” The rationale for me to use “Remote” instead of “Robotic” or “Telepresence” is that the differences between these two are the worlds where the bodies are.

¹⁴ It is also worth acknowledging that Paulos and Canny (Paulos and Canny, 1998b; Paulos, 2001) introduced the term “tele-embodiment” as a human-centered description of robotic telepresence. However, their usage of the term embodiment is not, to my knowledge, grounded in phenomenology. Meanwhile, I would argue that there is nothing in the term “tele-embodiment” that is not compatible with “remote embodiment.”

¹⁵ This discussion about virtual vs. physical could also relate to the previously mentioned issue regarding presence, and how studies in presence are mainly related to virtual worlds while robotic telepresence is in the physical world.

immersion in the other body, they can quit and go back to only embodying their original carnal body.

Applying the concept of re-embodiment to robotic telepresence is about learning to use the robot in a process of a remote embodiment of the robotic body. The robotic body exists in the same world as the user's carnal body; thus, the remote embodiment is essentially to learn how to use the body in the same physical reality as the user. Consequently, the actions mediated through the robot have the same consequences as the actions performed by the user's carnal body.

4.2.1.1 Social embodiment

Embodiment has a distinct social dimension as we are embodied in the world together with other people. As Dourish argues, *“we act in the world by exploring the opportunities for action that it provides to us – whether through its physical configuration or socially constructed meanings.”* Building up from this base, Dourish goes on to argue that embodiment is the common element that tangible interaction and social computing have in common (Dourish, 2001)

When we appropriate technology then, we do so in an inherently social lifeworld. Therefore, the embodiment itself is a part of the “social world.” Given previous accounts of re-embodiment, it is reasonable to analyze the reshaping of the body image as a process of re-embodiment. The new body image affects how we view ourselves and how others perceive us in the social world.

Going back again to the blind man's cane example, Merleau-Ponty argues that a fully sighted man can learn to use the cane so that it becomes “in hand” (Merleau-Ponty, 2013). However, his line of argument has been criticized based on his own construct. A blind man has a different form of embodiment in the world, and the blind man's blindness has direct social consequences (Reynolds, 2018). The societal distinction of able/disabled directly relates to the embodiment of normate/non-normate bodies. “Being blind” means that the lived body is non-normate.

While disability is a topic for medicine, it is clear that the experience of embodying a non-normate body forms another degree of understanding (Ghirotto, 2020; Paterson and Hughes, 1999; Reynolds, 2018). The non-normate body has social consequences; particularly, the intersubjectivity becomes altered.

This discussion of the non-normate body connects directly to the previous discussion regarding “re-embodiment of self” because it is a re-embodiment of a new physio-motoric structure and adaptation to an altered social situation and change in intersubjectivity.

A non-normate body is thus the embodiment of a body that is not in line with the normate body. Remote embodiment per se is, arguably, one kind of embodiment of a non-normate body; however, again, with the crucial difference in that such embodiment is temporary and that the embodiment of one's carnal body exists in parallel.

4.1 Activity Theory and Phenomenology

Multiple works discuss various consistencies between activity theory and phenomenology (Hasse, 2013; Kaptelinin and Nardi, 2012, 2006; Kosaka, 2013; Macdonald, 2000). This section is a summary of the debates, focusing on these two traditions' consistencies and their complementary nature.

The most significant similarities between activity theory and phenomenology focus on the individual and reject Cartesian dualism (Kaptelinin and Nardi, 2006). Both see the individual and the world as inseparable. Activity theory considers the relationship between humans and the world through the lens of activities, while phenomenology sees it through being. Furthermore, both perspectives regard humans as intentional subjects. Activity theory uses the concept of "object-orientedness," while phenomenology uses the concept of "intentionality."

Hasse (2013) combined activity theory and phenomenology to investigate the use of comforting robots (named *Paro*) in the field of dementia care. Her discussion highlights through the use of activity theory that the robot changes the activity structure and through phenomenology shifts the perspective to highlight the changed identity of the caregiver. This example beautifully highlights how activity theory and phenomenology can complement each other.

Both activity theory and phenomenology acknowledge experience. However, in phenomenology the question regarding the subject's experience is central, whereas activity theory has not – at least up until now – been giving experience the same level of attention.

Embodiment is interesting because it acknowledges the human body more than other related HCI theories (see for example Dourish, 2001; Svanæs, 2013). While arguably too much focus on human bodily capabilities can limit the perspective it has the potential to understand certain types of interactions (such as robotic telepresence).

Transparency is a concept that is evident in both activity theory and phenomenology; while the concept practically means the same thing in the two spheres, usage is slightly different. In activity theory, transparency is when the

tool is being used on a level of operation while the human focus is on higher goals. Thus, activity theory connects this to the operation concept, and the tool is transparent in its mediational role in the action.

Transparency in phenomenology is an *experience* of transparency (Ihde, 1990). Transparency is a prerequisite for embodiment relationships to be successful. However, not all tools enable an embodiment relationship, for example, a map. By combining transparency and operation, it is possible to explain the embodiment of a tool in this way: *a human uses a tool as an extension of her body to such a degree of efficiency that they have operationalized lower-level interactions to the extent that their focus is on higher-level goals.*

Both activity theory and phenomenology describe a state where humans can use artifacts to the point where it becomes transparent. Another similarity is that both describe how a breakdown causes the user to focus on the tool itself (Koschmann et al., 1998).

Combining these two perspectives leads to the conclusion that while the tool needs to afford an embodiment, it also needs to be designed so that humans can operationalize it. Consequently, if a breakdown on the operational level occurs, it also results in a breakdown of the embodied relation.

It is essential to acknowledge that this comparison between activity theory and phenomenology is made from the perspective of HCI. Consequently, they might appear more similar than they are. In addition, the purpose of the theoretical foundations which have been laid here is to understand how humans connect to the world through a robot.

4.1 Summary of Theoretical Foundations

Based on the theoretical foundations that have been laid, I would like to highlight the essential factors required to understand how humans are connected to the world through a robot.

Humans are inherently social beings; this is apparent from our direct social interactions and the socio-cultural foundation of many of our motives (Leontiev, 1977). Within human–human communication, there occurs a process of intersubjectivity that aims at creating an understanding of each other. Robotic telepresence thus acts in a social world, where the social dimension exists both between the individuals and as a part of the activity.

Humans act with intentionality stemming from their needs (Dourish, 2001); thus, technology is a mediator for meaningful human activities (Bødker, 1991). The

introduction of technology changes the context, i.e., changes the lifeworld (Ihde, 1990) and transforms activity (Tchounikine, 2017). Thus, the transformative nature of technology is manifested both in the objective and subjective world. This further highlights the fact that the introduction of robotic telepresence means reshaping the objective nature and the experience of an activity.

When effectively used, technology becomes transparent (Kaptelinin and Nardi, 2006) and can be seen as part of the human body (Ihde, 1990). Thus, learning to use a tool effectively means users become unaware of the tool. Commonly, this experience has, from a phenomenological point of view, been described as “embodiment.” However, the traditional notion of embodiment is about tools close to the body and not to other bodies. Thus, in light of this I introduce the concept of remote embodiment to describe the relationship between humans and the robot and how this relation connects the human to the world.

5 Research Design and Method

The methodological challenge of this thesis revolves mainly around the diversity of the studied phenomena. The necessary breakdown of the main research question into four sub-questions illustrates the phenomena's diversity. Furthermore, I look at the research questions from the dual perspectives of activity theory and phenomenology. This chapter consists of three main sub-sections: methodological considerations, study design and methodology, and the methods used in the five studies. The first sub-section highlights the overarching methodological considerations caused by both a subjective and objective side to the phenomena and deploying a research-through-design approach.

The second sub-section addresses how the methodological considerations are addressed. The last sub-section outlines the methodological decisions in each of the five studies presented in this thesis.

5.1 Methodological Considerations

Two significant aspects influence the methodological considerations in this thesis: the objective and subjective side of the phenomena and research through design.

The subjective and objective side of the phenomena

There is a subjective and objective side to the phenomena of robotic telepresence. The subjective side is what people experience/feel when using robotic telepresence, while the objective side is what is happening in the world.

This dual perspective directly relates to developments in HCI, wherein the focus of the initial wave of research was mainly on the objective side of the phenomena (Bannon, 1995; Bødker, 2015); the third wave's focus on user experience clearly conveys an interest in the subjective experiences of users (Bødker, 2006; Hassenzahl and Tractinsky, 2006).

The subjective and objective perspective permeates all the research sub-questions. The previously presented theoretical foundation arguably results in this dual perspective due to both activity theory and phenomenology acknowledging the need to understand both dimensions (while, as discussed, having different starting points).

The logical consequence of the dual nature of the research question(s) is that the methodological considerations need to account for both dimensions. The subjective side to the research question is addressed by asking the participants

about their experience, and the objective side is addressed through measurements and observations.

In addition, and since the data will be interrelated because the subjective experience is of something objective, it is essential to make methodological decisions that enable an analysis that can combine them. Therefore, I strive to collect subjective and objective data from the same situation.

Research through design

A research-through-design approach is employed, motivated by the aim to contribute to the design and usage of robotic telepresence. Consequently, a fundamental standpoint is that robotic telepresence is something that this research aims to change. According to Simon, "*Design is to devise courses of action aimed at changing existing situations into preferred ones*" (1996). Thus, studying robotic telepresence technology is to study existing and preferred situations and if a design changes the existing situation to a preferred one.

Research through design is an approach that aims at contributing both by designed artifacts and knowledge (Fallman, 2003; Hevner and Chatterjee, 2010; Zimmerman et al., 2007). The research-through-design approach places the researcher in a unique position where it is possible to combine technological opportunities with theories and empirical evidence into research artifacts (Zimmerman et al., 2007). Insights from studying such research artifacts then contribute to theoretical, technological, and empirical knowledge (Zimmerman et al., 2007).

Another possible contribution of a research-through-design approach is to improve design concepts (Höök and Löwgren, 2012; Stolterman and Wiberg, 2010). Strong design concepts make it possible to capture immediacy knowledge vertically grounded in theory (Höök and Löwgren, 2012).

The methodological consequence of said research-through-design approach is that the research revolves around the design and evaluation of robotic telepresence, where the designs build on previous designs and research, and the evaluation is of the design.

The stance in research through design also results here in combining hypothesis testing and exploration – the combination between hypothesis testing and exploration results in a mixed-methods approach.

5.2 Study Design and Methodology

The study design and methodology consist of multiple studies employing various forms of controlled experiments and scenarios. Conducting multiple studies is motivated by the inability to address all research sub-questions in a single study due to the diversity of the phenomena under consideration.

This sub-section outlines the underlying rationale for designing the controlled experiment and exploratory scenarios and how mixed methods are utilized for hypothesis testing and exploration. Lastly, the subjective and objective dimensions of the phenomena also affect how the mixed-method approach is implemented.

Controlled experiments and exploratory scenarios

This thesis is influenced by controlled experiments (Blandford et al., 2008) and exploratory methods focusing on the user experience (UX), such as experience prototyping (Buchenau and Suri, 2000), sketching user experience (Buxton, 2010), and user enactments (Odom et al., 2014, 2012).

Controlled experiments in HCI have their roots in psychology (Blandford et al., 2008). Controlled experiments are studies that isolate the phenomenon of interest and build on a hypothesis that there is a relationship between the dependent and independent variable(s) (Blandford et al., 2008).

One issue with hypothesis testing and controlled experiments is that they are limited in aiding the researcher to explore beyond their preconceived hypotheses because all of the data provided are the data that aims to address the hypothesis. For example, early usability evaluation can be harmful because it focuses only on the particular design, not the idea (Greenberg and Buxton, 2008). Therefore, in the studies included in this thesis, efforts have been made to combine controlled experiments with exploratory UX methods.

The exploration is twofold. Firstly, it seeks to explore if there are unanticipated differences between situations. Secondly, exposing potential users to novel situations primes them to imagine alternative futures, i.e., the participants can highlight possible future designs and use cases.

A common problem in HCI studies is the degree of "reality" needed to reach an acceptable degree of ecological validity. In short, the best source of data for addressing the research questions is from the "real world" for an extended period. However, the problem with such an approach is that constructing such a study makes it infeasible. This is especially problematic when evaluating novel designs since real-world usage puts higher quality demands on the prototype than a

controlled experiment. Thus, the aim is to conduct controlled experiments with realistic scenarios.

Mixed method

A mixed-method approach is taken, meaning that both qualitative and quantitative data are collected and integrated (Creswell and Clark, 2017). The mixed-method approach is adopted because quantitative methods enable hypothesis testing with a high degree of significance, which is extremely useful when comparing conditions. Meanwhile, qualitative methods enable an understanding of the situation beyond the experimenter's preconceived dimensions (Patton, 2014).

The previously mentioned dimensions of hypothesis testing/exploration and the subjective/objective side of the phenomena are related to using qualitative or quantitative methods (see discussion in Creswell and Clark, 2017; Patton, 2014). However, there is no exclusive and definitive relation between these.

Measuring the subjective experience of robotic telepresence is done by asking the participants about their beliefs. In the studies reported on, subjective data has been collected via quantitative methods, i.e., observations, semi-structured interviews, and questionnaires designed to afford quantitative analysis.

Measuring objective dimensions is mainly conducted with quantitative methods such as measuring performance indicators (for example time) and counting events that repeat. However, qualitative observations and interviews are complementary for understanding the quantitative and objective methods. These have an exploratory dimension, making it possible to see unanticipated behaviors.

To a large degree, hypothesis testing is carried out by utilizing quantitative methods. Through quantitative methods, it is possible to understand the significance of the findings. The interviews and observations are complementary as they enable understanding if some aspects are overlooked when forming the hypotheses.

As previously touched upon, I rely heavily on qualitative methods for exploration. Interviews especially enable an exploration of the design- and use-space by asking questions related to this.

Exploration of the quantitative data also has value since it is possible to find unanticipated relations. This is especially relevant when exploring quantitative subjective data, since it can uncover relationships previously not reported.

However, quantitative exploration needs to be handled with care since it runs the risk of appearing as hypothesis testing and thus hypothesizing after the results are known (HARKing) (Cockburn et al., 2018; Kerr, 1998; Wasserstein and Lazar, 2016).¹⁶ In the studies looked at, I have sought to clarify when exploratory analyses have been made and handle the findings with caution.

Hypothesis testing was based on the dependent variables being affected by the independent variables and the participants' backgrounds. The quantitative exploration included how different dependent variables correlated with each other.

Lastly, mixed methods enable a more thorough analysis since it is possible to use quantitative data to highlight patterns and qualitative data to understand the underlying reasons for such patterns. In particular, the interviews play a crucial role: they work as triangulation¹⁷, offer cross-validation, an understanding of the underlying reasons, and explorations. As will be seen in the next section, interviews are a part of the methodology of all studies conducted.

5.1 Study Methods

Five studies addressed the four research sub-questions. As seen in Table 4, each sub-question corresponds to one study, with that sub-question as its primary focus. Each study does, however, touch secondarily on multiple questions. As a result, each question is addressed in the main by one primary study, with three or four secondary studies also being of relevance.

¹⁶ At its most fundamental, it is about the process of trying to find statistically significant relationships in a dataset by looking at all possible combinations – without taking in to account that the probability of finding significant relations increases with the number of relationships explored – and presenting this as a regular hypothesis testing.

¹⁷ The use of activity theory and phenomenology suggests theoretical triangulation (Patton, 2014). However, while some aspects of theoretical triangulation are evident in this work, it has never been my intention to conduct theoretical triangulation.

Table 4: Summary of the studies related to research sub-questions. Each study relates primarily to one question and secondarily to multiple questions.

	Study 1	Study 2	Study 3	Study4	Study 5
	Probing the Design Space of a Telepresence Robot Gesture Arm with Low Fidelity Prototypes	Non-technical Users' first Encounters with a Robotic Telepresence Technology: An Empirical Study of Office Workers	Evaluating Input Devices for Robotic Telepresence	Performance, Power, and Place: User Experience of Contactless Object Manipulation in Robotic Telepresence	Exploring the Relationship Between Physical Presence, User Experience, and Task Parameters in Robotic Telepresence
How does the user operate the robot?	–	Secondary	Primary	Secondary	Secondary
How does robotic telepresence support performing actions in a distant physical location?	Secondary	Secondary	Secondary	Primary	Secondary
How does the pilot experience being in the robotic body?	Secondary	Secondary	Secondary	Secondary	Primary
How does the pilot experience being in the robotic body?	Primary	Secondary	–	Secondary	Secondary

All studies have secondary relations to the research questions because it is impossible to isolate a single question in a realistic study scenario. For example, all of the studies (2–5) where participants were asked to pilot the robot involved how they operated it, but only study 3 focused on this particular aspect.

Table 5 summarizes the methods used in the five studies. From an overarching perspective, it is worth noting that four of the five studies employed a mixed-mode approach. The exception, study 2, only employed quantitative methods.

A pilot study preceded every study. The motivation for a pilot study is to evaluate the overall methodology and train the experimenters. In cases where the pilot study resulted in significant methodological changes, another pilot study was then conducted to evaluate the changes.

The aim of *study 1* was to explore the design space of gesture arms for robotic telepresence. In study 1, the participants used six different lo-fi prototypes to design gestures suitable for three different scenarios. The quantitative data consisted of gestures and participants' scores and ratings of each gesture. The qualitative methods were interviews and observations.

Table 5: Overview of methodology in studies

Name	Context / Scenario	Design/ Prototype	Qualitative Methods	Quantitative methods
“Probing the design space of a telepresence robot gesture arm with low fidelity prototypes.”	<i>Gesture elicitation and assessment of a set of robot arm designs in different fictional use scenarios.</i>	Six different lo-fi gesture arms.	Interviews (S) Observations (S & O)	Gesture Elicitation (S) Rating (S) Ranking (S)
“Non-technical users’ first encounters with a robotic telepresence technology: An empirical study of office workers.”	<i>User study where participants piloted the robotic telepresence system in a realistic social scenario.</i>	–	Interviews (S) Observations (S & O)	–
“Evaluating Input Devices for Robotic Telepresence.”	<i>User study aimed to assess the impact of a remotely controllable system for extending the robotic telepresence-pilot’s action space in a realistic social scenario involving participants acting as a local user and pilot.</i>	Prototype of Double Remote Control system for enabling pilots to control objects in the local environment.	Interviews (S) Observations (S & O)	Ranking (S) Ranking (S)
“Performance, Power, and Place: User Experience of Contactless Object Manipulation in Robotic Telepresence”	<i>Comparison study of four different input devices in a controlled experiment where the participants, acting as a pilot, tried to complete a track as fast as possible and without failures with each input device.</i>	Prototype of feet input control consisting of a dance pad.	Interviews (S) Observations (S & O)	SDS ¹ (S) Ranking (S) Cognitive Load Questionnaire(S) ² Secondary Task (O) Time (O)
“Exploring the Relationship Between Physical Presence, User Experience, and Task Parameters in Robotic Telepresence”	<i>Controlled experiment for comparison of how two different factors, track difficulty and type of goal, affect the sense of presence for pilots.</i>	–	Interviews (S) Observations (S & O)	SDS ¹ (S) Performance (O)

1) Semantic Differential Scale

2) Influenced by NASA-THX

(S) Subjective, (O) Objective, and (S & O) Subjective and Objective

The role of the quantitative data was mainly to understand what designs and gestures were most suitable. The interviews and observations complemented the quantitative data by offering further insights into the participants' thoughts. All

of the data in study 1 was based on the participants' beliefs, i.e., the data was subjective, an immediate consequence of focusing on the social feasibility of gesture arms.

The aim of *study 2* was to explore potential users' anticipated appropriation of robotic telepresence. This was done by letting the participants experience robotic telepresence that could help them form an opinion on how they could appropriate the system. For the participants to get such an experience, they used a robotic telepresence system to explore a remote location with the help of an experimenter acting as a guide. During each scenario, all of the participants conducted a given set of tasks. The tasks were not explicitly known to the participants at the outset. Instead, the tasks were a natural consequence of how the researcher demonstrated the room for them. The primary data was collected by using semi-structured interviews. Field observations complemented the interviews by providing insights into how they managed the scenario.

Study 3 focused on how four input devices affect the performance of driving a track with the robotic telepresence system. The robotic telepresence system supported three devices: mouse, keyboard, and game controller. The fourth device, a dance mat, was specially configured to be used for this study. Study 3 relied extensively on quantitative data that included performance measurement, cognitive load (estimated and measured), ranking, and rating of the devices. The main analytical focus was on comparing the effects of the input devices. All participants used all of the devices; thus, it was possible to utilize both within- and between-subject comparisons. Within-subject comparisons were used to understand the difference of input devices, whereas the between-subject comparison was useful in investigating if particular previous experiences affected the performance.

The function of the interviews was twofold: to further develop an understanding of the participant's experience of the input devices and to explore the design space of input devices by idea elicitation.

The aim of *study 4* was to explore a prototype system called “Double Remote Control” (DRC). With the DRC system, the pilot of a robotic telepresence system can remotely control objects in the immediate environment, such as the TV, game pieces, blinds, and doors. The prototype consisted of a GUI containing all of the objects that the pilot could control. When the pilot conducted an action via the DRC GUI, an experimenter acting as a “Wizard of Oz” figure received a

notification.¹⁸ The “Wizard” then completed the action while striving to remain hidden from the participants.

The system was evaluated by having participants acting as pilots play a game with participants who acted as local users. All of the participants experience the scenario with and without DRC, but only as pilots or local users.

The data were gathered using a more developed version of a Semantic Differential Scale questionnaire previously introduced in Danielsson and Björnfot (2017), interviews, and video-recorded observations. The quantitative data was mainly used to compare the experience of DRC to non-DRC, but also to see differences in the experience between the user groups. The observations were used to analyze how the participants acted in the scenarios, and the interview was used to gain further insights into the users’ thoughts about the system.

Study 5 explored what factors affected the experience of presence. The participants conducted actions on a hard or easy track and focused on the outcome or process. Each of the participants did all of the combinations of track and goal. The primary measurement was a semantic differential scale developed for measuring presence. A within-subject comparison was used to compare results for each run, and a between-subject comparison to explore if previous experience affected the result. The subjective measurements were also compared to performance measurements to explore if performance correlated with the experience of presence.

Participants

Table 6 presents an overview of the participants in the studies. The total number of participants was 86, with 53 males and 33 females. The majority of the participants were connected to the university, i.e., either employed there or attending as students. Study 2 is an exception to this pattern, where only one participant was connected to the university.

The studies utilized a convenience sampling and snowballing method for recruiting participants. In four of the five studies, the selection criteria were that the participants should not have previous experience with robotic telepresence systems. In study 2, the selection criteria were that they should be non-technical office workers, i.e., not “IT-professionals.”

¹⁸ This study was preceded by a rather extensive pre-study (Kaptelinin et al., 2017) in which a TV and a lamp were manipulated through the DRC system without any involvement from the “Wizard”. However, the Wizard proved to be less prone to bugs and was thus applied for all actions in the main study.

Table 6: Participants in the studies

Study	Selection Criteria	N (Male/ Female)	Mean Age (min–max)
Study 1: “Probing the design space of a telepresence robot gesture arm with low fidelity prototypes.”	No experience of robotic telepresence	18 (10/8)	25 (21–30)
Study 2: “ <i>Non-technical users’ first encounters with a robotic telepresence technology: An empirical study of office workers.</i> ”	Non-technical office workers No experience of robotic telepresence	8 (5/3)	39 (25–56)
Study 3: “ <i>Evaluating Input Devices for Robotic Telepresence.</i> ”	No experience of robotic telepresence	16 (9/7)	34 (23–29)
Study 4: “ <i>Performance, Power, and Place: User Experience of Contactless Object Manipulation in Robotic Telepresence</i> ”	No experience of robotic telepresence	32 (20/12)	30–39 ¹
Study 5: “ <i>Exploring the Relationship Between Physical Presence, User Experience, and Task Parameters in Robotic Telepresence</i> ”	No criteria	12 (9/3)	20–39 ¹
Total		86 (53/33)	

1) The exact age was not collected

6 Summary of Studies

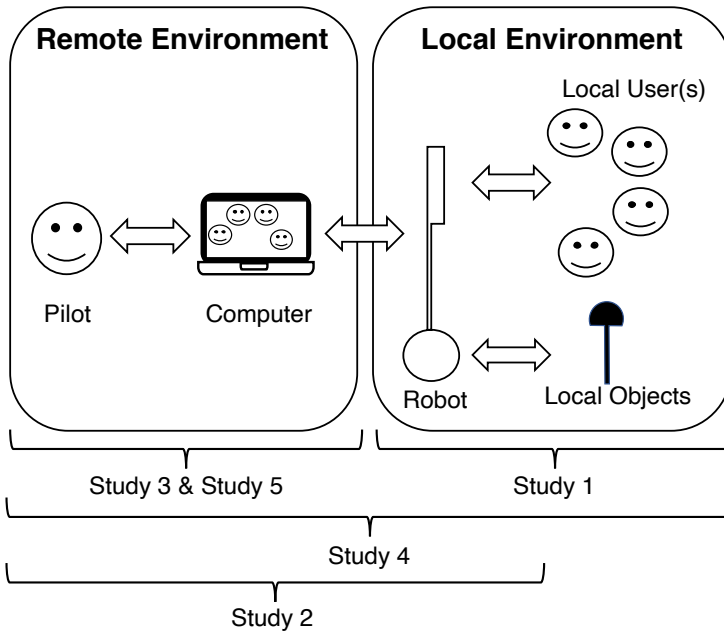


Figure 11: The five studies described in this thesis

This thesis describes five published studies. As shown in Figure 11: The five studies described in this thesis, these studies examined different aspects of robotic telepresence, focusing on the novel affordances that the technology provides to the users. Investigating different aspects of robotic telepresence was necessary to meet the aim of this thesis to formulate a holistic understanding of the phenomenon.

Study 1, “Probing the design space of a telepresence robot gesture arm with low fidelity prototypes.” described an exploration of how to design arm gestures for better communication between the local users and the pilot.

Study 2, “Non-technical users’ first encounters with a robotic telepresence technology: An empirical study of office workers.” focused on how potential users anticipated the benefits of, and initially appropriated robotic telepresence technology.

Study 3, “Evaluating Input Devices for Robotic Telepresence.” focused on the lower levels of interaction between the pilot and the telepresence robot by examining how various input devices affect performance and the experience.

Study 4, *“Performance, Power, and Place: User Experience of Contactless Object Manipulation in Robotic Telepresence”* examined how the pilot's increased action possibilities altered both local people's and pilot's overall experience. The added capability of the pilot was in the form of an Internet of Things-inspired remote controlling system.

Study 5, *“Exploring the Relationship Between Physical Presence, User Experience, and Task Parameters in Robotic Telepresence”* examined how various variables affected the feeling of presence. The primary variable was how the focus of a task affected the users' experience of spatial presence.

In the following sections, I outline each of the five studies' motivations, aims, key findings, and conclusions.

6.1 Study 1

Study 1 is reported in paper 1 *“Probing the design space of a telepresence robot gesture arm with low fidelity prototypes.”*

This study addressed the robotic telepresence system's lack of gesture capabilities by introducing and exploring the potential of using a gesture arm. One of the first telepresence robots reported, PRoP, had an arm (Paulos and Canny, 1998a). However, such arms have not, as yet, been adopted by commercial systems. Thus, in 2022, no commercial robotic telepresence systems have any gesture capabilities of this kind.

This study is linked to the research question concerning social embodiment and how local users perceive the robot. The study explored how to design a gesture arm to improve the social interaction between the pilot and the local user. The arms were designed to evaluate and explore what types of arms and gestures are suitable, as well as reveal important socio-cultural parameters, when designing robotic telepresence for social activities.

Thus, this study links to the main research question by exploring the potential improvements in the connection between the pilot and local users. The challenging aspect of creating gestures is that they need to communicate the pilot's intention with clarity and without being confusing whilst being experienced positively (so seen as polite and safe).

This study's rationale was to employ a simple, fast, and inexpensive method for understanding how to design a robot arm suitable for a telepresence robot. To do this, 18 participants evaluated six different lo-fi prototypes. These prototypes had

different joints: four had variations of human-like joints, and the remaining two had a sliding function (see Figure 12).

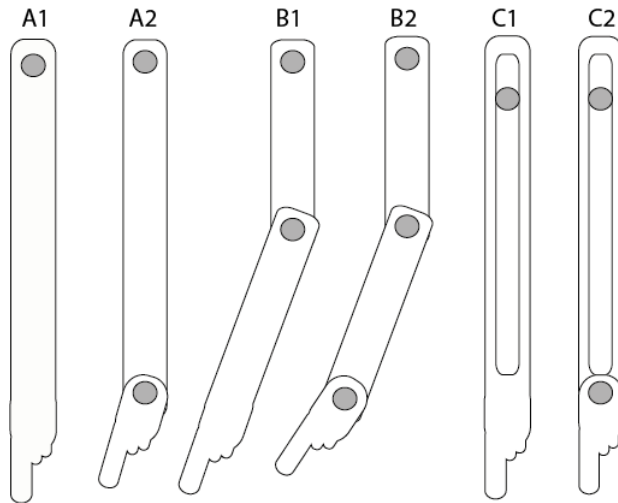


Figure 12. The designs of arm

While the most simplistic designs were the easiest to implement, they also worked as design provocations that uncovered underlying social aspects. The elicited gestures show that the participants preferred human-like gestures and uncovered their strategy to avoid the cultural problems of the simplistic designs' limited capabilities.

The participants created gestures with the six prototypes for three different scenarios. The participants rated the gesture's clarity, confusability, politeness, and perceived safety and ranked them from best to worst.

The social scenarios explored in this study were pointing at a person to gain their attention, signaling that a door was about to open, and pointing to an object of interest.

The study's result was that the most preferred arm was the one most similar to a human arm, not surprising given that the arm was to be used for social interaction and related to human-human interaction. Furthermore, the degree to which the participant transferred the social norms of human-human conversation into the robotic telepresence-mediated interaction was higher than anticipated. Thus, the design challenge is to create gestures that convey the message in a socially acceptable manner.

The key findings from this study were:

- The participants preferred the arm with the same number of joints as a human arm.
- The participants related the robot gestures directly to how they use their own body.
- Gestures mediated by the robot have social importance; the participants want the robot's gestures to be polite.

Making use of a human-like arm can be problematic. First, such a design can exhibit the “uncanny valley” problem, that is, people perceive the device as an expressionless human. Second, human likeness might convey human-like capabilities in cases where the actual capabilities are far below those of a human.

This study concluded that a lo-fi prototype seemed to provide an efficient way of exploring the design space of robotic gestures and that the participants related the gestures created to human gestures.

6.2 Study 2

Study 2 is reported in paper 2 *“Non-technical users’ first encounters with a robotic telepresence technology: An empirical study of office workers.”*

Since robotic telepresence technology is novel, most of the studies carried out included first-time users. However, no studies have focused on first-time users' initial appropriation of the robotic telepresence systems and their anticipation of future appropriation. This study aimed to address the research gap in novel users' initial appropriation and explore their predicted usage, so how their initial usage unfolded and how they wanted to use the robot.

Study 2 relates to the main research question by investigating what activities potential users want to use the system for, and their initial experience. Thus, participants' initial experience of being connected to the world through a robotic telepresence system worked as a primer to possible uses of the technology.

Study 2 aimed to understand how non-technical users initially appropriated the technology and wanted to incorporate it in their activities. The main focus was to understand the use space. A scenario to allow the user to try all the telepresence robot features and understand its capabilities was developed for this study.

All participants quickly learned to use the robot efficiently and completed the scenario without any significant issues – overall, the participants found the experience fun and exciting. The participants experienced a high degree of spatial

presence, and felt as if they were in the same environment as the robot. However, they struggled with understanding their "robot body," for example, how big they were.

The participants suggested multiple design improvements, and many of these related to the lack of body awareness. To improve body awareness, the participants suggested various improvements relating to proprioception and somatization as well as functionality to support the overall understanding of the robot body's position in relation to other physical objects. The challenge of judging distances caused social insecurity as the participants were afraid to invade the local user's personal space by standing too close.

The participants highlighted two kinds of anticipated appropriations tasks where the system enabled activities in the physical environment (e.g. exploring a warehouse) and tasks aimed toward social activities (e.g. elderly-care).

The key findings from this study were:

- The participants experienced a high degree of spatial presence: "It felt like being there."
- The participant had difficulty understanding their robotic body: "How big am I?" This made the pilot socially insecure.
- The participants had multiple suggestions on how to improve and use the system.
- The anticipated appropriations was motivated by either directing actions towards objects in the remote environment or social interaction.

The experience of high levels of spatial presence in contrast to the poor understanding of their body highlights a distinction between two types of experiences. Notably, this distinction is uncovered by the use of robotic telepresence because, in our everyday life, we do not have a separation between our carnal body and its location. In the light of this, it is understandable that the participants suggested improvements that would bring the experience closer to that of everyday being.

It is also evident that the lack of understanding of their body resulted in social insecurities, both in not knowing how they appear and not knowing if they positioned themselves in a socially acceptable way.

The division between using the system for actions directed toward the local environment or directed at people is notable. While one does not exclude the other, it is evident that actions directed towards the environment have received

less attention in the literature than social actions. Using the robot as a system for visiting locations is a function that no other established technology can provide. In contrast, numerous different technologies support social communication.

This study concluded that the use space consists of both socially and physically centered activities and indicates a distinction between the feeling of being in the body and spatial presence.

6.3 Study 3

Study 3 is reported in paper 3 *“Evaluating Input Devices for Robotic Telepresence.”*

The related research chapter describes commercial robotic telepresence systems that support different input devices. However, there have been no studies that have investigated how different input devices affect the pilot’s performance and experience.

Study 3 was an investigation into the effect input devices have on people’s experience and performance. Although the scope of this study covered just lower-level aspects, it resonated well with the concepts of transparency (phenomenology) and operation (activity theory). An understanding of the lower-level aspect being studied is fundamental for understanding the usage from an experience *and* activity perspective.

Study 3 relates to the main research question because it focused on the fundamental aspects of robotic telepresence usage, including task performance and user experience. Also, the measurement of cognitive load indicates the degree of cognitive effort needed to operate the robot and how much additional functionality might be possible for that cognitive load.

This study compared four input devices (keyboard, mouse, game controller, and dance pad) to understand the differences between the devices and which input devices are better suited to certain conditions. Furthermore, the study investigated how a user’s previous experience affected their usage. The study also highlighted promising future design directions based on the participants’ suggestions and use of the four input devices.

The participants drove around a track four times with a different track variation and input device each time. While driving, they also undertook a secondary task designed to measure the participant’s cognitive load.

The results yielded insights into which device to use, when, and by whom. It also gave insights into the further development of input devices.

The key findings from this study were:

- The game controller and keyboard performed best in the primary task, and the mouse and dance pad worst.
- The game controller's performance correlated with experience. The keyboard was best for non-experienced users.
- The keyboard had the lowest cognitive load.
- The participants believed the dance pad to be a fun and enjoyable alternative, but the implementation resulted in significant performance flaws, which produced frustration.
- The mouse was regarded as unsuitable for the task but yielded a good performance.

This study concluded that a game controller is the best device for experienced gamers or long-time users. In other cases, the keyboard is preferred. While using the body for controlling the device is a promising way forward, the lack of precision and accuracy, compared to the hands and fingers, could cause it to be inefficient.

6.4 Study 4

Study 4 is reported in paper 4 “*Performance, Power, and Place: User Experience of Contactless Object Manipulation in Robotic Telepresence*”

As highlighted in related research, one major issue with the current feature level of robotic telepresence technology is that the pilot has limited action possibilities in the local environment. Consequently, the pilot needs the help of local users to carry out everyday tasks such as opening doors, turning on lights, and operating projectors and monitors.

Study 4 investigated how local users and pilots are affected when the pilot can control objects in the local environment. Consequently, this study's broad scope and the rich data collection touch on more or less all of the leading research sub-questions, but with a particular emphasis on how the robotic telepresence supports carrying out actions in a distant physical location.

Study 4 relates to the research question by altering the connection with the local environment, which it does by introducing objects controlled by the pilot. A prototype system was created called Double Remote Control (DRC). The DRC system provides action capabilities to the pilot by adding remotely controllable

actuators in the local environment. The pilot controls these actuators with the same computer used for controlling the robotic telepresence system (see Figure 13). The system allowed the pilot to operate blinds, lights, TV screens, and open/close doors. However, the pilot still needed assistance moving chairs and rolling a die.

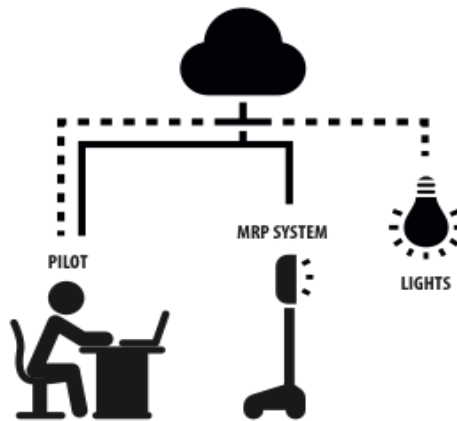


Figure 13: Sketch of the Double Remote Control system that enables the pilot to manipulate objects in the local environment. Note the term Mobile Remote Presence (MRP) is used instead of “Robotic telepresence”.

The study compared how pilots and local users experienced the DRC and non-DRC setups in a shared activity. The aim was to test the idea's feasibility and how it affected social interaction and users' experiences.

The key findings from this study were:

- The proposed DRC system is a promising solution for increasing a pilot's action possibility
- The DRC gave an increased degree of independence
- The DRC affected the social connection negatively
- The pilot related to DRC-controlled objects as abstract representations in the DRC setup.
- The local user believed the DRC to be almost “like magic.”

The DRC altered the social context. The participants approved of the pilot's increased independence. However, the pilot's increased independence resulted in less social interaction when carrying out various tasks, making the participants experience the DRC setup as less sociable. This suggests that there is a negative correlation between independence and social interaction.

This study concluded that apparent small technological changes altered the social context.

6.5 Study 5

Study 5 is reported in paper 5 *"Exploring the Relationship Between Physical Presence, User Experience, and Task Parameters in Robotic Telepresence."*

This study investigated which aspects affect the experience of physical presence when using robotic telepresence, by comparing two dependent variables: track difficulty and type of goal. The track was either difficult or easy. The goal was to either focus on carrying out the task as quickly as possible (outcome) or reducing the number of stops and turns, as well as avoiding collisions (process). The two dependent variables resulted in four conditions for each participant session.

Study 5 examined which parameters affect the sensation of physical presence. In particular, the study investigated how the structure of an activity affected the participant's subjective experience of being present as a robot. Therefore, the study's aim linked experience and activity.

The hypothesis was that participants would feel a heightened sense of presence in the outcome rather than in the process goal condition. The hypothesis was that the participants would focus their attention through the robot on the outcome condition and this would lead to a heightened sense of presence. In the process condition, the focus would be on the input device.

There was no significant difference between the two conditions or the difficulty of the tracks. However, gaming experience did correlate with the experience of presence, and the degree of presence did not correlate with performance. Furthermore, the gamers had an experience of being at both the remote and local settings *simultaneously*. The underlying causes of experienced gamers feeling a higher sense of presence are unclear. However, it might be because they are used to feeling a presence in the virtual world while playing, or that people predisposed to feel presence enjoy playing games to a larger degree.

The key findings in this study were:

- The feeling of being present correlated with self-reported gaming experience i.e. gamers felt a higher degree of presence.
- There was no significant correlation between performance and the feeling of being present.
- Task focus did not significantly affect the feeling of presence.

- The participants reported a higher presence during the last rounds than the first.

This study concluded that there are indications of a complex relationship between presence and the measured variables, which are particularly interesting for further research.

7 Remote Embodiment and Robotic Telepresence

In this chapter, the research questions of this thesis are addressed by applying the concept of remote embodiment to the findings from the five studies summarized above. The main points of the analysis are:

- After a short time, all the participants acting as a pilot could operate the robot to the extent that they could focus on actions in the local environment.
- The pilot feels ownership of the robotic body but struggles to form a body image.
- The pilot experiences spatial presence but lacks an awareness of the local environment.
- Social interaction through robotic telepresence relates to all the points above because it is a high-level action. In addition to this, social interaction also includes another user group, namely local users.

In the remainder of this chapter, these key findings will be discussed and analyzed as parts of the concept of remote embodiment.

7.1 Remote Embodiment

As previously addressed in the Theoretical Foundations chapter, an embodiment relationship between a person and an artifact forms when the person includes the artifact into their body schema (Merleau-Ponty, 2013). When a person embodies an artifact, it becomes a “transparent” extension of their body that influences their relationship with the world (Ihde, 1990). Thus, the artifact mediates the person's relationship with the world.

Re-embodiment is a term used to describe the embodiment of an artifact that results in a “new” body. This body can either be a “new” version of the person's existing body or another body that is either virtual or physical (Dolezal, 2009). As the prefix “re” suggests, it is about an embodiment that happens “again,” where the “first” embodiment is our carnal body. Re-embodiment where the tool extends the user, such as a wheelchair, alters the person's body schema and body image.

There are two forms of re-embodiment into another body: virtual and remote embodiment. Virtual embodiment refers to the presence of a person in a virtual environment through a virtual body. (see section Re-embodiment in the chapter Theoretical Foundations).

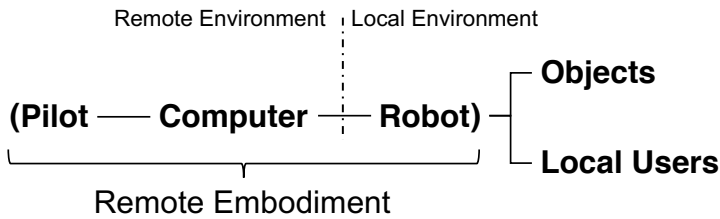


Figure 14: Remote Embodiment and robotic telepresence system

Remote embodiment refers to the presence of a person in a distant environment (usually referred to as "local environment") through another physical body, for instance, as a telepresence robot. In other words, when a person uses a robotic telepresence system, the person is remotely embodied as a *robot* in a *local environment* (see Figure 14).

Thus, to use Ihde's line of reasoning (see Ihde, 1990), a remote embodiment is the relationship between a person and body in another location (Person – Remote Body), where an artifact mediates a form of connection between the person and the body. In the case of remote embodiment and robotic telepresence, the person takes the role of a pilot, the remote body is a robot, and the connection between the pilot and the robot is mediated using a computer (see Figure 14).

Since the robot is a mediator between the person in the world, it shapes their connection to the world. The robot-mediated connection to the world is different from the "unmediated" connection due to the difference in perception and possible actions. While we can experience the mediating robot as "transparent," it only mediates a fraction of our corresponding unmediated experience.

When a person is remotely embodied as a robot, they feel present and carry out intentional and meaningful acts in a local environment. The process of becoming remotely embodied as a robot, therefore, is about learning to operate the system to such an extent that the person has body ownership, image, and schema of the robot, as well as a spatial presence in the local environment.

The local environment, where the person is present as a robot, often contains other people, referred to as local users. Robotic telepresence systems mediate interactions between a pilot and the local user(s). Therefore, remote embodiment occurs in a social context where the person remotely embodied as a robot is a social actor.

While remote embodiment is a phenomenological concept, it is also a type of mediation that I argue can be analyzed using an activity theoretical perspective.

The purpose of a person to be remotely embodied as a robot is to carry out meaningful activities through the robot. Thus, “remote embodiment” is a “type” of mediation that robotic telepresence technology enables. While this type of mediation is novel, it is still feasible to understand it as a tool defined by activity theory, that is, a tool that mediates between the subject and the object (Leontiev, 1977).

Appropriation of a tool into an activity changes the activity and the tool (Tchounikine, 2017). Since robotic telepresence offers the novel capability of being remotely embodied, it is hard to predict how the appropriation of robotic telepresence transforms an activity solely based on historical appropriations of other technologies.

The rationale for using activity theory and phenomenological perspectives is to understand how the introduction of robotic telepresence results in a multi-level transformation of human activity and experience. To form a holistic view of humans’ connections to the world through a robot, it is necessary to understand a remote embodied pilot as someone who carries out *intentional* and *meaningful* acts.

The rest of this chapter addresses the research question through the concept of remote embodiment. The first section focuses on how users operate the robot. Based on operating the robot, the following section discusses the aspect of a remote embodiment. The last parts discuss how robotic telepresence pilots carry out meaningful and intentional actions directed towards objects or people in the local environment¹⁹.

7.2 Operating Robotic Telepresence Technology

To be remotely embodied as a robot, the person needs to operate the system efficiently.

When applying activity theory, efficient control of the robot is considered to be when the pilot focuses on actions *mediated* by the robot rather than on the robot (Kaptelinin and Nardi, 2006). From a phenomenological point of view, efficient control is when the tool is “present-at-hand” (Heidegger, 2010) and experienced as transparent (Ihde, 1990).²⁰

¹⁹ It is worth acknowledging that the line between actions directed towards objects or local users is, as will be seen in the next chapter, somewhat blurred.

²⁰ As discussed in the section “Activity Theory and Phenomenology”, the activity theoretical concept of operation has similarities to the phenomenological concept of “present-at-hand”.

Overall, the robotic telepresence system used in studies 2 to 5 was easy to operate efficiently. The majority of the participants were first-time users who only needed a short training period before each session to complete the scenarios successfully and without any significant issues.

The participants operated the robot while carrying out actions such as navigating (study 4), speaking to local users (studies 2 and 4), and carrying out a secondary task (study 3). When carrying out higher-level tasks, it would appear that controlling the robot is an operation (Kaptelinin and Nardi, 2006) and that the users experience the technology as transparent (Ihde, 1990).

Study 3 provides clear evidence that input devices affect performance and cognitive load. While performance is crucial for task completion and avoidance of breakdown in the activity, the degree of cognitive load predicts how complex the mediated task can be. This study showed that the performance and cognitive load correlated positively, that is, participants performing well in the primary task had a lower cognitive load.

Study 3 did not conclude that a single device was better than the rest; instead, it highlighted that the game controller is preferred by experienced gamers or long-time users. For the others, the keyboard is the better choice performance-wise. The results indicated that the game controller could require more cognitive effort for experienced users than the keyboard. However, due to the limited sub-population of experienced game controller users, it was impossible to test the hypothesis. Furthermore, the results highlighted that the cognitive load can vary between similar operations.

A finding in study 3 was that lower-level interaction could be more or less suitable depending on the user's experience of how well the input commands correspond to the resulting movements of the robot. The dance pad had significant performance weaknesses, but some participants believed it added a degree of embodiment because their movements aligned well with the robot's movements. Participants suggested alternative input devices akin to the dance pad, including a moving mat, full-body tracking, and a balance pad. Utilizing the body in the interaction could be preferable as a tighter coupling between the body and technology allows for a deeper degree of embodied interaction/perception (Svanæs, 2013).

By contrast, the participants disliked the mouse because they felt that the modality was “wrong” for the task²¹. In addition, the participants also underestimated their performance with the mouse.

Study 4 increased the interaction complexity for the pilot by introducing both the “Double Remote Control” system for manipulating objects in the local environment and a shared presentation. The participants still completed the experiment without significant issues or breakdowns.

One particular finding in study 4 illustrates the importance of a good design of the lower levels of interaction. The participants frequently tried to move the robot when having the wrong window activated after issuing a DRC command, causing them to be confused over the robot's lack of response i.e. there was a breakdown of the operation “moving the robot” (see Kaptelinin & Nardi, 2006). What is noteworthy is that the participants repeated the breakdown multiple times, and the only learning effect observed was that the participants became faster at recognizing and solving the issue.

In conclusion, after a short period, the participants learned to operationalize the control of the robotic telepresence system used in these studies (Beam+ from SuitableTech). Some cases of breakdowns caused the operationalized control to become an action, but this was mainly related to the use of multiple systems in Study 4. Thus, the results indicate that the robotic telepresence system fulfills the prerequisite of remote embodiment, that is, that the control of the robot is transparent.

7.3 Being remotely embodied as a robot in a local environment

Remote embodiment is being in a robotic body in a local environment.

When the pilot can control the robot on an operational level, the technology becomes transparent, and enables the pilot to become remotely embodied as a robot.

There are two factors to being remotely embodied: being in the body and being in the local environment. However, since being in a body also means being in the body's environment, the two factors are interrelated to the extent that they are merely different starting points for analyzing remote embodiment. Furthermore,

²¹ One interesting observation was that in study 4, the participants could choose to use the mouse or keyboard, and all but one chose to use the mouse. However, I am hesitant to draw any concrete conclusions since the selection was not something that was adequately controlled in the study.

the robotic body's capabilities influence the experience of the local environment²².

Being the robot

Remote embodiment is a relationship between a person and a remote body in the real physical world. The pilot is simultaneously remotely embodied and embodied in their own body; these bodies exist in the same world but at geographically different positions.

The participant reported they had trouble understanding their robotic body. The lack of understanding resulted in insecurities, mainly seen in study 2 but evident in all studies where the participants used the robot. These insecurities included not knowing the robot's size, not knowing how close to an object the robot was, the consequences of bumping into objects, fear of falling with the robot, and general insecurity about the consequence of various actions and how they looked to others.

As highlighted in study 2, a limitation of robotic telepresence systems is that the only mediated senses are vision and hearing. These two senses are extremely flexible but become significantly limited when mediated. Vision and hearing interplay with our body's movement: we can respond instinctively to events in the room. Thus, vision and hearing are related to our bodily reflexes. Consequently, the mediation of vision and hearing is limited compared to how we experience it in our everyday lives.

A remote embodied relationship alters how the person perceives the world. In a robotic body, the action capabilities are radically different from those we are used to in our everyday lives. However, the lack of understanding of the robot's body in the remote location resulted in general insecurity about how close the robot was to physical objects and the consequences if the robot bumped into these objects.

One participant suggested that seeing herself in a mirror would have helped her understand the body better. Another strategy was to replace the need for bodily awareness by adding functionality to avoid breakdowns. Such functionality included employing sensors that would yield visual feedback when the pilot is potentially operating the robot in an undesired way, for example, driving too close to objects.

²² Gibson (Gibson, 1979) took, for example, a stance where the body affects the affordances in the environment.

It is hard for the pilot to understand their robotic body because the interface does not provide the same senses as our body, nor does it provide a third-person perspective. It is possible to provide the pilot with a third-person perspective, for example, using an external camera or, as one of the participants mentioned, a mirror.

It would be unreasonable to believe that it is feasible to reach a level of remote bodily understanding and awareness that is even close to our own. This is partly because of the senses mentioned above and partly because we constantly train our bodily awareness by existing in our carnal bodies. That is why remote embodiment will always be different from our embodied existence in our carnal bodies.

In conclusion, the participants experience a lack of connection to the robotic body they control, both regarding its body image and as a physical object in the local environment. However, they feel they have ownership of the robotic body. The participants related the issue to the lack of senses related to the body and forming a third-person image.

Being in the Local Environment as a Robot

Being remotely embodied means perceiving and carrying out meaningful activities in the local environment.

There are three factors relating to being in the local environment that have been observed: the sense of spatial presence, the understanding of the local environment, and the ability to act in the local environment. A crucial aspect of remote embodiment is the potential to carry out meaningful and intentional activities.

The lower-level operations occur in the pilot's location while the actions mediated by the system unfold in the robot's location, that is, the local environment. Thus, understanding the local environment is crucial for the pilot to carry out robotic telepresence-mediated actions. In other words, to carry out actions efficiently in the local environment, the pilot needs to feel as if they are "being" in the local environment. In study 2, a comment such as "it feels like being there" illustrated this sensation.

While the pilots struggled to understand their remote body, they experienced a *spatial presence* in the local environment, that is, they felt as if they were in the local environment. The pilots' experience of the local environment relates to their ability to operate and understand the robot because the robot mediates their entire experience of the environment.

In study 5, the variables that affect the degree of spatial presence relating to the participant's previous gaming experience were isolated. It was found that a high degree of spatial presence correlated positively with gaming experience and time. However, the degree of spatial presence did not correlate with task performance. Furthermore, the gamers seem to have the ability to feel present in both the local and remote environments and see the remote environment as unreal (i.e. as in a virtual world). Thus, this suggests a learning aspect that is also transferable from other areas where the users feel a spatial presence.

One participant contrasted the reported double-presence by describing that "events" in the remote environment made the feeling of presence disappear. Such events in the remote environment cause a shift of the pilot's focus from the robot's location to the body's location. This scenario is a breakdown in the experience of remote embodiment, that is, a breakdown in phenomenological terms. However, it is not necessarily a breakdown in an activity theory sense; instead, such a scenario can either be part of the activity in focus or part of another activity that the pilot needs to carry out.

A feeling of spatial presence does not necessarily mean that the pilot understands the surroundings. The robotic system is relatively inflexible in its affordance to create an overview of a physical environment. The robot's field of view is smaller than that of our eyes, and the entire body of the robot needs to move for the pilot to shift their visual focus.

What was evident was that creating a holistic understanding of the environment, followed by a general awareness of changes in said environment, was something that the participants driving the robot did *as a conscious act*. As observed in studies 2 and 4, the participants employed various strategies to understand the remote environment, including rotating and positioning themselves in a corner. Another example of the same aspect was that during study 4, participants stated that they would reposition themselves to see better. This reposition was a conscious act that, in many cases, would likely have been a subconscious act if the participants had been there in person.

Perceiving the environment is closely connected to acting in the environment. That robotic telepresence enables humans to be in a distant location is one of the main affordances highlighted by the participants in study 2. The ability to move freely around an environment offers the crucial ability to form an understanding, a "gut feeling," of the location such as when inspecting production lines and other tasks of a similar nature. In other words, the participants highlighted the value of being physically present regardless of social interaction and their limited action capabilities. In this case, perceiving the environment was described as an action rather than an operation.

Affording the pilot the ability to carry out actions in the local environment was the key aim of the system proposed in study 4. The approach is feasible for increasing a pilot's independence and can expand the possible use space for telepresence systems. From a functional point of view, the participants appreciated that the system provided a higher degree of independence for the pilots.

The design of the Double Remote Control system did not incorporate any form of action related to the robot. Thus, the pilot did not need to carry out any robotic body actions for issuing a Double Remote Control command. This separation from the actions and the body resulted in the pilot viewing the objects as abstract. For example, the pilots knew that they needed to start the TV monitor and to do this, they only needed to locate and press the correct button. Thus, the pilots did not need to know where the TV monitor was.

Two additional factors could have influenced this behavior. First, the effort of locating the objects in the environment could be too great for the pilot to bother doing. Second, a pilot's goal of completing the tasks overshadowed the purpose, thus creating a situation where the pilot could complete the tasks without actually understanding the consequences.

In conclusion, robotic telepresence affords the pilot the ability to be and act in another physical location. Robotic telepresence helps the user into feeling a connection to the local environment, that is, a sense of spatial presence. This spatial presence is a valuable feature and a unique experience. While the system supports this sensation, development is still needed to increase a pilot's awareness of the surroundings.

7.4 Social Remote Embodiment

Robotic telepresence works as a mediator of human interaction. In this social interaction, the pilot is remotely embodied and the local user is in-person.

One of the primary purposes of robotic telepresence systems is to mediate social interaction. The fact that the pilot is represented by the robot and the local user is in-person creates a fundamental difference between the two humans' experience of social interaction.

The pilot's interface is a computer, whereas the local user's interface is the robot. Additionally, the pilot controls the robot. Thus, the pilot is represented by, and is responsible for, the actions mediated by the robot. Thus, the pilot, computer, and robot form one unit. The pilot is the agent in the system, the computer is the mediator between the pilot and the robot, and the robot is the pilot's

representation in the local environment. The representational aspect is both social and physical. The robot is an artifact that enables the pilot to be spatially and physically present in the local environment. At the same time, the robot is the pilot's representation in any social interaction.

Social interaction is a high-level action that builds on the pilot's experience of being in the robot and the environment²³. Therefore, it will become evident that this section, to a large extent, builds on previous sections in this chapter.

Social body

The robotic body is a social representation of the remotely embodied pilot.

Our body plays a crucial role in our social being, from explicit gestures to questions concerning intersubjectivity: the robotic body plays a similar role when we are remotely embodied.

In everyday human-human interaction, our bodies represent us, both from the mere presence of the body as a physical object to how we efficiently use and perceive it in social situations. We can empathize with others by looking at their bodies; with little effort, we can see if people are happy, sad, or worried. The robotic body is the pilot's physical representation, and it is through this body the pilot conveys social cues. Thus, it is about understanding each other's experiences. The pilot will understand the local users' general experience of being in the local environment because the pilot has a point of reference in their own embodied being. However, the pilot's lack of body image shows that they are unsure about their body's social signals. As the pilot does not know how they look, they cannot understand how they are perceived and experienced by the local user.

While the participants displayed body image-related social insecurity in studies 2 and 4, no clear social "breakdowns" occurred. However, in study 2, the participants did not believe that it would be any problem to interact using the robotic telepresence system with people they already knew. However, the participants were concerned to appear "strange" when meeting people they did not know.

In study 2, one participant was unsure if she positioned the robot at a socially acceptable distance from the local user, highlighting worries that she was too close. Knowing the distance to objects was highlighted in the previous section as an essential aspect when navigating an environment, and in this case, it has a social dimension.

²³ Thus, neither can be separated from the pilot's operation of the robot.

However, judging the acceptable distance to a local user is complex. First, the pilot needs to understand the actual distance to the local users. Second, the pilot needs to understand the local users' experience, something challenging given that they struggle to form a correct body image. Furthermore, because the robot is another body, the acceptable social distance might not be the same as between two humans. Last, the correct social distance can differ between the local user and the pilot.

Furthermore, the robotic telepresence systems provide little help for local users to understand the pilot's situation, highlighted by local users asking if the pilot perceives particular things in the environment. For example, in study 4, the local user frequently asked the pilot if they could see the die²⁴.

As discussed in study 1, we also need to understand the degree of remote embodiment communicated to the local user, that is, the local user needs to understand what the pilot can control. This understanding is crucial because a mismatch might lead to awkward situations.

The local users tend to have a dualistic view of the pilot-robot relationship as both a robot and a human. Thus, it seems that they can see the pilot as both a person and a robot. This dualism was evident in study 4, where local users commonly refer to “the robot” carrying out various environmental actions while still interacting directly with the pilot.

The local users can only see the result of the pilot's operation and not the bodily movements that caused it. The low-level interactions occur in another physical location. It was most evident in study 4 where the participants, who acted as local users, were surprised by changes in the room caused by the pilot operating the Double Remote Control system.

Thus, two processes occur: the pilot needs to understand how the robotic body represents them in the eyes of the local user, while the local users need to understand the pilot's experience of the remote embodiment. This process is the same type of intersubjectivity that occurs when two humans interact in person but with the crucial difference that the two humans can better understand each other from the basis of their body embodiment.

Based on this knowledge, it is reasonable that the participants in study 1 experience robotic gestures as yielding a similar degree of cultural significance as common gestures. In study 1, the participants were positive towards human

²⁴ It is noteworthy that the Beam+ robot displays what the pilot sees on the monitor, however, this did not assist the participants acting as local users. A reasonable explanation is that the area displaying what the pilot sees was too small.

likeness, wanted deictic gestures that resembled those of humans, and believed that the politeness of the gestures was crucial. Additionally, many of the improvements that participants suggested (mainly in studies 1, 2 and 4) related to being more similar to a human, best summarized by the quote "The more human, the better."

Remote Embodiment in social activities

A pilot's remote embodiment enables them to carry out social activities in the local environment.

Social activity is when two or more people share a common or compatible motive or goal. However, for this section, activities in a social context are also included to avoid excluding relevant activities. Robotic telepresence enables a unique social interaction by introducing an ability for remote users to undertake a social activity in a distant environment.

In study 2, some participants indicated that the opportunity to have such social interactions with people in one specific location was a positive one. Such scenarios included visiting a warehouse to look at objects and interacting with a person who happened to be present.

The social interaction can either focus on a particular person or on whoever is in the physical location of interest. At one extreme, the location is secondary, and what matters is meeting a specific person, similar to making a telephone call. At the other extreme, the focus is the physical location, and anyone who just happens to be nearby becomes included in the pilot's activity.

In this thesis, study 4 is the most relevant study to consider when analyzing remote embodiment and social activities since the scenario in study 4 was a joint activity between a pilot and a local user. Study 4 examined how pilots and local users experienced alterations in the pilot's action capabilities. The pilot's action capabilities were increased by introducing a "Double Remote Control" system (DRC). The results indicated that while the DRC system is a feasible solution for improving a pilot's independence, the participants experience it as less sociable.

Thus, robotic telepresence enables novel forms of technology-mediated activities. However, as the following sub-section will highlight, social activities have a high degree of complexity that affects the rules and division of labor.

Model of scenario in study 4

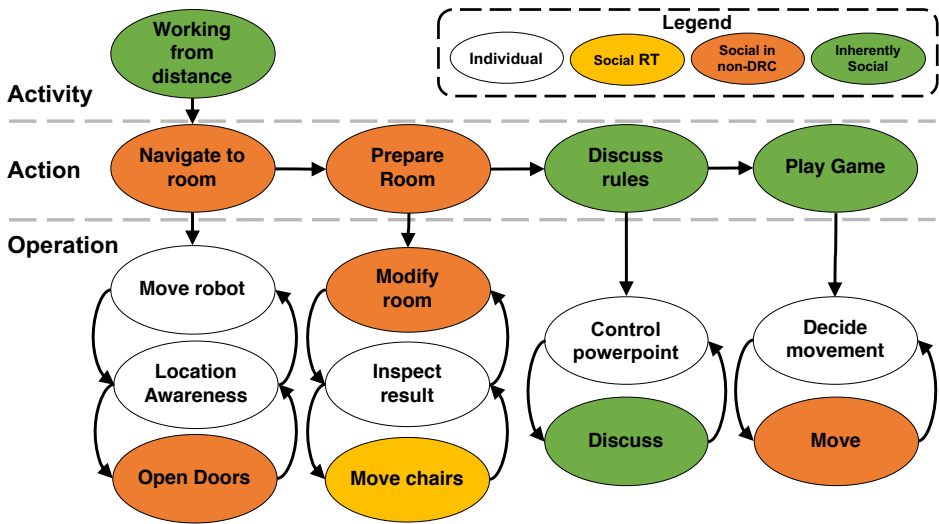


Figure 15: Outline of scenario part 1 in study 4

Figure 15 outlines the first part of each session in Study 4²⁵. There are four categories of acts: individual, social robotic telepresence, social in non-DRC, and inherently social.

Here, *individual actions* are actions that did not involve any social interaction. *Social robotic telepresence* refers to acts where the pilot needed help from a local user to carry them out. The DRC system addressed some of the pilot's action limitations, causing these acts to be non-social in the DRC condition and social in the non-DRC-condition; these acts are "*Social in non-DRC*." In other words, an act that is "social in non-DRC" is a consequence of the division of labor. Last, *inherently social* acts are social according to their nature, that is, the act involved another human.

The designed *activity* in this study was *inherently social*; the purpose of the scenario was for the pilot and local user to form a mutual understanding of a game and then play against each other.

Two *actions* are social in DRC, and two are inherently social. The *navigation to the room* consisted of three operations that occurred repeatedly: move the robot, create awareness, and open doors. Open doors was the only operation in the action that needed either actions using the DRC or through a human. However,

²⁵ The second part of each session began with a discussion of the rules and playing games, and ended with restoring the setup of the room and navigating to the robot's room.

what is not entirely evident is that, in the non-DRC condition, *navigating to the room* occurred in a social context because a local user walked to the room with the pilot.

In the *prepare room action*, the pilot needed assistance in every operation when using the robotic telepresence system, and only needed assistance in moving the chairs when using the DRC system. This combination of social and non-social operations creates a gray zone. In the non-DRC scenario, it was common for the pilot to explain the action's goal and state that there were a number of tasks. For the DRC scenario, the pilot only needed the local user to remove the chairs while not indicating any other deeper underlying goals. Thus, the pilot included the local user in the activity to a higher degree in the non-DRC than in the DRC condition.

The double remote control system design produced no relationship between the robotic body and the connected objects. Consequently, the local user could not perceive the pilot's actions and action space. If the pilot had carried out the activity in person, the local user would have anticipated the acts by observing the pilot's movements. This anticipation contrasts with the local user's feeling of "surprise" in the DRC condition. In a sense, the pilot had a set of actions that were invisible to the local user.

Furthermore, it was evident that some of the pilots treated the objects they could control as "abstract representations," but when instructing local users, the same objects would be "concrete." Thus, the design of the double remote control system invited the pilot to have an abstract relationship with the controllable objects in the local environment.

It is also possible to analyze the situation through the ideas of "power" and "control." The pilot had control of the scenario because they had the directions and were assigned a leadership role. This role became increasingly evident in the DRC scenario, where any interactions with the local user during the modification of the room operation(s) were superfluous. The role of the leader is a role with more power. However, in the non-DRC scenario, the pilot's power did not extend to the local environment. In the DRC condition, the power in the local environment shifts from the local user to the pilot. Suddenly, the pilot can alter the states of objects in a room where they are not physically present.

Discussing rules is an inherently social act, where the pilot controls the object of interest, for example, the PowerPoint presentation. The pilot could see the PowerPoint presentation on the same monitor used for controlling the robot, while the local user could see it on a TV. When seeing an abstract representation

of the object, a minority of the pilots positioned themselves in the same way they would have if looking at the presentation in person.

The game is also an inherently social act where the local user throws a die in both scenarios and moves the game piece for the pilot in the non-DRC scenario. Some participants did not understand that the pilot could control the game piece, a remote-controlled toy car without there being physical design features indicating that it could be remotely controlled.

During the last two actions, the pilot and local users frequently asked each other if they could see the result of their actions. The pilot asked the local user if the view of the PowerPoint presentation was good, and the local user asked if the pilot could see the result of the die throw. Thus, it is evident that the two user groups had difficulty understanding each other's perception of the environment.

From a remote embodiment perspective, study 4 could thus highlight the following aspects. First, the current remote embodiment of robotic telepresence limits the pilot and local user in understanding each other's situation. Second, the design of the double remote control system caused a detachment between the controllable objects and the bodily manifestation of the pilot. This detachment caused the pilot to treat the objects as abstract representations and reduced the local users' understanding of what would happen.

Third, while the pilot had less visibility of the local environment than the local user, the pilot had a virtual representation of the environment which was inaccessible to the local user. Thus, the pilot's control of the environment was invisible to the local user.

Last, the introduction of the DRC affected the division of labor and the degree of power that the local user had over the space. The effect of this caused the DRC scenario to appear less social. During both the conditions, the participants were unable to cooperate as if both were there in person because the power in the room was never equally divided.

7.5 The humans' connection to the world through remote embodiment

The main research question is: *"How does robotic telepresence influence humans' connection to the world?"*

Robotic telepresence is a new kind of mediation, quite different from earlier technologies. Humans connect to the world by being remotely embodied as a robot in a local environment. Being remotely embodied means being in the world

in another physical body. Remote embodiment is a relationship between the pilot and the robotic system.

Sub-question 1: *How does the user operate the robot?*

The participants who acted as pilots in the studies efficiently operated the robotic telepresence robot. The pilot's ability to undertake actions while driving the robot indicated that they had operationalized the system's controls. A pilot needs to efficiently operate the system to become remotely embodied as a robot.

Sub-question 2: *How does robotic telepresence support performing actions in a distant physical location?*

The pilots of robotic telepresence experience a spatial presence of being in the local environment. Thus, they feel as if they are “being there.” The sense of being in the local environment is related to being in the body. When sensing a being in the body, the body becomes transparent. However, the line is not clear per se because the feeling of spatial presence does not require an awareness of a body.

While current robotic telepresence systems limit a pilot's ability to modify objects in the local environment, the double remote control system proposed in study 4 presents promising possibilities for further development. A key aspect is that such a system might cause a perceived separation between the object's physical position and the pilot.

Sub-question 3: *How does the pilot experience being in the robotic body?*

Being in the body is a foundation of remote embodiment because this unification of the pilot and the robot enables meaningful activities to be carried out in the local environment. Remote embodiment includes three aspects related to the robotic body: ownership of the body, body image, and body schema. A pilot experiences a sense of ownership of the robot, referring to it as “theirs.” However, pilots are insecure about their body image and proximity to objects. The lack of body image and schema causes limitations in a pilot's experience of being in the robot body. Hence, it limits a pilot's experience of remote embodiment in the robot.

Sub-question 4: *How does robotic telepresence support social interaction?*

What is evident from this thesis is that social interaction is a critical element of being remotely embodied. Social interaction is a higher-level activity that relates to all three previous questions.

Remote embodiment has social meaning. The way the robot looks, how the pilot thinks they look, and the actions in the environment all have social consequences. Social remote embodiment differentiates itself from the regular usage of communication technology because it includes bodily and action elements.

When answering the questions through the concept of remote embodiment, it also becomes apparent that remote embodiment is a fundamental capability that makes it possible to use the system in many kinds of activities. Furthermore, remote embodiment can be a design concept and something that can be designed *for*.

In the next chapter, the future of robotic telepresence will be discussed by exploring remote embodiment as a design concept and by analyzing robotic telepresence as a multipurpose technology.

8 Possible Futures of Robotic Telepresence

This chapter focuses on the potential future development of robotic telepresence and research based on the previous chapter's concept of remote embodiment.

I employ two main perspectives on remote embodiment in this chapter. The first perspective is that remote embodiment is a fundamental quality of robotic telepresence that enables users to appropriate robotic telepresence in multiple activities. However, each activity has different requirements and, since there are numerous activities, there will be conflicting requirements. In the next section, I outline the issue and discuss potential ways forward.

The second perspective considers remote embodiment as a design concept vertically grounded in the phenomenological concept of the same name. By looking at it from this point of view, I discuss how to design for remote embodiment.

In the last sections of this chapter, I discuss additional research and theoretical questions involving robotic telepresence.

8.1 Robotic telepresence is a multipurpose technology

Digital technology can be more flexible than traditional technologies (Karanasios et al., 2021). However, people incorporate robotic telepresence into extant artifact ecologies, thus providing those systems with the ability to be remotely embodied as a robot in a local environment.

Remote embodiment can be used for many activities, thus making robotic telepresence a multipurpose technology. Consequently, further development of robotic telepresence will be constantly addressing conflicting requirements between these different kinds of activities. The issue of conflicting requirements is present in many types of multipurpose artifacts. So, in a sense, the problem category is not new. However, I argue that addressing the conflicting requirements is essential to realize the full potential of robotic telepresence systems.

The studies in this thesis, together with the Overview of Robotic Telepresence systems chapter, highlighted multiple different types of activities and actions where users could appropriate robotic telepresence technology (see Figure 16).

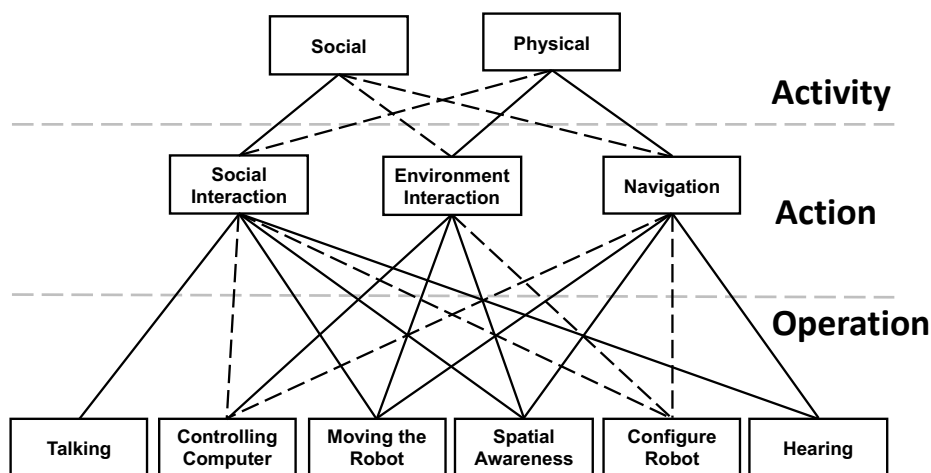


Figure 16: This sub-activity class diagram is an attempt to structure the classes of acts, and their relationships, as identified in the studies and related research.

The resulting model, shown in Figure 16, contains two classes of activities, three classes of actions, and six classes of operations²⁶. The relationships between the acts are either strong (black line) or weak (dotted gray line). A strong relationship is one that is inherent to the act; for example, navigation is related to moving the robot. A weak relationship is when an act is not inherently a part of it but can still occur, for example, when configuring the robot.

The model illustrates that there are relationships between all the different classes. This interconnectedness results in a high degree of complexity because every relationship is a dependency: altering one class affects all the other classes.

Socially-centered or Physically-centered activities

There are two main classes of activities: socially- and physically-centered. *Socially-centered* activities are activities where the robotic telepresence system's primary purpose is to mediate social interaction between humans, for example, office work and elderly care. *Physically-centered* activities are activities where the robotic telepresence system's primary purpose is to mediate the pilot's interaction with the physical environment, for example, visiting a museum, surveillance, and maintenance.

I use the term "centered" to indicate that many activities contain social and physical interaction but, often, one type can be considered as primary, and the

²⁶ It is worth noting that this model is created for the purpose of analyzing the multipurposeness of robotic telepresence. Consequently, the classes of acts outlined are not the only classes that exist, and it is worth acknowledging that the model is, to a large degree, subjective.

other secondary. For example, surveillance of a location is primarily a physically-centered activity. If a security guard uses a robot for surveillance and sees a person, the security guard will interact with the person. Similarly, elderly care is a socially-centered activity that contains physical aspects, for example, when the caregiver is helping their patient to find a particular physical object in the local environment.

During appropriation of a tool in an activity, instrumentalization can occur. Instrumentalization is when the appropriation of a tool changes the activity (Tchounikine, 2017). Thus, the activity changes when a tool offers different ways to mediate between the subject and the object.

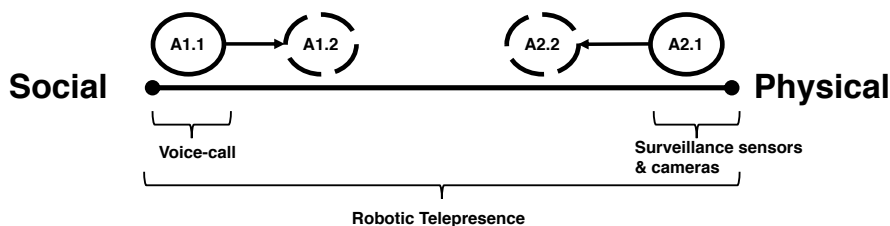


Figure 17: Illustration of how robotic telepresence causes instrumentalization of an activity that is either social or physical to become socially-centered or physically-centered.

Consequently, it is reasonable to believe that since robotic telepresence provides the users the capability to be remotely embodied, that is, carrying out physically-centered and socially-centered activities, its use will transform the activity into a blend of the two (see Figure 17).

For example, physically-centered activities mediated by surveillance cameras and sensors can be transformed into activities with more social interaction due to robotic telepresence technology, as illustrated by activity A2 in Figure 17. Similarly, activities mediated through voice calls or video-conference systems can transform into activities that contain physical factors when introducing robotic telepresence technology, as illustrated by activity A1 in Figure 17.

In conclusion, I argue that a robotic telepresence system has the potential to transform an activity by adding elements of physical or social interaction.

The diverse requirements within activities

The definition of a particular act is its class and relationship to other acts. A particular action *consists* of operations and is part of an activity, that is, a specific action is inseparable from the related operations and activity. Consequently,

actions and operations present different demands depending on what act they are a part of.

The definition of a particular act is its class and relationship to other acts. A particular action *consists* of operations and is part of an activity, that is, a specific action is inseparable from the related operations and activity. Consequently, actions and operations present different demands depending on what act they are a part of.

These conflicting demands are especially evident on the operational level, where an operation can belong to multiple different classes of actions that can occur after each other. For example, if the pilot first navigates to a place and then meets a person, the action changes from navigation to social interaction while the operation “moving the robot” continues.

As mentioned in Theoretical Foundations, the three classes of acts correspond to the three analytical questions “why?” (activity), “what?” (action), and “how?” (operation) (Bødker and Klokmoose, 2012).

Table 7: Operation requirements based on actions

Action Operation	Physical Interaction	Social Interaction	Navigation
Moving the Robot	High resolution	High resolution	Fast
Spatial Awareness	Object detection	Local User detection	Understanding the room Local user Detection
Hearing	Omni angle	Focus on the Local User	Omni angle

As shown in Table 7, multiple operations have conflicting requirements depending on what action they are a part of. The class “moving the robot” is straightforward from an operation perspective (“how?”): the user uses an input device to issue commands that will alter the robot’s position in the room. However, the operation alters when analyzing it as a part of a class of actions (“what?”), that is, Social Interaction, Environment Interaction, and Navigation.

During social interaction, *moving the robot* includes small re-positionings to direct attention to either a local user or object of interest. *Moving the robot* in the environment interaction is also about detailed movements, but only aimed at the environment. *Moving the robot* in the navigation action is different: it focuses on longer sequences of movements to reach a particular place.

The three action requirements for the operation are contradictory. The pilot wants to move the robot from point A to B as fast as possible in the navigation action. However, driving the robot quickly reduces precision, making it harder to carry out small re-positionings needed in the environment and social interaction actions. Many robots solve this issue by making it possible for the pilot to configure the maximum speed of the robot.

Conflicting requirements of different actions can exist within one activity because activities can contain many actions. Furthermore, operations from different actions can co-occur. Such a co-occurrence of operations was illustrated in study 2, where the participants could navigate while speaking with a local user. In this case, the participants carried out the operation “moving the robot” as a part of a navigation action while carrying out a “social interaction” action.

Addressing the conflicting requirements

Addressing the conflicting requirements can be achieved in multiple ways. First and foremost, the class of activity impacts the prioritizing actions taking place in the activity. If the activity is of the class “social presence,” it would be reasonable to downplay the requirements presented by navigation and environment interaction and, instead, prioritize social interaction because it is the prioritized action in the activity.

The problem with letting the activity determine the requirements is that the activity might fall between the definitions of physical and social activities, and the robotic telepresence system might be used for more than one activity.

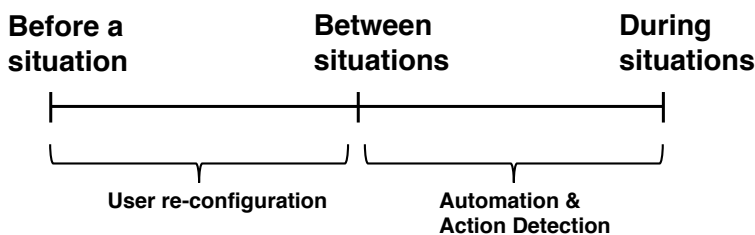


Figure 18: Different types of re-configuration occurrences and how these can either be addressed by user re-configuration or Automation & Action Detection

An additional approach is to divide the problem into how much time there is to adapt to changes (see Figure 18). Conflicts can be solved before a situation²⁷, between two different situations, and during the situation. For example, if a user is going to inspect a physical location where there are no people, there will be no requirements for social interaction so beforehand, the user can configure the system to best suit the activity. If the user then knows there will be another situation involving social interaction afterward, there is room to plan a re-configuration (see Figure 18). However, if the user first carries out movements as part of a social interaction and then needs to move a long distance, suddenly it is the navigation that is of utmost importance, and the robotic system needs to move faster.

I propose two strategies for solving these issues, enabling user re-configuration and various degrees of automation. User re-configuration builds on instrumentalization (Tchounikine, 2017) and focuses mainly on configuring the system for an activity (for example, before a situation). It partially covers issues between and during situations. The various degrees of automation focus mainly on the scenarios in the situation and partly on scenarios between situations.

Design for user re-configuration

User re-configuration means that the user can change the system to support the activity or action better. In particular, such a re-configuration is about changing how the user carries out operations. User re-configuration aligns with other multipurpose digital artifacts, including smartphones, smartwatches, TVs, personal computers, and web browsers. For example, when a person buys a smartphone, there are some basic applications that the smartphone designers believe are needed by the majority of the users. However, on top of this, the user can install apps that suit their particular needs. As a result, the user has instrumentalized the smartphone in various activities not accounted for in the original design. The user then re-configures the smartphone by switching between applications.

Re-configuration is an action of the user, and for this action not to disturb the main activity, it is preferable to carry out the re-configuration before or between situations.

Robotic telepresence systems can be re-configured using software and hardware. Software re-configurations involve changing the system's settings only with the user's input through the interface. Thus, software re-configurations include

²⁷ Note, I intentionally use the term “situation” instead of activity, action or operation, to highlight that any form of change, regardless of whether it is a change of activity (motives), action (goals), or operation (environment), can cause this type of requirement conflict.

changing the robot's maximum speed and height. Hardware re-configurations mainly involve changing the input and output devices used.

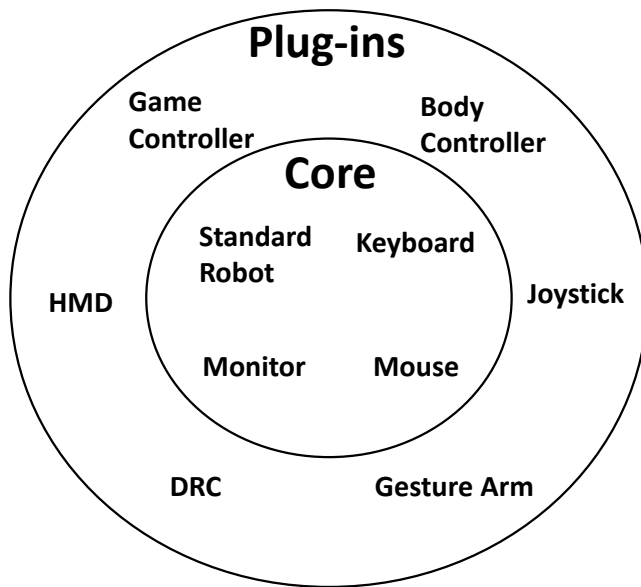


Figure 19: Plug-in Architecture of robotic telepresence systems

Thus, a possible way forward is to design robotic telepresence systems based on a plug-in architecture (see Figure 19). The platform/core is the robot and control software, and the input/output interfaces and various other additional features are plug-ins. Building such a plug-in architecture also involves separating the details from the overarching structure (Martin, 2017).

The core features are the features that the majority of activities require. Plug-in features are those that reduce the use space, reduce the potential user group, or are expensive. As Study 3 indicated, the default input device should be a keyboard. Still, if a higher degree of precision is needed and the user uses the system every day, a game controller would work better. However, since a game controller might be harder to use and is less common than a keyboard, it should be viewed as a plug-in.

As another example, head-mounted displays (HMDs) would be suitable for certain physically-focused activities because they enable users to look around by moving their heads. However, as an HMD hides the user's eyes, it might be less suitable for socially-focused activities.

Imagine the following scenario: a user is to inspect an environment. To do this, they use an HMD and a game controller (configuration before the situation). After the inspection, they talk to a colleague, so replace the HMD with a computer monitor so that the colleague can see their face when they are talking (re-configuration between situations). While speaking with the colleague, the user needs to show some pictures taken during the inspection to share on the screen. To operate the other artifact of the computer, they use the keyboard and mouse. As it is inconvenient to have to repeatedly pick up and put down the game controller, they use the keyboard to control the robot (re-configuration in the situation).

Design for automation

Automation in robotic telepresence is, to a large degree, synonymous with automating actions and automating re-configuration. Automating actions means that whole actions become operations. For example, navigation can be largely automated by enabling the pilot to select the location on a map, and have the robot drive there autonomously. Automated software re-configuration involves adapting the input sensitivity or loudness of the speaker to the context and situation.

Arguably, it is possible to create algorithms that can detect what class of action the pilot is carrying out and adapt the configuration based on the information. As an example, if the pilot is driving for long distances in a semi-open environment without talking to anyone, it would be reasonable to classify this as a navigation task. If the pilot is stationary, talking, and a face is detected in the video stream, the pilot is probably undertaking social interaction. Last, if the pilot is driving slowly with minor re-positionings without talking and only looking at the environment, the action is probably one of interaction with the local environment.

8.2 Design for Remote Embodiment

Designing for remote embodiment focuses on the unique aspect that makes robotic telepresence a multipurpose tool. To use Höök and Löwgren's (2012) ideas, remote embodiment can be a strong design concept that is vertically grounded in phenomenology.

In the introduction of this thesis, I introduced two influential contributions with radically different perspectives. Minsky (1980) envisioned a seamless connection between our body and a robot, while Paulos & Canny (1998) instead focused on providing “human skills” without relating the robot to the human body. While current robotic telepresence systems resemble the latter, the vision of a seamless connection between humans and robots is advancing ever closer.

In these two strategies, I see the future of robotic telepresence. On the one hand, we use cars efficiently without them resembling our bodies. The design of cars is, in many ways, far superior to that of our bodies when it comes to traveling. In addition, multiple features enable us to become better car drivers, including anti-skid systems, anti-lock braking systems and various servos.

On the other hand, the pilot uses robotic telepresence systems in similar situations as their carnal body. Thus, the carnal body is probably superior in many social situations. However, if development of robots with human-likeness is feasible, would it be advantageous to do so? Returning to the statement from a participant, “The more human, the better,” we can ask whether this is true. One issue is the uncanny valley problem (Wang et al., 2015), where a robot resembles a human to such a high degree that, at first glance, it is hard to see that it is a robot but, after a short while, the motion and the expressionless body become strange to observe. The second issue is that such a robot would still be an inaccurate representation of the pilot.

A focus on remote embodiment does not necessarily entail choosing a path. Designing for remote embodiment means designing a robot suitable to provide a user remote embodiment in a local environment. This suitability includes the possibility of efficiently controlling the robot.

There are examples of how to design for remote embodiment based on the studies in this thesis. In study 4, the pilot controlled the objects as abstract representations. Consequently, the local users were surprised by changes in the environment. An alternative would be that the pilot needed to look at the object to manipulate it. This solution would connect the robotic body to the object resulting in increased visibility for the local user and a concrete representation for the pilot. The concrete representation for the pilot would position both the object and the action in the local environment.

Another path is to connect the pilot's carnal body to the robot. In study 2, one such connection was provided by the dance pad. However, it was limited to only having the users control the robot with their feet. It is reasonable to connect the body to the robot and utilize our natural reflexes and movements. As previously argued, vision and hearing are closely connected to how we move our bodies.

For example, Thellman et al. (2017) connected a human to a humanoid robot using a head-mounted device and gesture tracking, making it possible for the pilot to control the robot's arms, eyes and head. Giving such flexibility to the pilot could improve their awareness of the body and the spatial presence in the local environment, and increase the experience of remote embodiment.

While such a design has some issues, it is still an approach for exploring the design space. Such a design can improve embodied perception (Svanæs, 2013) by combining vision and hearing with bodily movements. However, HMD devices cover the person's eyes, making it harder to experience natural social interaction. An alternative would be to utilize eye-tracking by having the robot carry out minor corrections so that it was always centered at the place the pilot was looking at during social interactions. This would make it possible for the local users to understand what the pilot is interested in.

Designing for remote embodiment is also about designing for breakdown avoidance. While the obvious candidate of such design would be on the lower-level interactions, it is also worth realizing that the previously mentioned automatization is a viable solution. Through automatic re-configuration, the robot will have the settings most suitable for the situation without any action from the pilot. The most suitable system settings are the ones that are least prone to causing a breakdown or a manual re-configuration.

I argue that the two ways forward for robotic telepresence can both use the concept of remote embodiment and are relevant. I believe that the vision described by Minsky (1980), in line with Thellman et al. (2017), has the potential to provide insights that can then be realized in robotic telepresence systems described in this thesis, such as one that builds on the ideas of Paulus & Canny (1998).

The robotic telepresence system will always create a filter between the pilot's carnal body and the robot. In current robotic systems, only the sense of vision and hearing are partly mediated. However, this is not necessarily an issue. Instead, it can be framed as a design opportunity. When a pilot is remotely embodied as a robot, they do not feel pain if the robot is hit, they do not run the risk of damaging their hearing or sight from noise or light, and they will not be exposed to harmful elements, airborne particles or weather.

Robot as a non-normate body

The big question regarding social interaction is how it will unfold as people become increasingly accustomed to robotic telepresence systems. While the nature of the question is empirical, it is still feasible to think about it based on the concept of remote embodiment.

The pilots are remotely embodied in the world through a robotic body. The robotic body is different from a human body, and thus the embodiment differs. The robotic body alters the action capabilities in the world and thus how the pilots experience the world. A pilot's experience is not evident to the local user.

The pilot of a robotic telepresence system has been described as being “a person with disabilities” (Takayama and Go, 2012), while in study 4, it was the case that the pilot had additional hidden capabilities that directly affected the local user. Relating to this, study 4’s participants saw the body as *both* a robot and a human.

The pilot’s robotic body is a non-normate body with different capabilities to a human body. Thus, the robotic body is another kind of body. From this perspective, it is not strange that the pilot can control objects in the room while having difficulty understanding the distance between the robot and an object in its proximity. Furthermore, it also captures the robot/human dualism.

Acknowledging that remote embodiment is being in a non-normate body acknowledges the immediate difference between the pilot and the local user. However, while it remains true that the pilot is in a non-normate body, it is crucial to acknowledge that the pilot’s remote embodiment is temporary. Thus, while remote embodiment has several connections to the embodiment of a non-normate physical body, the latter is permanent.

The robotic body that the pilot remotely embodies will always be a non-normate body regardless of how human-like the body becomes. Consequently, remote embodiment as a concept is not about moving towards human likeness; instead, it is about enabling intersubjectivity between the local user and the pilot embodied in the non-normate body.

8.3 Theory and Robotic telepresence

To understand how users connect to the world through robotic telepresence, I introduced the concept of remote embodiment. To understand further remote embodiment, I applied activity theory and phenomenology. These two traditions have been proven to be compatible and complementary in the context of this thesis (see section 4.1 Activity Theory and Phenomenology for further elaboration). They have two vastly different points of departure: phenomenology focuses on experience while activity theory is object-oriented. Arguably, most theoretical works on human and technology relationships can be categorized into one of these points of departure.

However, the broadness of their scope and the resulting volume of work and development in diverse fields is also a commonality between phenomenology and activity theory. Such a large body of knowledge is both intriguing and overwhelming. As a researcher in a narrow field that is not focused on the core questions of either respective tradition, one is always running the risk of either oversimplifying concepts or walking down paths that divert too far from one’s subject.

Thus, I believe that there are still areas in both traditions that can aid further theorization of remote embodiment. From phenomenology, the concept of intersubjectivity and "being-with-others" is, to a large degree, overlooked in this thesis. These two concepts might deepen the analysis of the asymmetry between the pilot and the local user, and might have the potential to elaborate further on the connection between being remotely embodied and being in a non-normate body (see Ghirotto, 2020; Paterson and Hughes, 1999; Reynolds, 2018).

From activity theory, I believe that the analysis of remote embodiment could be widened by including aspects of learning and more thoroughly analyzing how the pilot and local users' activity systems differ and relates in various situations and setups, i.e., utilizing Engeström (1987) to a further extent. Furthermore, robotic telepresence is a technology that further challenges the traditional activity theory definition of a tool due to its impact on the activity systems and its multi-purpose nature building further on Karanasios et al. (2021).

While the similarities and differences between phenomenology and activity theory have been discussed and demonstrated in this thesis, there are, arguably, still insights about how these can be combined.

Finally, the research on presence has only been touched on in this thesis due to its focus on virtual reality. Still, I believe that the research focusing on presence (see for overview Riva et al., 2014) has potential for further understanding of the phenomena of robotic telepresence, especially since efforts have been made to connect presence to activity theory (Riva et al., 2011) and phenomenology (Dolezal, 2009).

9 Conclusion

This thesis aims to address the research question: *“How does robotic telepresence influence humans’ connection to the world?”*

By addressing the research question, this thesis contributes to the research and design in the field of Human-Computer Interaction in general, and robotic telepresence in particular.

In conclusion, robotic telepresence systems offer the users a unique connection to the world. In this thesis, I have introduced the concept of remote embodiment to capture the connection’s uniqueness and place this into already existing theoretical foundations. I have framed remote embodiment as a phenomenon, feature, and design concept.

Remote embodiment is a phenomenon where users experience an embodiment in a physical body separated from their own body. Remote embodiment is similar to re-embodiment but distinguishes itself by being temporary and physical.

As remote embodiment is a fundamental feature for acting in a local environment, the users can use the systems in multiple activities, that is, robotic telepresence is a multipurpose system. As previously discussed, this causes various points of conflict that are evident during the appropriation process and in-situ user activities. Therefore, I suggested that the development of robotic telepresence should utilize a plug-in architecture and support action detection that enables automatic re-configurations.

As a design concept, remote embodiment is an effort to increase further the experience of being embodied in a distant physical body. While it is impossible to deny that exploration of remote embodiment as a concept will relate to the human body, it is still crucial to acknowledge that designing for remote embodiment is not synonymous with designing a human-like robot.

Future research related to this thesis could consist of continuing the studies presented, exploring remote embodiment, and further theorizing of robotic telepresence.

The continuation of the studies I wish to highlight are: a) further development and study of the robotic gesture arm introduced in study 1, b) comparison between user’s anticipated appropriation and their actual appropriation, d) continuing exploration of how remotely connected devices, introduced in study

4, further expand a pilot's action space, and e) further understand what factors affect the sense of presence in robotic telepresence.

There are various factors relating to the design concept of remote embodiment to explore, especially the question about human-likeness: "what are the benefits of human-likeness?" At present, we still have the technological capabilities to move far beyond the current commercial robotic telepresence system (see Thellman et al., 2017). Consequently, I believe that explorations built on full-body immersion into humanoid robots could yield valuable insights into human-robot interactions in general, and robotic telepresence in particular. Such design explorations could provide feedback to the more commercially available technologies.

To understand robotic telepresence, I used phenomenology and activity theory. While similarities and differences are discussed in this thesis and related works, I still believe that interesting findings are to be found that would make a combination between these even more fruitful and accessible.

Last, this thesis touched on the research of presence without connecting it with the concept of remote embodiment. The research on presence has mainly been concerned with virtual presence, but it seems that this body of research could benefit the understanding of remote embodiment. Thus, I believe that future research should explore the similarities and differences between virtual and physical presence to provide further understanding and development of remote and virtual embodiment technologies.

10 References

- Aaltonen, I., Niemelä, M., Tammela, A., 2017. Please Call Me?: Calling Practices with Telepresence Robots for the Elderly, in: Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction. ACM, Vienna, Austria, pp. 55–56.
- Aas, S., Wasserman, D., 2016. Brain–computer interfaces and disability: extending embodiment, reducing stigma? *Journal of medical ethics* 42, 37–40.
- Activity Theory: The Encyclopedia of Human-Computer Interaction, 2nd Ed. [WWW Document], n.d. . The Interaction Design Foundation. URL <https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/activity-theory> (accessed 1.26.17).
- Agah, A., Tanie, K., 1999. Multimedia Human-Computer Interaction for Presence and Exploration in a Telemuseum. *Presence: Teleoper. Virtual Environ.* 8, 104–111. <https://doi.org/10.1162/105474699566071>
- Agarwal, R., Levinson, A.W., Allaf, M., Markov, D., Nason, A., Su, L.-M., 2007. The RoboConsultant: Telementoring and remote presence in the operating room during minimally invasive urologic surgeries using a novel mobile robotic interface. *UROLOGY* 70. <https://doi.org/10.1016/j.urology.2007.09.053>
- All, S., Nourbakhsh, I.R., 2001. Insect Telepresence: Using Robotic Tele-Embodiment to Bring Insects Face-to-Face with Humans. *Auton. Robots* 10, 149–161.
- Allom, V., Mullan, B., Hagger, M., 2016. Does inhibitory control training improve health behaviour? A meta-analysis. *Health Psychology Review* 10, 168–186. <https://doi.org/10.1080/17437199.2015.1051078>
- Anderson, P.L., Zimand, E., Hodges, L.F., Rothbaum, B.O., 2005. Cognitive behavioral therapy for public-speaking anxiety using virtual reality for exposure. *Depression and anxiety* 22, 156–158.
- Ava Robotics [WWW Document], n.d. URL <https://www.avarobotics.com> (accessed 6.17.20).
- Bamoallem, B.S., Wodehouse, A.J., Mair, G.M., Vasantha, G.A., 2016. The impact of head movements on user involvement in mediated interaction. *Comput. Hum. Behav.* 55, 424–431.
- Bannon, L.J., 1995. From human factors to human actors: The role of psychology and human-computer interaction studies in system design, in: *Readings in Human–Computer Interaction*. Elsevier, pp. 205–214.
- Bateson, G., 2000. Steps to an ecology of mind: Collected essays in anthropology, psychiatry, evolution, and epistemology. University of Chicago Press.
- Bazzano, F., Lamberti, F., Sanna, A., Gaspardone, M., 2019. The impact of field of view on robotic telepresence navigation tasks. *Communications in Computer and Information Science* 983, 66–81. https://doi.org/10.1007/978-3-030-12209-6_4
- Bazzano, F., Lamberti, F., Sanna, A., Paravati, G., Gaspardone, M., 2017. Comparing Usability of User Interfaces for Robotic Telepresence, in:

- Richard, P., Yamaguchi, T., Braz, J. (Eds.), proceedings of the 12th international joint conference on computer vision, imaging and computer graphics theory and applications (visigrapp 2017), vol 2. Scitepress, av d manuell, 27a 2 esq, setubal, 2910-595, portugal, pp. 46–54. <https://doi.org/10.5220/0006170300460054>
- Beam Quickstart Guide, n.d.
- Beer, J.M., Takayama, L., 2011. Mobile Remote Presence Systems for Older Adults: Acceptance, Benefits, and Concerns, in: Proceedings of the 6th International Conference on Human-Robot Interaction, HRI '11. ACM, New York, NY, USA, pp. 19–26. <https://doi.org/10.1145/1957656.1957665>
- Berri, R., Wolf, D., Osório, F., 2014. Telepresence Robot with Image-Based Face Tracking and 3D Perception with Human Gesture Interface Using Kinect Sensor, in: Proceedings of the 2014 Joint Conference on Robotics: SBR-LARS Robotics Symposium and Robocontrol. IEEE Computer Society, pp. 205–210.
- Besmer, K.M., 2015. What Robotic Re-embodiment Reveals about Virtual Re-embodiment. *Postphenomenological Investigations* 55.
- Blandford, A., Cox, A., Cairns, P.A., 2008. Controlled experiments 15.
- Bødker, S., 2015. Third-wave HCI, 10 years later-participation and sharing. *interactions* 22, 24–31.
- Bødker, S., 2006. When second wave HCI meets third wave challenges, in: Proceedings of the 4th Nordic Conference on Human-Computer Interaction: Changing Roles. ACM, pp. 1–8.
- Bødker, S., 1991. Through the Interface: A Human Activity Approach to User Interface Design. L. Erlbaum Associates Inc., Hillsdale, NJ, USA.
- Bødker, S., Andersen, P.B., 2005. Complex mediation. *Human-computer interaction* 20, 353–402.
- Bødker, S., Klokmoose, C.N., 2012. Dynamics in artifact ecologies, in: Proceedings of the 7th Nordic Conference on Human-Computer Interaction: Making Sense Through Design. ACM, pp. 448–457.
- Bødker, S., Klokmoose, C.N., 2011. The human–artifact model: An activity theoretical approach to artifact ecologies. *Human–Computer Interaction* 26, 315–371.
- Botvinick, M., Cohen, J., 1998. Rubber hands ‘feel’ touch that eyes see. *Nature* 391, 756–756.
- Boudouraki, A., Fischer, J.E., Reeves, S., Rintel, S., 2021. “I Can’t Get Round”: Recruiting Assistance in Mobile Robotic Telepresence. *Proc. ACM Hum.-Comput. Interact.* 4. <https://doi.org/10.1145/3432947>
- Brayda, L., Mollet, N., Chellali, R., 2009. Mixing Telerobotics and Virtual Reality for Improving Immersion in Artwork Perception, in: Proceedings of the 4th International Conference on E-Learning and Games: Learning by Playing. Game-Based Education System Design and Development, Edutainment '09. Springer-Verlag, Berlin, Heidelberg, pp. 62–73. https://doi.org/10.1007/978-3-642-03364-3_8
- Brey, P., 2010. Philosophy of technology after the empirical turn. *Techné: Research in Philosophy and Technology* 14, 36–48.
- Buchenau, M., Suri, J.F., 2000. Experience prototyping, in: In Proceedings of the 3rd Conference on Designing Interactive Systems: Processes,

- Practices, Methods, and Techniques (DIS '00). ACM, New York, NY, USA, pp. 424–433. <http://dx.doi.org/10.1145/347642.347802>
- Burgard, W., Cremers, A., Fox, D., Hahnel, D., Lakemeyer, G., Schulz, D., Steiner, W., Thrun, S., 1999. Experiences with an interactive museum tour-guide robot. *ARTIFICIAL INTELLIGENCE* 114, 3–55. [https://doi.org/10.1016/S0004-3702\(99\)00070-3](https://doi.org/10.1016/S0004-3702(99)00070-3)
- Burgard, W., Trahanias, P., Hähnel, D., Moors, M., Schulz, D., Baltzakis, H., Argyros, A., 2003. Tele-Presence in Populated Exhibitions Through Web-Operated Mobile Robots. *Auton. Robots* 15, 299–316. <https://doi.org/10.1023/A:1026272605502>
- Buxton, B., 2010. Sketching user experiences: getting the design right and the right design. Morgan Kaufmann.
- Carroll, J.M., 2013. Human computer interaction-brief intro. *The Encyclopedia of Human-Computer Interaction*, 2nd Ed.
- Cesta, A., Cortellessa, G., Orlandini, A., Tiberio, L., 2016. Long-term evaluation of a telepresence robot for the elderly: methodology and ecological case study. *International Journal of Social Robotics* 8, 421–441.
- Chen, J.Y., Haas, E.C., Barnes, M.J., 2007. Human performance issues and user interface design for teleoperated robots. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 37, 1231–1245.
- Cheng, X., Jia, Y., Su, J., Wu, Y., 2019. Person-following for Telepresence Robots Using Web Cameras, in: *IEEE International Conference on Intelligent Robots and Systems*. pp. 2096–2101. <https://doi.org/10.1109/IROS40897.2019.8967645>
- Chua, Y., Tee, K.P., Yan, R., Li, L., Huang, Z., 2012. Towards More Engaging Telepresence by Face Tracking, in: *Proceedings of the Workshop at SIGGRAPH Asia, WASA '12*. ACM, New York, NY, USA, pp. 137–141. <https://doi.org/10.1145/2425296.2425320>
- Claudio, G., Luca, G., Luce, L.M., 2017. Interaction design for cultural heritage. A robotic cultural game for visiting the museum's inaccessible areas. *DESIGN JOURNAL* 20, S3925–S3934. <https://doi.org/10.1080/14606925.2017.1352895>
- Clemmensen, T., Kaptelinin, V., Nardi, B., 2016. Making HCI theory work: an analysis of the use of activity theory in HCI research. *Behaviour & Information Technology* 35, 608–627.
- Clotet, E., Martinez, D., Moreno, J., Tresanchez, M., Palacin, J., 2016. Assistant Personal Robot (APR): Conception and Application of a Tele-Operated Assisted Living Robot. *SENSORS* 16. <https://doi.org/10.3390/s16050610>
- Cobalt Robotics [WWW Document], n.d. URL <https://cobaltrobotics.com/> (accessed 6.17.20).
- Cockburn, A., Gutwin, C., Dix, A., 2018. Hark no more: on the preregistration of chi experiments, in: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. pp. 1–12.
- Cohen, B., Lanir, J., Stone, R., Gurevich, P., 2011. Requirements and design considerations for a fully immersive robotic telepresence system, in: *Proceedings of HRI 2011 Workshop on Social Robotic Telepresence*. ISTC-CNR Rome, Italy, pp. 16–22.

- Coltin, B., Biswas, J., Pomerleau, D., Veloso, M., 2012. Effective semi-autonomous telepresence, in: Röfer, T., Mayer, N.M., Savage, J., Saranlı, U. (Eds.), Robot Soccer World Cup XV. Springer-Verlag, pp. 365–376.
- Cosgun, A., Florencio, D.A., Christensen, H.I., 2013. Autonomous person following for telepresence robots, in: Proceedings - IEEE International Conference on Robotics and Automation. pp. 4335–4342.
<https://doi.org/10.1109/ICRA.2013.6631191>
- Creswell, J.W., Clark, V.L.P., 2017. Designing and conducting mixed methods research. Sage publications.
- Danielsson, K., Björnfot, P., 2017. A semantic scale for evaluating the UX of a MRP system, in: Proceedings of the European Conference on Cognitive Ergonomics 2017. ACM, Umeå, Sweden, pp. 59–60.
- De Preester, H., 2011. Technology and the body: The (im) possibilities of re-embodiment. *Foundations of science* 16, 119–137.
- Dolezal, L., 2009. The remote body: The phenomenology of telepresence and re-embodiment. *Human Technology: An Interdisciplinary Journal on Humans in ICT Environments*.
- Double Robotics [WWW Document], n.d. URL
<https://www.doublerobotics.com/> (accessed 5.7.18).
- Dourish, P., 2001. Where the action is: the foundations of embodied interaction. MIT press.
- Eagleton, T., 2007. Ideology: An Introduction, New and Updated Edition. ed. Verso, London ; New York.
- Edwards, A., Edwards, C., Spence, P.R., Harris, C., Gambino, A., 2016. Robots in the classroom: Differences in students’ perceptions of credibility and learning between “teacher as robot” and “robot as teacher.” *COMPUTERS IN HUMAN BEHAVIOR* 65, 627–634.
<https://doi.org/10.1016/j.chb.2016.06.005>
- Engeström, Y., 1987. Learning by expansion. Helsinki: Orienta Konsultit.
- Engeström, Y., Sannino, A., 2021. From mediated actions to heterogenous coalitions: four generations of activity-theoretical studies of work and learning. *Mind, Culture, and Activity* 28, 4–23.
- Fallman, D., 2011. The new good: exploring the potential of philosophy of technology to contribute to human-computer interaction, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 1051–1060.
- Fallman, D., 2003. Design-oriented human-computer interaction, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 225–232.
- Farina, G., 2014. Some reflections on the phenomenological method.
- Fitter, N.T., Joung, Y., Hu, Z., Demeter, M., Matarić, M.J., 2019. User Interface Tradeoffs for Remote Deictic Gesturing, in: 2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN). IEEE Press, pp. 1–8.
<https://doi.org/10.1109/RO-MAN46459.2019.8956354>
- Frauenberger, C., 2019. Entanglement HCI The Next Wave? *ACM Trans. Comput.-Hum. Interact.* 27, 2:1-2:27. <https://doi.org/10.1145/3364998>

- Ghirotto, L., 2020. Phenomenology and Physical Disability: for a Non-normate Body Policy. *Encyclopaideia* 24.
- Gibson, J.J., 1979. *The Ecological Approach to Visual Perception*.
- Giraff, n.d. URL <http://www.giraff.org/> (accessed 5.7.18).
- Goodridge, D., Marciniuk, D., 2016. Rural and remote care: Overcoming the challenges of distance. *CHRONIC RESPIRATORY DISEASE* 13, 192–203. <https://doi.org/10.1177/1479972316633414>
- Greenberg, S., Buxton, B., 2008. Usability evaluation considered harmful (some of the time), in: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. pp. 111–120.
- Halverson, C.A., 2002. Activity theory and distributed cognition: Or what does CSCW need to DO with theories? *Computer Supported Cooperative Work (CSCW)* 11, 243–267.
- Hasse, C., 2013. Artefacts that talk: Mediating technologies as multistable signs and tools. *Subjectivity* 6, 79–100.
- Hassenzahl, M., Tractinsky, N., 2006. User experience-a research agenda. *Behaviour & information technology* 25, 91–97.
- Hayamizu, A., Imai, M., Nakamura, K., Nakadaï, K., 2014. Volume Adaptation and Visualization by Modeling the Volume Level in Noisy Environments for Telepresence System, in: *Proceedings of the Second International Conference on Human-Agent Interaction, HAI '14*. ACM, New York, NY, USA, pp. 67–74. <https://doi.org/10.1145/2658861.2658875>
- Heidegger, M., 2010. *Being and time*. Suny Press.
- Heidegger, M., 1977. *The question concerning technology*.
- Herring, S.C., Fussell, S.R., Kristoffersson, A., Mutlu, B., Neustaedter, C., Tsui, K., 2016. The Future of Robotic Telepresence: Visions, Opportunities and Challenges, in: *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems, CHI EA '16*. ACM, New York, NY, USA, pp. 1038–1042. <https://doi.org/10.1145/2851581.2886423>
- Heshmat, Y., Jones, B., Xiong, X., Neustaedter, C., Tang, A., Riecke, B.E., Yang, L., 2018. Geocaching with a Beam: Shared Outdoor Activities Through a Telepresence Robot with 360 Degree Viewing, in: *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, CHI '18*. ACM, New York, NY, USA, p. 359:1-359:13. <https://doi.org/10.1145/3173574.3173933>
- Hevner, A., Chatterjee, S., 2010. *Design research in information systems: theory and practice* (Vol. 22). Springer. DOI 10, 978–1.
- Hevner, A.R., March, S.T., Park, J., Ram, S., 2004. Design science in information systems research. *MIS quarterly* 75–105.
- Höök, K., Löwgren, J., 2012. Strong concepts: Intermediate-level knowledge in interaction design research. *ACM Transactions on Computer-Human Interaction (TOCHI)* 19, 1–18.
- Hutchins, E., 1995. *Cognition in the Wild*. MIT press.
- Ihde, D., 2012. Postphenomenological re-embodiment. *Foundations of science* 17, 373–377.
- Ihde, D., 2010. *Embodied technics*. Automatic Press/VIP.
- Ihde, D., 2009. *Postphenomenology and technoscience: The Peking university lectures*. Suny Press.

- Ihde, D., 2002. *Bodies in technology*. U of Minnesota Press.
- Ihde, D., 1990. *Technology and the lifeworld: From garden to earth*. Indiana University Press.
- Izumi, M., Kikuno, T., Tokuda, Y., Hiyama, A., Miura, T., Hirose, M., 2014. Practical Use of a Remote Movable Avatar Robot with an Immersive Interface for Seniors, in: *Proceedings of the 8th International Conference on Universal Access in Human-Computer Interaction. Aging and Assistive Environments - Volume 8515*. Springer-Verlag, Berlin, Heidelberg, pp. 648–659. https://doi.org/10.1007/978-3-319-07446-7_62
- Johnson, S., Rae, I., Mutlu, B., Takayama, L., 2015. Can You See Me Now?: How Field of View Affects Collaboration in Robotic Telepresence, in: *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15*. ACM, New York, NY, USA, pp. 2397–2406. <https://doi.org/10.1145/2702123.2702526>
- Jones, B., Maiero, J., Mogharrab, A., Aguliar, I.A., Adhikari, A., Riecke, B.E., Kruijff, E., Neustaedter, C., Lindeman, R.W., 2020a. FeetBack: Augmenting Robotic Telepresence with Haptic Feedback on the Feet, in: *Proceedings of the 2020 International Conference on Multimodal Interaction, ICMI '20*. Association for Computing Machinery, New York, NY, USA, pp. 194–203. <https://doi.org/10.1145/3382507.3418820>
- Jones, B., Zhang, Y., Wong, P.N.Y., Rintel, S., 2021. Belonging There: VROOM-Ing into the Uncanny Valley of XR Telepresence. *Proc. ACM Hum.-Comput. Interact.* 5. <https://doi.org/10.1145/3449133>
- Jones, B., Zhang, Y., Wong, P.N.Y., Rintel, S., 2020b. VROOM: Virtual Robot Overlay for Online Meetings, in: *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems, CHI EA '20*. Association for Computing Machinery, New York, NY, USA, pp. 1–10. <https://doi.org/10.1145/3334480.3382820>
- Jouppi, N.P., Iyer, S., Thomas, S., Slayden, A., 2004. BiReality: Mutually-immersive Telepresence, in: *Proceedings of the 12th Annual ACM International Conference on Multimedia, MULTIMEDIA '04*. ACM, New York, NY, USA, pp. 860–867. <https://doi.org/10.1145/1027527.1027725>
- Kang, B., Hwang, I., Lee, J., Lee, S., Lee, T., Chang, Y., Lee, M.K., 2018. My Being to Your Place, Your Being to My Place: Co-present Robotic Avatars Create Illusion of Living Together, in: *Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services*. ACM, Munich, Germany, pp. 54–67.
- Kaptelinin, V., 2016. Supporting Referential Gestures in Mobile Remote Presence: A Preliminary Exploration, in: *Proceedings of the 14th International Conference on Inclusive Smart Cities and Digital Health - Volume 9677*. Springer-Verlag, Wuhan, China, pp. 262–267.
- Kaptelinin, V., Bannon, L.J., 2012. Interaction design beyond the product: Creating technology-enhanced activity spaces. *Human-Computer Interaction* 27, 277–309.
- Kaptelinin, V., Björnfot, P., Danielsson, K., Wiberg, M.U., 2017. Mobile Remote Presence Enhanced with Contactless Object Manipulation: An Exploratory Study, in: *Proceedings of the 2017 CHI Conference*

- Extended Abstracts on Human Factors in Computing Systems. ACM, pp. 2690–2697. <https://doi.org/10.1145/3027063.3053204>
- Kaptelinin, V., Nardi, B., 2012. Activity theory in HCI: Fundamentals and reflections. *Synthesis Lectures Human-Centered Informatics* 5, 1–105.
- Kaptelinin, V., Nardi, B.A., 2006. Acting with technology: Activity theory and interaction design. MIT press.
- Karanasios, S., Nardi, B., Spinuzzi, C., Malaurent, J., 2021. Moving forward with activity theory in a digital world. *Mind, Culture, and Activity* 1–20.
- Kerr, N.L., 1998. HARKing: Hypothesizing After the Results are Known. *Pers Soc Psychol Rev* 2, 196–217. https://doi.org/10.1207/s15327957pspro203_4
- Khojasteh, N., Liu, C., Fussell, S.R., 2019. Understanding Undergraduate Students’ Experiences of Telepresence Robots on Campus, in: Conference Companion Publication of the 2019 on Computer Supported Cooperative Work and Social Computing, CSCW ’19. Association for Computing Machinery, New York, NY, USA, pp. 241–246. <https://doi.org/10.1145/3311957.3359450>
- Kiran, A.H., 2015. Four dimensions of technological mediation. *Postphenomenological Investigations* 123.
- Kiselev, A., Kristoffersson, A., Melendez, F., Galindo, C., Loutfi, A., Gonzalez-Jimenez, J., Coradeschi, S., 2015a. Evaluation of using semi-autonomy features in mobile robotic telepresence systems, in: 2015 IEEE 7th International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM). Presented at the 2015 IEEE 7th International Conference on Cybernetics and Intelligent Systems (CIS) and IEEE Conference on Robotics, Automation and Mechatronics (RAM), pp. 147–152. <https://doi.org/10.1109/ICCIS.2015.7274564>
- Kiselev, A., Loutfi, A., 2012. Using a mental workload index as a measure of usability of a user interface for social robotic telepresence. Presented at the 2nd Workshop of Social Robotic Telepresence in Conjunction with IEEE International Symposium on Robot and Human Interactive Communication 2012.
- Kiselev, A., Scherlund, M., Kristoffersson, A., Efremova, N., Loutfi, A., 2015b. Auditory Immersion with Stereo Sound in a Mobile Robotic Telepresence System, in: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction Extended Abstracts. ACM, Portland, Oregon, USA, pp. 55–56.
- Koceska, N., Koceski, S., Zobel, P.B., Trajkovik, V., Garcia, N., 2019. A Telemedicine Robot System for Assisted and Independent Living. *Sensors* 19, 834. <https://doi.org/10.3390/s19040834>
- Kosaka, T., 2013. A foundation of a first-person perspective systems analysis.
- Koschmann, T., Kuutti, K., Hickman, L., 1998. The Concept of Breakdown in Heidegger, Leont’ev, and Dewey and Its Implications for Education. *Mind, Culture, and Activity* 5, 25–41. https://doi.org/10.1207/s15327884mca0501_3
- Kratz, S., Ferriera, F.R., 2016. Immersed Remotely: Evaluating the Use of Head Mounted Devices for Remote Collaboration in Robotic Telepresence, in: 25th IEEE International Symposium on Robot and Human Interactive

- Communication, RO-MAN 2016, IEEE RO-MAN. IEEE, 345 E 47TH ST, NEW YORK, NY 10017 USA, pp. 638–645.
- Kristoffersson, A., Coradeschi, S., Loutfi, A., 2013. A review of mobile robotic telepresence. *Advances in Human-Computer Interaction* 3. <http://dx.doi.org/10.1155/2013/902316>
- KUBI Telepresence Robot | Welcome [WWW Document], n.d. URL <https://www.kubiconnect.com/> (accessed 12.22.20).
- Kuzuoka, H., Oyama, S., Yamazaki, K., Yamazaki, A., Mitsuishi, M., Suzuki, K., 2000. GestureMan: a mobile robot that embodies a remote instructor's actions. p. 354. <https://doi.org/10.1145/358916.358986>
- Lazewatsky, D.A., Smart, W.D., 2011. A Panorama Interface for Telepresence Robots, in: *Proceedings of the 6th International Conference on Human-Robot Interaction, HRI '11*. ACM, New York, NY, USA, pp. 177–178. <https://doi.org/10.1145/1957656.1957719>
- Learn How to Control Ohmni [WWW Document], n.d. . OhmniLabs. URL <http://ohmnilabs.zendesk.com/hc/en-us/articles/115000852174> (accessed 4.27.20).
- Lee, M.K., Takayama, L., 2011. Now, I have a body: Uses and social norms for mobile remote presence in the workplace, in: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, pp. 33–42. <https://doi.org/10.1145/1978942.1978950>
- Leeb, R., Tonin, L., Rohm, M., Desideri, L., Carlson, T., Millan, J. del R., 2015. Towards Independence: A BCI Telepresence Robot for People With SevereMotor Disabilities. *PROCEEDINGS OF THE IEEE* 103, 969–982. <https://doi.org/10.1109/JPROC.2015.2419736>
- Leontiev, A.N., 1977. Activity and consciousness. *Philosophy in the USSR*, problems of dialectical materialism.
- Lister, T., 2020. Meaningful engagement via robotic telepresence: an exploratory case study. *Current Issues in Emerging eLearning* 6, 6.
- Liu, C., Ishi, C.T., Ishiguro, H., 2015. Bringing the Scene Back to the Tele-operator: Auditory Scene Manipulation for Tele-presence Systems, in: *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI '15*. ACM, New York, NY, USA, pp. 279–286. <https://doi.org/10.1145/2696454.2696494>
- Macdonald, P.S., 2000. Phenomenological factors in Vygotsky's mature psychology. *History of the human sciences* 13, 69–93.
- MantaroBot - MantaroBot TeleMe 2 -- Datasheet [WWW Document], n.d. URL <http://www.mantarobot.com/products/teleme-2/teleme-2-datasheet.htm> (accessed 6.9.20).
- Martin, R.C., 2017. *Clean architecture: a craftsman's guide to software structure and design*. Prentice Hall Press.
- Matsuda, A., Rekimoto, J., 2016. ScalableBody: A Telepresence Robot Supporting Socially Acceptable Interactions and Human Augmentation through Vertical Actuation, in: *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. ACM, Tokyo, Japan, pp. 103–105.
- McNelis, J., Schwall, G.J., Collins, J.F., 2012. Robotic remote presence technology in the surgical intensive care unit. *Journal of Trauma and*

- Acute Care Surgery 72, 527–530.
<https://doi.org/10.1097/TA.0b013e31822f7d3b>
- Mendez, I., Jong, M., Keays-White, D., Turner, G., 2013. The use of remote presence for health care delivery in a northern Inuit community: a feasibility study. *International Journal of Circumpolar Health* 72, 381–388. <https://doi.org/10.3402/ijch.v72i0.21112>
- Merleau-Ponty, M., 2013. *Phenomenology of perception*. Routledge.
- Minsky, M., 1980. *Telepresence*.
- Morales-Vidal, S., Ruland, S., 2013. Telemedicine in stroke care and rehabilitation. *Top Stroke Rehabil* 20, 101–107.
<https://doi.org/10.1310/tsr2002-101>
- Mori, M., MacDorman, K.F., Kageki, N., 2012. The uncanny valley [from the field]. *IEEE Robotics & Automation Magazine* 19, 98–100.
- Moyle, W., Jones, C., Cooke, M., O'Dwyer, S., Sung, B., Drummond, S., 2014. Connecting the person with dementia and family: A feasibility study of a telepresence robot. *BMC Geriatrics* 14. <https://doi.org/10.1186/1471-2318-14-7>
- Nakanishi, H., Murakami, Y., Nogami, D., Ishiguro, H., 2008. Minimum movement matters: impact of robot-mounted cameras on social telepresence, in: *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work*. ACM, pp. 303–312.
- Neustaedter, C., Singhal, S., Pan, R., Heshmat, Y., Forghani, A., Tang, J., 2018. From Being There to Watching: Shared and Dedicated Telepresence Robot Usage at Academic Conferences. *ACM Trans. Comput.-Hum. Interact.* 25, 33:1–33:39. <https://doi.org/10.1145/3243213>
- Neustaedter, C., Venolia, G., Procyk, J., Hawkins, D., 2016. To Beam or not to Beam: A study of remote telepresence attendance at an academic conference, in: *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*. ACM, pp. 418–431. <https://doi.org/10.1145/2818048.2819922>
- Newport, R., Pearce, R., Preston, C., 2010. Fake hands in action: embodiment and control of supernumerary limbs. *Experimental brain research* 204, 385–395.
- Ng, M.K., Primates, S., Giuliano, L., Lupetti, M.L., Russo, L.O., Farulla, G.A., Indaco, M., Rosa, S., Germak, C., Bona, B., 2015. A Cloud Robotics System for Telepresence enabling Mobility Impaired People to Enjoy the whole Museum Experience, in: *2015 10TH IEEE INTERNATIONAL CONFERENCE ON DESIGN & TECHNOLOGY OF INTEGRATED SYSTEMS IN NANOSCALE ERA (DTIS)*. IEEE, 345 E 47TH ST, NEW YORK, NY 10017 USA.
- Nielsen, C.W., Goodrich, M.A., 2006. Comparing the usefulness of video and map information in navigation tasks, in: *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction*. pp. 95–101.
- Niemelä, M., van Aershot, L., Tammela, A., Aaltonen, I., 2017. A Telepresence Robot in Residential Care: Family Increasingly Present, Personnel Worried About Privacy. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* 10652 LNAI, 85–94. https://doi.org/10.1007/978-3-319-70022-9_9

- Niemelä, M., van Aerschot, L., Tammela, A., Aaltonen, I., Lammi, H., 2019. Towards Ethical Guidelines of Using Telepresence Robots in Residential Care. *International Journal of Social Robotics*.
<https://doi.org/10.1007/s12369-019-00529-8>
- Odom, W., Zimmerman, J., Davidoff, S., Forlizzi, J., Dey, A.K., Lee, M.K., 2012. A Fieldwork of the Future with User Enactments, in: *Proceedings of the Designing Interactive Systems Conference, DIS '12*. ACM, New York, NY, USA, pp. 338–347. <https://doi.org/10.1145/2317956.2318008>
- Odom, W., Zimmerman, J., Forlizzi, J., Choi, H., Meier, S., Park, A., 2014. Unpacking the Thinking and Making Behind a User Enactments Project, in: *Proceedings of the 2014 Conference on Designing Interactive Systems, DIS '14*. ACM, New York, NY, USA, pp. 513–522. <https://doi.org/10.1145/2598510.2602960>
- O'Neill, L., Murphy, M., Gray, D., Stoner, T., 2001. An MRP System for Surgical Linen Management at a Large Hospital. *J. Med. Syst.* 25, 63–71.
- Orlandini, A., Kristoffersson, A., Almquist, L., Björkman, P., Cesta, A., Cortellessa, G., Galindo, C., Gonzalez-Jimenez, J., Gustafsson, K., Kiselev, A., 2016. Excite project: A review of forty-two months of robotic telepresence technology evolution. *Presence* 25, 204–221.
https://doi.org/10.1162/PRES_a_00262
- Paepcke, A., Soto, B., Takayama, L., Koenig, F., Gassend, B., 2011. Yelling in the Hall: Using Sidetone to Address a Problem with Mobile Remote Presence Systems, in: *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology, UIST '11*. ACM, New York, NY, USA, pp. 107–116. <https://doi.org/10.1145/2047196.2047209>
- Pang, W.-C., Wong, C.-Y., Seet, G., 2017. Exploring the Use of Robots for Museum Settings and for Learning Heritage Languages and Cultures at the Chinese Heritage Centre. *PRESENCE-TELEOPERATORS AND VIRTUAL ENVIRONMENTS* 26, 420–435.
https://doi.org/10.1162/PRES_a_00306
- Papadimitriou, C., 2008. Becoming en-wheeled: the situated accomplishment of re-embodiment as a wheelchair user after spinal cord injury. *Disability & society* 23, 691–704.
- Park, C.H., Ryu, E.-S., Howard, A.M., 2015. Telerobotic Haptic Exploration in Art Galleries and Museums for Individuals with Visual Impairments. *IEEE TRANSACTIONS ON HAPTICS* 8, 327–338.
<https://doi.org/10.1109/TOH.2015.2460253>
- Paterson, K., Hughes, B., 1999. Disability studies and phenomenology: The carnal politics of everyday life. *Disability & Society* 14, 597–610.
- Patton, M.Q., 2014. *Qualitative Research & Evaluation Methods: Integrating Theory and Practice*, Fourth edition. ed. SAGE Publications, Inc, Thousand Oaks, California.
- Paulos, E., Canny, J., 1998a. PRoP: Personal Roving Presence, in: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '98*. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, pp. 296–303. <https://doi.org/10.1145/274644.274686>
- Paulos, E., Canny, J., 1998b. Designing personal tele-embodiment, in: *Proceedings. 1998 IEEE International Conference on Robotics and Automation (Cat. No. 98CH36146)*. IEEE, pp. 3173–3178.

- Paulos, E., Canny, J., 1997. Ubiquitous tele-embodiment: Applications and implications. *International Journal of Human-Computer Studies* 46, 861–877.
- Paulos, E.J., 2001. Personal tele-embodiment. University of California, Berkeley.
- Rabardel, P., 2001. Instrument mediated activity in situations, in: *People and Computers XV—Interaction without Frontiers*. Springer, pp. 17–30.
- Rabardel, P., Bourmaud, G., 2003. From computer to instrument system: a developmental perspective. *Interacting with computers* 15, 665–691.
- Rae, I., Mutlu, B., Takayama, L., 2014. Bodies in Motion: Mobility, Presence, and Task Awareness in Telepresence, in: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '14*. ACM, New York, NY, USA, pp. 2153–2162. <https://doi.org/10.1145/2556288.2557047>
- Rae, I., Neustaedter, C., 2017. Robotic Telepresence at Scale, in: *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, pp. 313–324. <https://doi.org/10.1145/3025453.3025855>
- Rae, I., Venolia, G., Tang, J.C., Molnar, D., 2015. A framework for understanding and designing telepresence, in: *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*. ACM, pp. 1552–1566.
- Redström, J., 2017. *Making design theory*. MIT Press.
- Reynolds, J.M., 2018. Merleau-Ponty, world-creating blindness, and the phenomenology of non-normate bodies. *Merleau-Ponty, world-creating blindness, and the phenomenology of non-normate bodies* 419–436.
- Riva, G., Waterworth, J., Murray, D., 2014. *Interacting with Presence: HCI and the Sense of Presence in Computer-mediated Environments*. Walter de Gruyter GmbH & Co KG.
- Riva, G., Waterworth, J.A., Waterworth, E.L., Mantovani, F., 2011. From intention to action: The role of presence. *New Ideas in Psychology* 29, 24–37.
- Rosenberger, R., Verbeek, P.-P., 2015. A field guide to postphenomenology. *Postphenomenological investigations: Essays on human-technology relations* 9–41.
- Roussou, M., Trahanias, P., Giannoulis, G., Kamarinos, G., Argyros, A., Tsakiris, D., Georgiadis, P., Burgard, W., Haehnel, D., Cremers, A., Schulz, D., Moors, M., Spirtounias, E., Marianthi, M., Savvaides, V., Reitelman, A., Konstantios, D., Katselaki, A., 2001. Experiences from the Use of a Robotic Avatar in a Museum Setting, in: *Proceedings of the 2001 Conference on Virtual Reality, Archeology, and Cultural Heritage, VAST '01*. ACM, New York, NY, USA, pp. 153–160. <https://doi.org/10.1145/584993.585017>
- Rueben, M., Groechel, T., Zhang, Y., Matarić, M.J., Ragusa, G., 2020. Increasing Telepresence Robot Operator Awareness of Speaking Volume Appropriateness: Initial Model Development, in: *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction, HRI '20*. Association for Computing Machinery, New York, NY, USA, pp. 421–423. <https://doi.org/10.1145/3371382.3378388>
- Sacks, O., 1994. *The Man Who Mistook His Wife for a Hat*. Summit: New.

- Selic, S., 2014. Potential mobile Telepresence System for health Care in rural Areas. *Zeitschrift Fur Gerontologie Und Geriatrie* 47, 75–75.
- Seymour, W., 2012. *Remaking the body: Rehabilitation and change*. Routledge.
- Simon, H.A., 1996. *The sciences of the artificial*. MIT press.
- Smith, D.W., 2018. Phenomenology, in: Zalta, E.N. (Ed.), *The Stanford Encyclopedia of Philosophy*. Metaphysics Research Lab, Stanford University.
- Special Statement [WWW Document], n.d. URL <https://chi2017.acm.org/special-statement.html> (accessed 10.1.20).
- Spinuzzi, C., 2020. “Trying to predict the future”: third-generation activity theory’s codesign orientation. *Mind, Culture, and Activity* 27, 4–18.
- Stahl, C., Anastasiou, D., Latour, T., 2018. Social Telepresence Robots: The role of gesture for collaboration over a distance, in: *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference*. ACM, Corfu, Greece, pp. 409–414.
- Standal, Ø.F., 2011. Re-embodiment: incorporation through embodied learning of wheelchair skills. *Medicine, Health Care and Philosophy* 14, 177–184.
- Stolterman, E., Wiberg, M., 2010. Concept-Driven Interaction Design Research. *Human–Computer Interaction* 25, 95–118. <https://doi.org/10.1080/07370020903586696>
- Svanæs, D., 2013. Interaction design for and with the lived body: Some implications of merleau-ponty’s phenomenology. *ACM Transactions on Computer-Human Interaction (TOCHI)* 20, 8.
- Takayama, L., Go, J., 2012. Mixing Metaphors in Mobile Remote Presence, in: *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work, CSCW ’12*. ACM, New York, NY, USA, pp. 495–504. <https://doi.org/10.1145/2145204.2145281>
- Takayama, L., Marder-Eppstein, E., Harris, H., Beer, J.M., 2011. Assisted driving of a mobile remote presence system: System design and controlled user evaluation, in: *Robotics and Automation (ICRA), 2011 IEEE International Conference On*. IEEE, pp. 1883–1889.
- Tchounikine, P., 2017. Designing for appropriation: A theoretical account. *Human–Computer Interaction* 32, 155–195.
- Telehealth Devices & Equipment - Vita | InTouch Health [WWW Document], n.d. URL <https://intouchhealth.com/telehealth-devices/intouch-vita/?gdprorigin=true> (accessed 6.17.20).
- Telepresence robot / system for video conferencing - Endurance [WWW Document], n.d. . EnduranceRobots. URL <http://endurancerobots.com/en/robots/telepresence-robot-rig-system/> (accessed 6.17.20).
- Telepresence Robot BotEyes [WWW Document], n.d. URL <https://boteyes.com/index.html> (accessed 4.27.20).
- The Webot telepresence robot . [WWW Document], n.d. URL <https://wicon.com/en/projects/webot> (accessed 6.17.20).
- Thellman, S., Lundberg, J., Arvola, M., Ziemke, T., 2017. What Is It Like to Be a Bot? Toward More Immediate Wizard-of-Oz Control in Social Human-Robot Interaction, in: *Proceedings of the 5th International Conference on Human Agent Interaction*. pp. 435–438.

- Tokuda, Y., Hiyama, A., Miura, T., Tanikawa, T., Hirose, M., 2013. Towards Mobile Embodied 3D Avatar As Telepresence Vehicle, in: Proceedings of the 7th International Conference on Universal Access in Human-Computer Interaction: Applications and Services for Quality of Life - Volume Part III, UAHCI'13. Springer-Verlag, Berlin, Heidelberg, pp. 671–680. https://doi.org/10.1007/978-3-642-39194-1_77
- Tomasello, M., 2010. Origins of human communication. MIT press.
- Tsui, K.M., Desai, M., Yanco, H.A., 2012. Towards Measuring the Quality of Interaction: Communication Through Telepresence Robots, in: Proceedings of the Workshop on Performance Metrics for Intelligent Systems, PerMIS '12. ACM, New York, NY, USA, pp. 101–108. <https://doi.org/10.1145/2393091.2393112>
- Tsui, K.M., Desai, M., Yanco, H.A., Uhlik, C., 2011. Exploring use cases for telepresence robots, in: Proceedings of the 6th International Conference on Human-Robot Interaction. ACM, pp. 11–18. <https://doi.org/10.1145/1957656.1957664>
- Tsui, K.M., Yanco, H.A., 2013. Design challenges and guidelines for social interaction using mobile telepresence robots. *Reviews of Human Factors and Ergonomics* 9, 227–301. <https://doi.org/10.1177/1557234X13502462>
- Vaughan, J., Kratz, S., Kimber, D., 2016. Look where you're going: Visual interfaces for robot teleoperation, in: 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN). IEEE, pp. 273–280.
- Venolia, G., Tang, J., Cervantes, R., Bly, S., Robertson, G., Lee, B., Inkpen, K., 2010. Embodied Social Proxy: Mediating Interpersonal Connection in Hub-and-satellite Teams, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '10). ACM, New York, NY, USA, pp. 1049–1058. <https://doi.org/10.1145/1753326.1753482>
- VGo Solution User Guide, n.d.
- Wang, S., Lilienfeld, S.O., Rochat, P., 2015. The uncanny valley: Existence and explanations. *Review of General Psychology* 19, 393–407. <https://doi.org/10.1037/gpr0000056>
- Wasserstein, R.L., Lazar, N.A., 2016. The ASA Statement on p-Values: Context, Process, and Purpose. *The American Statistician* 70, 129–133. <https://doi.org/10.1080/00031305.2016.1154108>
- Yanco, H.A., Drury, J., 2004. “Where am I?” Acquiring situation awareness using a remote robot platform, in: 2004 IEEE International Conference on Systems, Man and Cybernetics (IEEE Cat. No. 04CH37583). IEEE, pp. 2835–2840.
- Yang, L., Jones, B., Neustaedter, C., Singhal, S., 2018. Shopping Over Distance through a Telepresence Robot.
- Yang, L., Neustaedter, C., 2020. An Autobiographical Design Study of a Long Distance Relationship: When Telepresence Robots Meet Smart Home Tools, in: Proceedings of the 2020 ACM Designing Interactive Systems Conference, DIS '20. Association for Computing Machinery, Eindhoven, Netherlands, pp. 129–140. <https://doi.org/10.1145/3357236.3395467>
- Yang, L., Neustaedter, C., 2018. Our House: Living in a Long Distance Relationship through a Telepresence Robot. *Proceedings of the ACM on*

- Human-Computer Interaction 2, Article 190.
<https://doi.org/10.1145.3274459>
- Yang, L., Neustaedter, C., Schiphorst, T., 2017. Communicating Through A Telepresence Robot: A Study of Long Distance Relationships, in: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems, CHI EA '17. ACM, New York, NY, USA, pp. 3027–3033. <https://doi.org/10.1145/3027063.3053240>
- Zhang, G., Hansen, J.P., Minakata, K., 2019. Hand- and gaze-control of telepresence robots, in: Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications. ACM, Denver, Colorado, pp. 1–8.
- Zimmerman, J., Forlizzi, J., Evenson, S., 2007. Research through design as a method for interaction design research in HCI, in: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. ACM, pp. 493–502.