HABILITATION TREATISE

‘ONE CANNOT KNOW EVERYTHING’
— ON THE NEED OF EPISTEMOLOGICAL DIVERSITY IN HUMAN-CENTERED HRI RESEARCH

ASTRID WEISS • 2021

Habilitationsschrift

zur Erlangung der Venia Docendi für das Fach Human-Computer Interaction an der Naturwissenschaftlichen Fakultät der Paris-Lodron Universität Salzburg

For my mother

Preface

Human-Robot Interaction is a rapidly and exponentially expanding field worldwide, pushed not only within the research community, but also by substantial investments from the public and the private sector. The reasons for this are rooted in the industrial demand for increased automation on the one hand, but also in the lack of service personnel and the need to prepare the current generation for the demand of demographic change. The recently established Austrian Council on Robotics and Artificial Intelligence as an advisory board of the Austrian Ministry for Climate Action, Environment, Energy, Mobility, Innovation and Technology (BMK, formerly Austrian Ministry for Transport, Innovation and Technology) is representative for the high societal relevance of this research field.

Robots and AI are getting more and more deployed into the public and private sphere and interact with people that lack formal training in automation handling – which results in the need for the design of safe, intuitive, and trustworthy robots. Although this endeavour resembles the notion that originated in Human-Computer and Human-Machine Interaction in general, the understanding of human interaction with robots requires the consideration of unique and unprecedented aspects: for the first time in the history of mankind we have to consider the interaction with machines that act autonomously, with their programming subject to learning and their actions subject to artificial intelligence. Furthermore, other than when interacting with computers, the interaction with robots means interacting with machines that can physically act upon and react to the world around them, including the human.

This unprecedented challenge can only be addressed by means of an interdisciplinary approach and requires communication and joint efforts between various disciplines. Furthermore, simply addressing the challenges ahead by means of established methods derived from either one of the related fields (sociology, psychology, engineering, computer science, etc.) or from the perspective of historically more established intersections (e.g. human factors, HCI) seems insufficient in the light of these new challenges. Therefore, a habilitation from only the perspective of a sociologist, or a computer scientist, cannot address the integrated requirements of this emerging discipline.

As a trained social scientist with extensive work experience in interdisciplinary settings, I am uniquely qualified to address these new challenges – and I have proven this with my postdoc research on human-robot interaction since 2010. In a world that is driven by the hype for technological development, my work takes a critical stance on technological feasibility, aiming to shift the field's focus towards a human-centered perspective. Over the course of my still evolving career, I have constantly been working with engineers, computer scientists, interaction designers, psychologists, and representatives from other related disciplines, towards my goal to shape a technology development process that creates interactive systems, sustainably used by people in their actual everyday environments and contexts.

I base my research on theories on human behaviour and empirical studies of human-human and human-technology interaction. My aim is to use the results of these studies to derive novel interaction paradigms and interface designs. Of particular interest to me is to study if similar processes occur when we interact with robots as compared to other technologies or humans. A central theme of my research thereby is how to design interactive technologies in order to seamlessly become part of our everyday life, which involves understanding current usage practices and domains. To me, field work and in-situ studies are the most enriching type of research (more than traditional controlled laboratory studies in which I am also well-experienced), as understanding social practices gives us an idea on what type of technology could sustainably change everyday life.

Therefore, my research is unique and stands out in the international and national community such that it is directly performed where it is contextually situated: in the public, in the factory, in domestic and care settings, always involving specific target groups that would naturally interact in said settings (e.g. older adults in care facilities, operators and maintainers in factory settings, pedestrians in public space, families in their homes etc.).

Originally trained in sociology at the University of Salzburg (2000-2005, Mag.phil), I moved on to earn a PhD (Dr.phil) in the Human-Computer Interaction (HCI) Unit at the ICT&S Center at the University of Salzburg, already focusing on the then novel field of Human-Robot Interaction (HRI) within the FP6 EU-project robot@cwe. In this project, I developed an evaluation framework that supports user-centered assessment of HRI scenarios with respect to Usability, Social Acceptance, User Experience, and Societal Impact – the USUS framework. My framework now serves as basis and guidance for fellow researchers performing user-centered HRI studies worldwide (currently 140 citations). From 2010 to 2013, I was a postdoc and interdisciplinary team leader for a research group on “Adaptive Systems” at the ICT&S Center, University of Salzburg (HCI Group). Since 2013 I am a researcher at the TU Wien. Currently, I am a senior scientist at the Human-Computer Interaction Group led by Prof. Geraldine Fitzpatrick (Institute of Visual Computing and Human-Centered Technology, TU Wien).

Since finishing my Phd degree, I have been working in two internationally and five nationally funded projects addressing various aspects of human-robot interaction, two of which I was in a principal investigator role and five leader of my institution's activities. Between 2011 and 2020 I have advised five Phd (2 graduated) and five master students (all graduated) in HRI related projects. Furthermore, I have developed an interdisciplinary introductory course to HRI and taught nine other courses on HCI related aspects. In parallel to my academic research and teaching activities, I also serve as an expert Human-Robot Interaction Consultant for Blue Danube Robotics (http://bluedanuberobotics.com) and the FFG project RoboGen.

For my research on Human-Robot Interaction (HRI) I have been awarded with the highly competitive and prestigious Hertha-Firnberg and Elise-Richter grant from the FWF Austrian Science Fund, a postdoc programme aimed at promoting the university career of young female research scientists. In total, I have received around 1.5 M. Euro in funding for my work in human-robot interaction. Honoring my accomplishments and efforts to promote Human-Robot Interaction as an independent and essential research field, I was selected as one of the “25 women you need to know in robotics” by the robohub contributor group in 2013. Acknowledging my influence in the field, I was a selected speaker for Tedx TUWien in 2017 and was elected as member of the Young Academy of the Austrian Academy of Sciences in 2018.

Since February 2019, a habilitation in Human-Computer Interaction can be achieved cumulatively at the University of Salzburg. Furthermore, it is permitted through my FWF Elise-Richter fellow grant proposal to pursue a habilitation as an external candidate at the University of Salzburg, while continuing to work at TU Wien. Therefore, in pursuit of said cumulative habilitation in HCI, I have compiled the treatise at hand which provides evidence for the my scientific excellence. It consists of 19 articles published in peer-reviewed, high- ranked, (inter)national conferences, journals, and edited books. The treatise starts with an introduction of my work, providing an overview on how my work is connected and how my publications fit in the grander theme. Then, part one describes the common theme of the treatise *'One cannot know everything' — On the Need of Epistemological Diversity in Human-centered HRI Research*. It includes sections on how my interconnected research foci are represented in the corresponding articles. Copies of said articles are then made available in the second part. References that are not my contribution, but referred to as important publications in the field, are given as footnote citations; references of my co-authored publications are listed in an additional reference section and highlighted as bold in the text, if they are part of the cumulative treatise. The appendix includes my academic CV with a list of all my publications. Along with this treatise an adjunct habilitation dossier has been compiled. It shows how this habilitation treatise fulfils requirements of the University of Salzburg and, in specific, the guidelines for habilitations in HCI. It includes a declaration of shared work and my contributions to each co-authored publication, a classification of the articles in this treatise, and a more detailed explanations of my contributions to the field.

Although Human-Robot Interaction is not an established research discipline according to current scientific standards in Austria, my work has shown that a traditional, mono-disciplinary approach will not lead to applicable results. Furthermore, my publications and accomplishments not only testify that I am uniquely qualified to further the field, and educate new researchers in the necessary inter- and transdisciplinary perspective, it also demonstrates the huge societal and industrial interest in my work. Hence, this document should be considered as reassurance that I am qualified and deserving to be rewarded with a venia docendi in HCI at the University of Salzburg.

Acknowledgements

The road to my habilitation was definitely a long and winding one accompanied by so many coincidences, people, and robots. In the first place, I never really planned or intended to study Sociology, it simply happened. And as I love thinking about life and why we live it the way we do, Sociology turned out to be really my thing. I was also definitely into technology, as I have two older brothers and my father was responsible for the computer infrastructure at the Institute of Germanistik at the University of Salzburg. So there were computers, point-and-click adventures, and the Internet early on in our home. However, there was no specific fascination with R2D2 or any other Star Wars stories in my early childhood other than it was a movie trilogy I really liked. But no superior interest in robots.

It just happened by coincidence that my PhD research ended up to be part of an EU project on Human-Robot Interaction and that I continued that path for my postdoc, because I love doing research at the intersection of humans and machines. I also never really planned to enter academia (even though my whole family are/were academics at some point of their job life), it simply happened and it turned out to be this type of job that never really felt like actually having a job.

I cherish my profession for all its flexibility, variety, and constructiveness, also when other parts of the academic life frustrate me more and more, such as the publish-or-perish culture, fixed-term working contracts, and the constant need of proofing your ‘èxcellence” through awards, grants, and citations. However, if you are part of the problem you hopefully become part of the solution at some point.

One of these excellence-proofing criteria which is still alive in Austria, Germany, and France is the habilitation. And finally I managed to compile this master piece of 10 years post-doc life as Human-Robot Interaction researcher.

At this stage, I really want to say “thank you” to so many people who accompanied this way and it worries me while writing that I will miss to mention some of you by name. First of all a huge thank you to **all my co-authors and supervised students** who are part of this compendium. It would not exist without you. From here, I will

continue chronologically by thanking my closest collaborators at the formal ICT&S Center of the University of Salzburg, where I did my first postdoc (in alphabetic order): **Ilha Aslan, Regina Bernhaupt, Roland Buchner, Verena Fuchsberger-Staufer, Hermann Huber, Patricia M. Kluckner, Susanne Meerwald-Stadler, Alexander Meschtscherjakov, Nicole Mirnig, Christiane Moser, Martin Murer, Marianna Obrist, Gerald Stollnberger, Petra Sundström, Ewald Strasser, Manfred Tscheligi, Barbara Weixelbaumer, David Wilfinger, Jakub Zlotwoski**. I am deeply grateful that I had the chance to be part of such an amazing team! - Never forget where you are coming from...

Next, my sabbatical at the Universities of Amsterdam and Twente in the Netherlands: **Vanessa Evers** - thanks for being such an impacting mentor for my academic career in only six months!

Then, my FWF Hertha-Firnberg postdoc time at the V4R group at the TU Wien (in alphabetic order): **Markus Bajones, David Fischinger, Georg Halmetschlager-Funek, Paloma de la Puente, Christiana Tsiourti, Markus Vincze, Daniel Wolf**. Academic life with all its insecurities, failure moments, doubts is not always fun, but you guys were always there to share all the ups and downs with me. We managed together to survive the Hobbit project, get over rejected papers and proposals, and spend endless hours at the lab. However, we also new how to celebrate success or anything else that somehow could be celebrated: birthdays, summer, Christmas ... what a great time!

And finally my Elise Richter postdoc time at the HCI group at TU Wien (in alphabetic order): **Geraldine Fitzpatrick, Chris Frauenberger, Glenda Hannibal, Kay Kender, Anna Pillinger, Peter Purgathofer, Margit Pohl, Isabel Schwaninger, Katta Spiel, Monika Zauner**. When I came to your group, I have entered a completely new phase in my life. Balancing motherhood and academia was one of the biggest challenges I had to face in the last four years. However, all of you always helped me not only to critically reflect on my research, but also on life goals, priorities, and mindfulness. You all gave me the flexibility, space, and time that I needed, above all during the whole COVID-19 craziness. The time at your group was on the one hand the most productive one in my postdoc career, but on the other hand also the most reflective life-learning one. I have learned so much about myself - thank you.

I am grateful for having received all the external funding, especially that of the **7th EU Framework and the Austrian Science Fund (FWF)**. Beyond that I thank **all the team members of the IURO project and the Hobbit project** for developing two robotic platforms which allowed me to explore so many facets and dimensions on

Human-Robot Interaction. Similarly, I like to express my gratitude to all of the people who participated in the user studies and field trials. You made all of this research so adventuresome, diverse, and insightful. Thank you very much!

And then there are some fellow researchers who simply during collaborations on academic services became friends over the years: **Marc Hahnheide, Tamara Lorenz, Astrid Rosenthal von der Pütten, Maha Salem, Jim Young**. I am so grateful for all the discourses and chats we had over wine, beer, and coffee, as well as zoom, skype or any other video-conferencing system. And clearly also for every time you helped me out with an emergency review or any other academic services.

However, all of this academic life would not have turned out the way it did without the support from my friends and family. Very special thanks go to my parents **Gerlinde and Andreas Weiss**. You have always believed in me, fostered my curiosity and intellect, and supported my academic career. Another huge "thank you" needs to go out to all the people who take care of my kids so wonderfully that I never had to worry about their well-being while working: **Omi, Großvati, Maia and Melitta from Krabbelgruppe Papagno, Lena, Daniela, Michi from Elisabethinum Kindergarten**.

Finally, thank you **Harald, Alma, and Jakob** for everything you are to me: My family, my home, my cheerleaders! You were such a huge support during all the countless days and nights of drafting this habilitation. I look forward to all our new adventures, now that this chapter is finally closing.

Abstract

It is argued that Human-Robot Interaction is different not only from Human-Human Interaction, but also from Human-Computer Interaction. Lay people seem to perceive autonomous robots differently than most other computing technologies. Subsequently, in the existing research landscape of HCI, the inter-and transdisciplinary research field of HRI evolved. In the last decade more and more empirical evidence was gathered on the similarities and differences of how robots are perceived and treated by lay people. However, studying how people interact with robots is based on wicked problems (i.e. problems that resist complete definition and resolution). Like other fields, such as psychology, medicine, and also HCI, HRI is facing a "replication crises" with an increasing number of unreplicated studies. Thereby, HRI research involves a unique challenge: the vast variety of robots. We need to acknowledge that every prototypical robot design is inextricably related to the problem space, the technological readiness level, the design, and the people, including the researchers involved in the development. In other words, the problem and the solution are necessarily highly interdependent, subjective, and fluid.

The goal of this habilitation is to argue for more epistemological diversity in HRI research, going beyond the dominant positivist approaches stemming from experimental psychology and robotics. The perspective is given by empirical work categorized into three different approaches of human-centered HRI research, focusing on the type of knowledge it created and what was learned about how people interact with autonomous and intelligent robots. I aim for a (partly) chronological and reflective presentation of my work, with the ultimate goal to provide a call for more epistemological diversity in HRI research, including suggestions for additional quality criteria. The gained insights shall contribute to an approach for developing robotic technology that is meaningfully and sustainably integrated into everyday practice.

This habilitation treatise comprises ten papers with a main focus on empirical contributions and nine mainly methodological and/or theoretical contributions structured according to three different approaches human-centred HRI research can undertake: (1) human-translated approach, (2) stakeholder approach, (3) co-shaping approach. A con-

clusion and outlook section seeks to answer the question, how HRI research can address the replication crisis in future and how a more relativist, interpretivist, and critical perspective may help in addition to the established, traditional positivist approach.

Much of my research presented here has focused on exploring human-centered requirements for robot development and evaluation studies of newly developed robots or existing robot platforms amended for novel use cases. Through this research I accumulated substantial domain knowledge on the requirements for robots in public space, care robots for older adults, and cobots in industry. Moreover, my research lead to a deeper understanding of how we might undertake HRI studies and develop robotic systems. For instance, I have amongst others developed methods how to study robots in the industrial context and detailed methods descriptions for ethnographic studies in the domestic context. Finally, my HRI research has amended existing robotic platforms for the industrial context and informed the conceptualisation and implementation of two novel service robots. However, my theoretical and methodological contributions addressed the bigger underlying questions of how HRI research can engage with society. My research indicates that the main reason why robots and other technologies are perceived and treated differently, is that they have stronger agency in society – an aspect that due to the ubiquity of AI and machine learning becomes more and more relevant in the HCI community in general. Therefore, I aim to generate knowledge of robots that takes into account their potential influence on society.

Contents

Chapter 1

Habilitation Treatise — Introduction

1.1 The Human in Human-Robot Interaction Research

Human-Robot Interaction (HRI) as a field of research has developed in the early 1990s and can be characterized as a field of study dedicated to the understanding, development and evaluation of robot systems for use by or with humans[1]. Researchers with diverse disciplinary backgrounds have entered this field due to different motivations: Roboticists are often interested in building robots for specific applications and want to understand what happens when these robots come into contact with humans. Psychologists are interested in understanding human reactions to agent-like robotic systems. The motivation of cognitive scientists and AI researchers is to use robot platforms as test environments for their intelligent systems[2]. As Goodrich and Schulz put it: *"The HRI problem is to understand and design the interactions between one or more humans and one or more robots"*[3].

HRI is thus a rather young and, in its origins, interdisciplinary field of research. Opinions differ as to when exactly it became established, but it can at least be stated that the first international conference on "Human-Robot Interaction" was held in 2006. Likewise, opinions differ as to whether it is still only an inter- or transdisciplinary field of research or whether it is already evolving as its own independent scientific discipline, something I have discussed already in a Late Breaking Report published at the HRI conference in 2012 **[Wei12]**.

The *"human component"* of HRI research is what distinguishes the field from other

[1]Michael A Goodrich, Alan C Schultz, et al. "Human–robot interaction: A survey". In: *Foundations and Trends in Human–Computer Interaction* 1.3 (2008), pp. 203–275.

[2]Kerstin Dautenhahn. "Human-robot interaction". In: *The Encyclopedia of Human-Computer Interaction, 2nd Ed.* (2013).

[3]Michael A Goodrich, Alan C Schultz, et al. "Human–robot interaction: A survey". In: *Foundations and Trends in Human–Computer Interaction* 1.3 (2008), pp. 203–275, p.33.

technical disciplines, such as robotics, computer vision, and AI, however, its focus is often the development of (semi-)autonomous robotic systems which should successfully interact with lay people. User studies comprise, besides others (technical, design, methodological, and theoretical) a main contribution to the research field. Similarly, contributions can be distinguished according to a robot-centered, cognition-centered and human-centered perspective[4]; the research presented in this treatise mainly relates to the latter. In a German book chapter **[Wei20]**, I give an introduction to the typical problem spaces and application contexts for human-centered HRI research, closely related to the content presented in this treatise.

1.1.1 Human-Computer vs. Human-Robot Interaction

The origin of HRI can be considered to be the study of human-machine interaction (HMI) or human factors. The goal of HMI research was to optimize machines for the usage by humans. The development from machines as passive and user-input-dependent to autonomous and pro-active systems paved the way for HRI. In the industrial context, however, roboticists for a long time only considered experts as a target group for their systems, such as robot programmers and trained maintenance personnel. The objective was to achieve efficient automated machines. The vision of co-existing intelligent service robots[5] which will be surrounded by by people in unstructured environments, as well as the narrative of the multi-functional humanoid social robot changed the envisioned target group to inexperienced users. The objective shifted to the so-called "intuitive" and "natural" interaction (a notion that is already questioned by fellow researchers[6] and myself). In order to achieve this, the first interdisciplinary research projects on human-robot interaction were set up in the early 2000s, for which various disciplines were considered relevant, including artificial intelligence, robotics, computer science, sociology, developmental psychology, art/design, and human-computer interaction.

Kiesler and Hinds[7] already noted in their *Introduction To the Special Issue of Human-Robot Interaction* in the Human- Computer Interaction Journal that *"people seem to perceive autonomous robots differently than they do most other computer technologies"*. Similarly, Dautenhahn argued that *"human-robot interaction in very different from human-*

[4]Kerstin Dautenhahn. "Socially intelligent robots: dimensions of human–robot interaction". In: *Philosophical transactions of the royal society B: Biological sciences* 362.1480 (2007), pp. 679–704.

[5]Laurel D Riek. "The social co-robotics problem space: Six key challenges". In: *Robotics Challenges and Vision (RCV2013)* (2014).

[6]Kerstin Dautenhahn. "Human-robot interaction". In: *The Encyclopedia of Human-Computer Interaction, 2nd Ed.* (2013).

[7]Sara Kiesler and Pamela Hinds. "Introduction to this special issue on human-robot interaction". In: *Human–Computer Interaction* 19.1-2 (2004), pp. 1–8, p. 3.

human interaction, human-computer interaction, and human-animal interaction"[8].

What was mainly an assumption back then, has since been substantiated with empirical evidence. As it turns out, studying how people interact with robots is based on wicked problems (i.e. problems that resist complete definition and resolution). We need to acknowledge that every prototypical robot design is inextricably related to the problem space, the technological readiness level, the design, and the people, including the researchers involved in the development. In other words, the problem and the solution are necessarily highly interdependent, subjective, and fluid. Kiesler and Hinds have already pointed towards the fact that *"designing these robots appropriately will require a deep understanding of the context of use and of the ethical and social considerations surrounding this context."*[9]. Similarly, Dautenhahn argued that the *"exploring of the design space [. . .] is likely never to be completed"*[10].

In other words, the building of a thorough understanding of how different design choices affect the complex variety of human responses towards robots, such as cognitive, affective, emotional, and relational ones, is a broader challenge than initially expected by the research community. Like other fields, such as psychology, medicine and also HCI, HRI is currently confronted by a "replication crises" with an increasing number of unreplicated experimental studies. One of the unique challenges HRI research has to face (additionally to the common ones of too small and not representative samples) is the vast variety of robots. In acknowledging the HCI-research inspired stance on HRI of Kiesler and Hinds[11], I rather consider a human-robot interaction setting as a *technology probe*[12] that offers opportunities to learn about the relationship of people, data, and technology beyond learning about the system itself. Therefore, the aim of my human-centered HRI research is not only to derive specific design features for robots from empirical user research, but to construct knowledge through critical reasoning based on the gathered data and in relation to the backdrop of existing knowledge.

HCI already embraces the potentials of *relativism*[13]. HCI scholars are heavily involved

---

[8]Kerstin Dautenhahn. "Some Brief Thoughts on the Past and Future of Human-Robot Interaction". In: *J. Hum.-Robot Interact.* 7.1 (May 2018). DOI: 10.1145/3209769. p. 2.

[9]Sara Kiesler and Pamela Hinds. "Introduction to this special issue on human-robot interaction". In: *Human–Computer Interaction* 19.1-2 (2004), pp. 1–8, p. 4.

[10]Kerstin Dautenhahn. "Robots we like to live with?!-a developmental perspective on a personalized, life-long robot companion". In: *RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No. 04TH8759).* IEEE. 2004, pp. 17–22, p. 21.

[11]Sara Kiesler and Pamela Hinds. "Introduction to this special issue on human-robot interaction". In: *Human–Computer Interaction* 19.1-2 (2004), pp. 1–8.

[12]Hilary Hutchinson, Wendy Mackay, Bo Westerlund, Benjamin B Bederson, Allison Druin, Catherine Plaisant, Michel Beaudouin-Lafon, Stéphane Conversy, Helen Evans, Heiko Hansen, et al. "Technology probes: inspiring design for and with families". In: *Proceedings of the SIGCHI conference on Human factors in computing systems.* ACM. 2003, pp. 17–24.

[13]Paul Dourish. "Implications for Design". In: *Proceedings of the SIGCHI Conference on Human Factors*

in negotiating the fragile balance between, experimental studies, ethnographic accounts, design practices, and system development[14]. The *constructivist* perspective, which views people, social systems, and technological development as interrelated also recently found its way into HRI research[15]. Additionally, perspectives from Science and Technology Studies (STS) as well as Philosophy of Science enter the stage more and more[16]. Currently (in contrast to my earlier thoughts **[Wei12]**), I would therefore describe HRI projects as primarily transdisciplinary rather than interdisciplinary, since they are usually integrative in nature and transcend disciplinary boundaries.

In the following I want to categorize my human-centered HRI research according to three different approaches on how the human was considered: (1) human-translated, (2) stakeholder involvement, (3) co-shaping. I will briefly explain the differences between these approaches and present the related publications. I will focus on the type of knowledge created through this empirical work and what was learned about how people interact with autonomous and intelligent robots. This perspective will open the stage to critically reflect on the need for more epistemological diversity in HRI research going beyond dominant positivist approaches stemming from the origins of the research field in experimental psychology and robotics.

1.1.2 Human-Translated Approach

The human-translated approach was the first one I based my human-centered HRI research on after my PhD studies. In my PhD thesis, I developed an evaluation framework for human-centered HRI, called *USUS* [WBT11]: Usability, Social Acceptance, User Experience, and Societal Impact. This evaluation framework was aligned with traditional user-centered design (UCD) research, proposing empirically validated methods suitable for studying interaction with humanoid robots, such as behavioural measures, surveys, interviews, focus groups etc. In my first postdoc project *IURO* (an FP7 funded EU-project running from 2010-2013[17]) we used the USUS framework to develop and assess an autonomous robot for outdoor public spaces. The abbreviation IURO stands for: In-

in Computing Systems. CHI '06. Montréal, Québec, Canada: ACM, 2006, pp. 541–550. DOI: 10.1145/1124772.1124855.

[14] Saul Greenberg and Bill Buxton. "Usability evaluation considered harmful (some of the time)". In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM. 2008, pp. 111–120, p. 114.

[15] Selma Šabanović, Sarah M Reeder, and Bobak Kechavarzi. "Designing robots in the wild: In situ prototype evaluation for a break management robot". In: *Journal of Human-Robot Interaction* 3.1 (2014), pp. 70–88.

[16] Johanna Seibt. "'Integrative Social Robotics'-A New Method Paradigm to Solve the Description Problem And the Regulation Problem?" In: *Robophilosophy/TRANSOR*. 2016, pp. 104–115.

[17] https://astridweiss.net/services/iuro-2010-2013/

teractive Urban Robot; the main vision of the project was to develop a robot that can independently and autonomously navigate from a starting point to a designated place in a public urban environment by asking pedestrians for guidance to find its way. In order to inform the interaction design for IURO human-human interaction served as the basis.

Such an approach is commonly called human-translated: Interaction paradigms are explored through human-human observational studies and the findings are transferred to robot behaviour. Reasons for this approach are that people are already familiar with human-human behaviour strategies[18] and can intuitively relate to it[19]. The following Table 1.1.2 gives an overview on all empirical studies, I was involved in that followed this human-human translated approach.

Study Focus	**Gained Knowledge**	**References**
contextual analysis of user requirements for a mobile navigation robot in public space based on three human-human interaction studies	development of a communication structure, navigation principles and a context model on relevant (confounding) contextual variables	[MWT11], [BWT11], **[Wei+11]**,
laboratory Wizard-of-Oz experiments to explore how multimodal feedback can help a robot to carry out itinerary requests effectively and satisfactory for a human interaction partner and facial expressions; video-based online study to assess the quality of facial expressions of two different robot heads	for itinerary requests verbal feedback is most prominent but other feedback modalities may support the conversation by providing reassurance or positive emotions; the more human-like designed robot head was rated significantly higher with respect to anthropomorphism than the head using animal-like features, but the recognition rate did not differ	[Mir+12], [Mir+15]
field trial consisting of a series of six way-finding runs (over 24 hours of run-time in total) and interactions with approximately 100 pedestrians of which 36 included a full route dialogue in the city center of Munich, Germany	data gathered through observations, interviews, and surveys provides insights into usability, user experience, and acceptance of IURO and enabled deriving recommendations for the development of other socially interactive robots for public space	**[Wei+15]**

Table 1.1: Empirical research within the *IURO* project.

Studies following a human-translated approach, are based on the idea of human-human interaction as "gold standard" of intuitiveness. For instance the study on exploring the recognition rate of facial expressions in the IURO project [Mir+15] used the recognition of

[18] Cynthia Breazeal. "Toward sociable robots". In: *Robotics and autonomous systems* 42.3-4 (2003), pp. 167–175.

[19] Yusuke Okuno, Takayuki Kanda, Michita Imai, Hiroshi Ishiguro, and Norihiro Hagita. "Providing route directions: design of robot's utterance, gesture, and timing". In: *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*. ACM. 2009, pp. 53–60.

human facial expressions as baseline. In a later study on human perception of congruent and incongruent bodily emotional expressions of the humanoid robot Pepper, we mimicked the studied bodily expressions of humans as closely as possible **[Tsi+19]**. A common feature of human-translated studies performed in a laboratory setting is that participants and the research setting are not necessarily the intended users or the environment where the robot will eventually deployed.

Thought to an extreme a human-translated approach might assume that humans are interchangeable and fungible and always show the same reaction to the same stimuli, like a Turing Machine[20]. A highly debatable assumption! However, as research on Joint Action reveals some human behavior emerges inevitably during repetitive actions, such as movement synchronisation. This phenomenon was mostly studied in undirected tasks such as when two people are rocking in chairs next to each other[21] or walk side-by-side[22]. However, it was shown that it also emerges in goal-directed movements like in pick-and-place tasks, which are common activities of daily living[23]. In an article in the international Journal on Social Robotics, I discussed and reflected with my co-authors how we might beneficially use insights from neurocognitive and neurophysical rehabilitation in the behaviour design for social companion robots in the care domain **[LWH16]**.

In general in the human-translated approach, lay people are rather seen as informants and have limited space to express their own ideas and thoughts of how to shape the human-robot interaction. As participants' identities (e.g. gender, age, professional background etc.) are not in the focus of the human-translated approach, many studies involve university populations as convenience sample, which has often been criticized for the lack of diversity[24]. In the *IURO* project the basic challenge was already the definition of the primary target audience, as the project narrative was to develop a robot for "potentially every by-passer on the street"– a reoccurring design challenge for service robots intended as tour guides, receptionist, and front desk host. In the *IURO* project, this lead to the decision that every naive person engaging with the robot is suitable for involvement in

[20] Jerome Bruner. "Culture and human development: A new look". In: *Human development* 33.6 (1990), pp. 344–355.

[21] Michael J Richardson, Kerry L Marsh, Robert W Isenhower, Justin RL Goodman, and Richard C Schmidt. "Rocking together: Dynamics of intentional and unintentional interpersonal coordination". In: *Human movement science* 26.6 (2007), pp. 867–891.

[22] Niek R van Ulzen, Claudine JC Lamoth, Andreas Daffertshofer, Gün R Semin, and Peter J Beek. "Characteristics of instructed and uninstructed interpersonal coordination while walking side-by-side". In: *Neuroscience letters* 432.2 (2008), pp. 88–93.

[23] Tamara Lorenz, Alexander Mörtl, Björn Vlaskamp, Anna Schubö, and Sandra Hirche. "Synchronization in a goal-directed task: human movement coordination with each other and robotic partners". In: *2011 RO-MAN*. IEEE. 2011, pp. 198–203.

[24] Joseph Henrich, Steven J Heine, and Ara Norenzayan. "Most people are not WEIRD". in: *Nature* 466.7302 (2010), pp. 29–29.

our studies, but neither our requirement studies nor evaluation activities did depict the necessary range of diversity among participants, instead we followed "non-probabilistic sampling" **[Wei+15, p. 50]**: In other words, whoever was available for a study was a good fit. To compensate for that, it was decided early in the project to work with personas and scenarios [ZWT11]. However, those were only informed by available pre-existing data, and subsequently shaped by researcher assumptions only.

1.1.3 Stakeholder Involvement Approach

The typical stakeholder approach in HRI research considers the identities (e.g. age, gender, cultural background etc.) of primary users and their surroundings. The first projects in which I focused on involving primary users in the development were in two very different domains: (1) human-robot collaboration for the factory context and (2) care robot development for enabling independent living at home for older adults. In both domains, it seemed obvious and needed to involve target users (factory operators and older adults) in all stages of the UCD approach. In both domains, several studies were even conducted in the actual context-of-use, so either in the industrial environment or people's private homes.

In the *factory context*[25], studies were most often proof-of-concept usability studies of novel interaction paradigms based on representative tasks performed by target users. Relevant research aspects were also of methodological nature: How do we best benchmark these interaction paradigms from a human-centered perspective and how can we perform studies with high ecological validity in the factory setting without negatively impacting the production flow? Another aspect on human-robot collaboration I explored with factory operators was how their UX and Social Acceptance of different industrial robots deployed in the production line changes over time. These studies were done through interviews, surveys and focus groups and revealed relevant aspects that stay hidden in the usability studies, such as the impact the robot has on the work rhythm and work routine and that the perception of cobots without a safety fence slightly becomes more positive, but does not reach the same positive rating as a classical "caged" robot after 1.5 years of deployment. The following Table 1.1.3 gives an overview on empirical studies, I was involved in that followed a stakeholder approach in the *factory context*.

The basic idea of the EU-funded project Hobbit[26] (2011-2015) was to build a service robot that enables older adults to stay longer in their private homes. The main goal was

[25] https://astridweiss.net/services/industry-4-0-cobots/
[26] https://astridweiss.net/services/hobbit-2011-2015/

Study Focus	Gained Knowledge	References
proof-of-concept studies on the usability and acceptance of interaction modalities for industrial robotic prototypes in the context of human-robot cooperation; all studies involved target users and were conducted in the respective factory setting	methodological insights on stakeholder involvement of industrial partners; traditional touch panels for off-the-shelf robotic systems are still limited in terms of usability and acceptance and create an intermediate layer between the operator, the robot, and the work piece	**[Buc+12]**, **[Wei+16]**, [HW17]
Studies on long-term experience of industrial robots based on surveys and interviews filled in by operators after actually working with the robot in the production line	insights on how user experience factors change over time and impact factors that might foster and/or hinder uptake	[Buc+13], [WH16]

Table 1.2: Empirical research within the *factory context*.

to provide a "feeling of safety and being supported" while maintaining or increasing older adults' feeling of self-efficacy (one's own ability to complete tasks). Consequently, the functionalists within the project team focused on emergency detection (mobile vision and external sensor systems), handling emergencies (calming dialogues, communication with relatives, etc.) as well as fall prevention measures (keeping floors clutter-free, transporting small items, searching and bringing objects, and reminders). Moreover, high usability, acceptance, as well as a reasonable level of affordability were considered as key criteria to achieve a sustainable success of the robot. Like in *IURO*, the project was based on a UCD approach, but with the aim not only to involve primary users (i.e. older adults), but also secondary users (i.e. care givers, such as friends and relatives) in the process. The following Table 1.1.3 gives an overview on empirical studies, I was involved in that followed a stakeholder approach in the *Hobbit* project.

In the Hobbit project, we wanted to start from a widely open design space that was "only" limited through the project vision of the robot. Therefore, in the requirement analysis phase, the term "robot" was intentionally avoided. To identify the expectations for the design of the robot, only the term "helper" was used in an association study procedure. This was done *"to prevent technologically colored associations"* **[Hub+14, p. 105]**. Techniques such as imagining what a "perfect helper" would look like, should avoid the challenges with all the baggage that comes with the term robot, however, it also opened up for an unrealistic design space for the project. For instance, one participant stated: "This helper will need to clean my windows and water my plants". Amending a service robot with the capability of cleaning windows and watering plants, however,

Study Focus	Gained Knowledge	References
development of a new new method for stakeholder involvement to derive social role repertoires for adaptive human-robot interaction from user requirement studies	basic role concepts and specific preference clusters of primary users suitable for the parameterizations and development of Hobbit's adaptive behaviour and social role	**[Hub+14]**
laboratory study of the first prototype of the Hobbbit robot with 49 target users in three countries to explore the technical robustness (robot-centered perspective) and how older adults react to a care robot with reciprocal behaviour and how they perceive their relationship with it (human-centered perspective)	lessons learned how the first prototype needs to be extended for the field trials; proof-of-concept that reciprocal behaviour positively influences the perceived usability and ease of learning of a care robot	[Fis+14], [Lam+14]
field trial with 18 participants (aged 75 years and older) who lived with the autonomously interacting Hobbit robot for several weeks under real-world conditions	lessons learned regarding autonomous operation and adaptive behaviour coordination (robot-centered perspective); showing that all participants interacted with Hobbit on a daily basis, rated most functions as well-working, and reported that they believe that Hobbit will be part of future elderly care; the adaptive behaviour approach eased the interaction over time (human-centered perspective)	[Baj+18], **[Baj+19]**, [Pri+16]

Table 1.3: Empirical research within the *Hobbit* project.

would be a project on its own. When the first prototype of the robot was ready for lab trials that participant was frustrated when the robot did not have any functionality close to cleaning. Therefore, in a parallel project on UCD with children, we decided in our method development (the so-called 5-Step Plan) **[LWV15]**, to work with definitions of the terms 'technology' and 'robot' to give children a basic understanding of the boundary object 'robot' that was the core concept of our educational robotics approach.

The laboratory trials with the first prototype and the field trials with the second prototype were both performed with fully autonomous robots and participants were primary and secondary target users, the adaptivity of the robot was pre-scripted: In a first step, we created a set of parameters for a purely device-like robot, which showed no proactive behavior and a very companion-like *Hobbit* which approaches people (it was a project vision to show that a more adaptive robot is more accepted by the target audience). None of this was learned, but only parameterized, e.g., after inaction of a pre-specified number of hours, the robot proactively approaches the person interacting with it suggesting some entertainment functionalities. To enable the envisioned "intuitive" interaction, however, we need stakeholders to be actively involved also in the usage of the robot. We need to offer them ways of shaping the interaction by customizing and personalizing their robot. Human-robot interaction, is not designed to its end, when we do an evaluation with target users in the field, we rather need to think of ways how to design for the use of the robot.

1.1.4 Co-shaping Approach

In the co-shaping approach, human-robot interaction is no longer understood as a dyadic interaction between one robot and one human, but as a complex socio-technical arrangement of human and non-human agents such as robots, sensors, programs, or other devices embedded in their social context, whereby social context not only relates to stakeholders, domain, and cultural back-drop, but also to views, values, intuitions and expertise of researchers. In such an approach, humans are no longer just considered as informants, but as lay experts who collaboratively engage with researchers in the shaping of the human-robot interaction. Already in 2010, Šabanović[27]stressed that in this process society (the context of use and stakeholders involved) is an active shaper rather than a "passive receptor" of robots. In a co-shaping approach we therefore have to study how people engage with robots over time. Long-term field trials in which this is actually explored are still an exception and mainly done with commercially available robots, such as in my current project *SharedSpace*, in which I explore the technology adoption process of the Anki Vector robot. The following Table 1.1.3 gives an overview on empirical studies, I was involved in that followed a co-shaping approach in the *SharedSpace*[28] project.

In future co-shaping approaches need to go beyond long-term studies of people using commercially available robots in the field. To my conviction, the goals is to develop means for co-shaping in the development of highly capable service robots. Clearly it poses significant challenges when lay experts should be involved in the active decision making of HRI development projects. As experienced in the factory context as well as in the Hobbit project, traditional stakeholder involvement in HRI research does not allow to gradually equalize the footing of researchers and participants in robot design across sessions. Therefore, participants do not develop competency in interacting with it and subsequently do not feel more confident about their ability to contribute to design.

[27]Selma Šabanović. "Robots in society, society in robots". In: *International Journal of Social Robotics* 2.4 (2010), pp. 439–450.
[28]https://astridweiss.net/services/sharedspace-2018-2022/

Study Focus	Gained Knowledge	References
research agenda for a long-term study with a commercially available robot focusing on how the relationship evolves over time	method description of the underlying sociological ethnographic approach, the types of qualitative and quantitative data gathered, and an outline of its analysis following the Domestic Robot Ecology framework[29]	**[WH18]**
content analysis of forum and Reddit threads, as well as a Facebook group comparing anthropomorphic language use of owners of Amazon Echo Show, Q.bo One, and Anki Vector to understand how they relate to these different embodiment of a voice assistant	findings suggest that the life-likeness of the artifact is not predominantly linked to the appearance, but to its interactivity and attributed agency and gender	**[Wei+20]**
analysis if being isolated at home due to COVID-19 pandemic affected peoples' engagement with Vector (as we expected people to have more time with the robot and a higher need for entertainment and companionship)	insight that even though some new use cases emerged, the engagement with Vector did not increase, supporting previous findings that commercially available robots fail to engage owners in the long run	[TPW20]

Table 1.4: Empirical research within the *SharedSpace* project.

In turn, researchers are not becoming more conscious of different perspectives on robot capabilities, usefulness, and challenges and consider their expert knowledge as superior. I will further outline my thoughts on this in section 1.2.2.

1.2 Conclusion & Outlook

Among the interests of HRI researchers lies the development of robotic systems based on the understanding and interpretation of a domain problem. The knowledge that can be constructed through human-centered HRI research varies greatly in topic, form, and range, also depending on the chosen approach. As shown above, we can learn new things in several ways and at multiple scales about different contributions:

- ***Problem domains***: Research can generate knowledge on the relationships between people, technology, and data. For instance, through my HRI research, I offered the community accumulated substantial knowledge on the requirements for robots in public space, care robots for older adults, and cobots in industry.
- ***Methods and design guidelines***: Research can also lead to a deeper understand-

ing of *how* we might undertake HRI user studies and develop robotic systems. For instance, I have developed methods how to study robots in the industrial context and detailed method descriptions for ethnographic studies in the domestic context.

- ***Robotic systems***: Research can result in system knowledge informing the technological and aesthetic decisions of future systems. My HRI research has amended existing platforms for the industrial context and informed the conceptualisation and implementation of two novel robotic systems: *IURO* and *Hobbit*.

The robot itself as well as its interaction patterns are therefore an expression of knowledge *as well as* a technological response to contextual questions. Therefore, a prototypical robot design cannot be assessed as true or false, but rather as appropriate for addressing certain questions. Every prototypical robot design is inextricably related to the specific problem context, the technological readiness level, the design, and the people, including the researchers involved in the development.

For instance the qualitative meta analysis of 18 studies using the same type of Nao robot and the standardised Godspeed questionnaire, I performed together with Bartneck, demonstrated that people's reactions to a robot are influenced by its overall appropriateness for a given problem: *"Regarding the perceived intelligence scale, the interaction scenario, the complexity of the behaviour, the role of the robot, and the realism of the interaction seem to be relevant impact factors."* **[WB15, p. 385]** In other words, robots and researchers are necessarily highly interdependent, intertwined, and undetermined.

Similarly, I reflected with my co-author Rosenthal von der Pütten **[PW15]** the empirical work of MacDorman & Entezari[30] (2015) on the uncanny valley in a comment for the Interaction Studies journal. We argue and demonstrate based on related work that judgements on the uncanny valley are affected by personality traits and pre-experiences of people, robot behaviours and task context, and the fact if we measure reactions on a reflective or implicit level.

In other words, studying how people interact with robots is based on wicked problems (i.e. problems that resist complete definition and resolution) and as Dautenhahn phrases it: *"exploring of the design space [...] is likely never to be completed"* [Dau04, p. 21]. Subsequently, HRI research is currently facing a "replication crises" with an increasing number of findings derived from experimental user studies that cannot be replicated.

[30] Karl F MacDorman and Steven O Entezari. "Individual differences predict sensitivity to the uncanny valley". In: *Interaction Studies* 16.2 (2015), pp. 141–172.

1.2.1 Replication Crisis

The HRI research community actively discusses aspects of reliability, especially as they concern user studies. This replication crisis is not just the case in our field: similar discussions can also be observed for HCI, psychology, medicine, and all fields involving people in their research. Our key academic conference "The ACM/IEEE International Conference on Human-Robot Interaction" even offered a new submission theme in 2020: *Reproducibility in Human-Robot Interaction: "This theme targets research that makes a contribution supporting the science of HRI via reproducing, replicating or re-creating prior HRI/HRI-relevant work, and artifacts for HRI research, to help our community build a strong and reliable evidence base."*[31] This focus on reproducibility is not entirely new, given the call for replication studies was already found over ten years ago in the Call for Papers for the AISB (Artificial Intelligence and Simulation of Behaviour) 2009: New Frontiers in HRI symposium: *"To develop as a field, HRI findings obtained by one research group need to be replicated by other groups. Without any additional novel insights, such work is often not publishable. Within this category, authors will have the opportunity to report on studies that confirm or disconfirm findings from experiments that have already been reported in the literature. This category includes studies that report on negative findings."*[32]

The consequences is that the academic culture of the HRI research community implicitly accepts the doctrine that human-centered HRI research follows an approach of well-controlled experimental settings with large sample sizes allowing for inferential statistics, in the believe that this helps generating knowledge from the particular to the general. From my personal experience in the role of a reviewer and Programme Committee member it can be said that this ethic is fundamental not only among authors, but also reviewers. These often cite problems with generalisability/limitations, rather than with the contribution per se, as a reason to reject a paper. Indeed, a recent analysis of ACM/IEEE HRI papers published in the years 2013-2015 found that from the 101 published full papers 96 involved an experimental study, independent of the submission theme[33].

So what can be other fruitful ways to tackle the challenges that the wicked nature of the HRI research problem poses? In the following two subsections, I will briefly outline my personal research agendas for future service robot and cobot development that in my opinion needs a stronger notion of a participatory and a social practice stance to succeed

[31] http://humanrobotinteraction.org/2020/full-papers/

[32] http://www.aisb.org.uk/convention/aisb09/Proceedings/NEWFRONTIERS

[33] Paul Baxter, James Kennedy, Emmanuel Senft, Severin Lemaignan, and Tony Belpaeme. "From characterising three years of HRI to methodology and reporting recommendations". In: *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI).* IEEE. 2016, pp. 391–398.

in developing robotic technology that will be sustainably integrated into people's everyday surroundings.

1.2.2 Service Robot Development

People's active roles in the design of social service robots are closely related to constructing ways to make them substantively meaningful[34]. Future service robot development needs a participatory approach in order for these robots to be socially constructed by diverse groups of people representing the breadth and depth of the intended target audience. In the development of the *IURO* and *Hobbit* as potentially capable co-existing social service robots, the projects focused on technical challenges, such as enabling pro-active HRI in densely populated human environments [Bus+11] or grasping unknown objects [FWV15] and therefore considered lay people mainly as informants instead of experts in their lived experience. Even though these projects tried to involve stakeholders in all phases of development (conceptualizing, building, and using the robot), an actual culture of participation was only present during the conceptualization phase of these robots. A reflection on the reasons for that and the robot-specific challenges for a participation culture in HRI research are summarized in a recently accepted journal article in AI & Societies **[WS21]**.

But how can we embed a culture of participation also in the building and using phase of robots? I envision an approach that includes the following characteristics: (1) more balancing the power of researchers and lay experts in the decision making process, (2) considering the robot as a learning non-human actor in the socio-technical arrangement, (3) considering the sociality of the robot as an outcome of the interaction, rather than an interaction paradigm designed through social cues. A first outline how I envision this approach to work is presented in the following Table 1.2.2 .

Crucial for success in developing a culture of participation will be a change in power structures in projects following a participatory approach. Besides the dominance of engineering-based robot narratives and preference for lab-based studies in the HRI research field, the community also becomes more and more aware that people evaluate and categorize robots based on how they fit within their everyday life. I am aware that robot development projects which put emphasis on lay experts involvement also in the phases of building and using robots will face numerous obstacles: from cost, time, and required

[34] Hee Rin Lee, Selma Šabanović, Wan-Ling Chang, Shinichi Nagata, Jennifer Piatt, Casey Bennett, and David Hakken. "Steps toward participatory design of social robots: mutual learning with older adults with depression". In: *Proceedings of the 2017 ACM/IEEE international conference on human-robot interaction.* 2017, pp. 244–253.

Phase	Learning Aspects	Methods
Conceptualizing	Shared context understanding: (1) Researchers: Learn about the context of use and the experiences of stakeholders; how lay experts interpret existing robots and what they like/ dislike about them; everyday life challenges of stakeholders and how robots might be used to address them (2) Lay experts: Learn about capabilities and applications of existing robots	Excursions visiting existing robots, context tours, future workshops, speculative enactments
Building	Shared interaction understanding: (1) Researchers: Learn how lay experts make sense of the robot; how robot performs basic functionalities (2) Lay experts: Learn typical error situations of robots and possible mitigation strategies and how to teach the robot new behaviors (3) Robots: Learn basic functionalities and how people fix typical error situations	Incremental testing of the robot with lay experts, development of teaching interfaces, robot learning-by-demonstration
Using	Shared functionality: (1) Researchers: Learn what lay experts want to personalize/customize; how actual usage evolves over time (2) Lay experts: Learn how to make use and sense of the robot in everyday life, (3) Robots: Learn preferences of a specific person or group of people	long-term user research, personalisation and customization of the robot by lay experts

Table 1.5: Suggestion for a participatory approach for future service robot development.

management efforts to participant attrition and ethical concerns of their privacy, from the familiar high rate of mechanical robot failures to their unforeseen effects on daily living.

1.2.3 Cobot Development

The introduction of collaborative robots into manufacturing organizations is poised to revolutionize how work is done in industrial settings and how workers interact with a robotic "co-worker". However, there is a lack of research that focuses on the socio-technical arrangements of cobots, people with different social roles in the factory, and other networked technologies and thereby answers the bigger picture question if cobots provide the promised potential as flexible and adaptive workplace enhancers. It is timely to conduct more research through a social practice lens to understand how cobots are actually being applied in the manufacturing, how different human stakeholders experience the collaboration, and how this affects the socio-technical work environment. Such an understanding will on the one hand inform the design of improved future human-cobot interaction, and on the other hand the creation of good working conditions in future Industry 4.0.

Social practice and workplace studies are rigorous analysis of actual work practices developed within the research community of Computer Supported Cooperative Work (CSCW) - originally derived from phenomenological sociology[35]. Practice-oriented work explores *"[...] historical processes and performances, longer-term actions which persist over time, and which must be studied along the full length of their temporal trajectory [,][...], situated in time and space"*.[36] The idea of applying CSCW techniques to human-robot interaction is not new, however, so far it mainly focused on studying performance in human-robot teams in the context of urban search and rescue[37]. The strength of social practice and workplace studies is its potential to provide rich and detailed descriptions of collaborative work between human and non-human entities. For example in the health-care sector *"rich ethnographies have illustrated, often in meticulous detail that collaborative clinical work involves the ordering and coordination of tasks, which requires real-time processing of local information. They have shown that clinical knowledge is often tacit, context-bound and ephemeral rather than codifiable, transferable and enduring"*[38]. It is these types of insights we are lacking for Industry 4.0 involving cobots in order to challenge our narratives and enable actual robot-supported work.

In an article submitted to Transactions in Human Machine Systems (currently under the first round of revisions) we propose a roadmap for future studies on cobots in Industry 4.0, focusing on the (1) Individual Level, (2) Team Level, and (3) Organizational Level as unit of examination targeting open research questions in the areas of (a) interaction design, (b) use and adoption, and (c) structural impact. To conduct a meaningful study on *human-robot work* in Industry 4.0, it is necessary to consider different human and non-human actors, both in dyadic and larger constellations. A cobot is never used in isolation, neither from other networked technologies nor from humans, but is an integral part of the way work is performed and perceived. In other words, human-robot work is more than an isolated individual process between a synthetic average operator 4.0[39] and

[35]Kjeld Schmidt. "The concept of 'work'in CSCW". in: *Computer Supported Cooperative Work (CSCW)* 20.4-5 (2011), pp. 341–401.

[36]Kari Kuutti and Liam J Bannon. "The turn to practice in HCI: towards a research agenda". In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 2014, pp. 3543–3552.

[37]Jill L Drury, Jean Scholtz, and Holly A Yanco. *Applying CSCW and HCI techniques to human-robot interaction*. Tech. rep. MITRE CORP BEDFORD MA, 2006.

[38]Trisha Greenhalgh, Henry WW Potts, Geoff Wong, Pippa Bark, and Deborah Swinglehurst. "Tensions and paradoxes in electronic patient record research: a systematic literature review using the meta-narrative method". In: *The Milbank Quarterly* 87.4 (2009), pp. 729–788.

[39]David Romero, Johan Stahre, Thorsten Wuest, Ovidiu Noran, Peter Bernus, Åsa Fast-Berglund, and Dominic Gorecky. "Towards an operator 4.0 typology: a human-centric perspective on the fourth industrial revolution technologies". In: *proceedings of the international conference on computers and*

the cobot, but a collective process. With regard to the potential of cobots to change work in Industry 4.0 to the better, a view only on the human-robot dyad does not suffice, it needs to be extended it to a team and organizational level as units of examination.

1.2.4 Need for New Quality Criteria

As shown through the categorization of my empirical research into three different approaches, human-centered HRI research can draw from a range of disciplinary backgrounds and benefits from the epistemological diversity of positivist, relativist, constructivist, and interpretivist approaches (among others). As the field involves a myriad of phenomena that cannot all be studied effectively through controlled empirical studies, we need to reflect this in our choice of normative quality criteria. What determines *quality* in HRI research in this comparatively young research area is continuously negotiated. Reproducibility is clearly a critical part to further evolve the field for good reason: HRI findings obtained by one research group need to be replicated by other groups to further build on them. However, singularly focusing on reliability, replicability, and representativness as the main criteria for assessing quality in HRI research fails to address research that operates outside of positivist paradigms. I propose *rigour* as a high-level criterion that emphasizes the different qualities stemming from relativist, interpretivist, and critical perspectives in addition to positivist research.

The suggestion for a set of quality criteria is based on the six criteria for rigour as developed for visualisation design studies[40]. Additionally, I added *ethical reflection* as an additional criterion. In the following, I will briefly explain the criteria and their meaning for human-centered HRI research. Further details on how methods can guide researchers in meeting these criteria during planning, conduction, and reporting their work can be found in the original paper from visualisation design studies.

- ***Informed*** - *Existing knowledge informs design and facilitates new interpretations*: An informed HRI user study carefully considers the context of a situation against a backdrop of existing knowledge and uses it as an anchoring frame or point of contrast. Knowledge abstractions in the form of design guidelines cannot inform an HRI design in all detail. For instance, a design recommendation such as: *"proactive dialogues with robots in public space [should] start with a social sub-dialogue to increase the social acceptability of the interaction."* **[Wei+15, p. 54]** can only

industrial engineering (CIE46), Tianjin, China. 2016, pp. 29–31.

[40] Miriah Meyer and Jason Dykes. "Criteria for rigor in visualization design study". In: *IEEE transactions on visualization and computer graphics* 26.1 (2019), pp. 87–97.

serve as a starting point in the development of dialogue-based interaction scenarios. We are just starting as a research field to systematically structure the knowledge we created through our own empirical work. Based on insights on failure types in human-robot interaction we developed a taxonomy on their impacts on trust and thereby identified research gaps that the community still needs to close **[Tol+20]**.

- ***Reflexive*** - *Positionality is adequately presented and reflected*: The knowledge of HRI user studies is highly embedded in the personal positionality of researchers. To account appropriately for this, we need to embrace the situatedness and multitude of the knowledges we create. We need to recognise opportunities for further investigations. As part of planning an HRI user study, we might ask questions such as: Why am I doing this specific study? What are my expectations, my underlying assumptions? What is my moral and ethical stance towards the problem under study?

- ***Abundant*** - *Complexity is accounted for by involving many different perspectives*: Abundance is highly relevant for HRI user studies, as the multidimensionality we observe in human responses towards robots calls for the respective studies to be complex, flexible, and multifaceted. HRI researchers might want to reflect on questions such as: Did I spend enough time gathering diverse data? Is the data enough to derive meaningful claims? However, the most critical question with this respect is: How much is enough? Plausible knowledge claims give fellow researchers the confidence to use them as basis for their work.

- ***Plausible*** - *Knowledge claims are evidenced, appropriate, and persuasive*: The plausibility of very particular knowledge claims: *"a participant interacted with the robot that way"* is heavily reliant on an appropriate methodological conduct, as well as an explication of the researcher's subjective perspective. Therefore, we see a strong link between plausibility and reflexivity. The plausibility of more general claims: *"this is a meaningful human-robot interaction scenario"* relies on a process of analytic generalisation, which can be achieved e.g. through replication studies, but also contrasting and synthesising of multiple studies, observations, contexts, types of data etc. (as we did it for our trust taxonomy **[Tol+20]**.

- ***Resonant*** - *The research inspires understanding and invites action*: Resonating research inspires and affects others to use the generated knowledge; it has a palpable impact on existing and future research. This can be achieved through *transferability* (i.e. identifying knowledge that might hold in different, but comparable contexts) and *evocative reports* (i.e. presenting insights in a thought provoking way that inspires a new understanding of a known phenomenon. HRI research has already tremendously benefited from *resonant* knowledge production often conducted in a mix of quantitative and qualitative approaches and will continue in future.

- ***Transparent*** - *The reporting invites scrutiny*: From a relativist stance, transparency cannot be produced through reproducibility. Instead, researchers are expected to provide a detailed description of *what* was done, *how* it was done and above all *why* it was done. Transparency particularly pertains potentially implicit decisions in the interaction design of a robot. Hence, this criterion is also heavily intertwined with reflexivity. One example is the summary publication on the *IURO* project in which we outlined all the development steps leading to the overall design of the robot. Moreover, we made design decisions explicit **[Wei+15, p. 45f]**: *"Based on the findings [...], it was decided that the IURO robot should appear as a combination of human-oriented perception cues with an anthropomorphic, but not entirely humanoid appearance. Therefore, we aimed to combine a humanoid robot head with a functionally designed body. As the robot did not need to be able to grasp or manipulate anything we decided against equipping it with an arm with a pointing hand. Thereby, we wanted to avoid wrong expectations if the robot had a hand, but would not be able to grasp with it. Instead,we used a pointer, mounted on the head of the robot, for indicating directions."*

- ***Ethical*** - *The developed design follows ethical conduct*: Much recent research has demonstrated that the development of robots touches on societal contexts that require ethical considerations, in addition to aspects of robot functionality and usability. A considerable problem is, however, that the ethical discourse (in ethics councils and governmental thinktanks) and the academic practice and development of (social) robotics platforms and interactions are rarely integrated[41]. This quality criterion goes beyond *"passing the ethical advisory board"*, though. Ethical conduct needs to permeate planning, execution, and reporting of HRI research. Procedural ethics are inherently tied to the design of intelligent systems for and with people. Ethical issues might, for example, concern funding sources or project focus. In other words, ethical conduct encompasses the fundamental stance towards the situation under study. One outcome of reflection from the use case of the Hobbit project was the Triple-A Model **[HWR16]**, which is intended as a framework for researchers

[41] Johanna Seibt. "'Integrative Social Robotics'-A New Method Paradigm to Solve the Description Problem And the Regulation Problem?" In: *Robophilosophy/TRANSOR*. 2016, pp. 104–115.

and developers to incorporate ethics in the user-and robot-centered design of social companion robots. Similarly, Seibt[42] developed the method of Integrative Social Robotics (ISR), a proposal for how to generate responsible, i.e. culturally and ethically sustainable social robotics. A hands-on example is the development of HRI use cases for the SMOOTH project using the method of ISR. Together with Fischer and colleagues we developed use cases for a social service robots in an elderly care home **[Fis+20]**. We used qualitative methods (interviews, ethnographic observation, focus groups, prototyping) in a participatory design approach and assessed our work following the five ISR principles to ensure ethical development of their social robot.

In summary, this treatise demonstrates that *HRI research benefits from epistemological diversity created through different approaches how to perform human-centered HRI research.* Our research field is concerned with multiple types of phenomena that require assessment from a range of different perspectives and approaches. Every human-centered HRI study operates in a situated problem space with the goal to *add* to the existing body of knowledge. This can be achieved through human-translated, stakeholder involvement, and co-shaping approaches, whereas every approach has its benefits and downsides. The positivist stance in the HRI research field might root back to its origins in robotics and cognitive sciences. Often roboticists "*pursued pure science such as chemistry, mathematics or physics, and then got started with engineering studies by chance*"[43]. Many perceptual, cognitive, and computational phenomena can be and should be studied effectively through controlled, empirical studies, where objectivity, repeatability, and prediction are valued and efforts are made to remove bias and errors. However, when studying people with the aim to understand their complex, contextualised, social reactions to dynamic settings in the interplay with robotic systems, a *relativist* perspective might be more suitable.

[42]Johanna Seibt. "'Integrative Social Robotics'-A New Method Paradigm to Solve the Description Problem And the Regulation Problem?" In: *Robophilosophy/TRANSOR*. 2016, pp. 104–115.

[43]EunJeong Cheon and Norman Makoto Su. "Integrating Roboticist Values into a Value Sensitive Design Framework for Humanoid Robots". In: *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*. HRI '16. Christchurch, New Zealand: IEEE Press, 2016, pp. 375–382. p. 378.

References

[Baj+19] Markus Bajones, David Fischinger, Astrid **Weiss**, Paloma De La Puente, Daniel Wolf, Markus Vincze, Tobias Körtner, Markus Weninger, Konstantinos Papoutsakis, Damien Michel, and et al. "Results of Field Trials with a Mobile Service Robot for Older Adults in 16 Private Households". In: *J. Hum.-Robot Interact.* 9.2 (Dec. 1, 2019). DOI: 10.1145/3368554.

[Baj+18] Markus Bajones, David Fischinger, Astrid **Weiss**, Daniel Wolf, Markus Vincze, Paloma de la Puente, Tobias Körtner, Markus Weninger, Konstantinos Papoutsakis, Damien Michel, et al. "Hobbit: Providing fall detection and prevention for the elderly in the real world". In: *Journal of Robotics* 2018 (Jan. 1, 2018). DOI: 10.1155/2018/1754657.

[Buc+12] Roland Buchner, Nicole Mirnig, Astrid **Weiss**, and Manfred Tscheligi. "Evaluating in real life robotic environment: Bringing together research and practice". In: *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. Sept. 1, 2012, pp. 602–607. DOI: 10.1109/ROMAN.2012.6343817.

[BWT11] Roland Buchner, Astrid **Weiss**, and Manfred Tscheligi. "Development of a context model based on video analysis". In: *2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. Mar. 1, 2011, pp. 117–118. DOI: 10.1145/1957656.1957689.

[Buc+13] Roland Buchner, Daniela Wurhofer, Astrid **Weiss**, and Manfred Tscheligi. "Robots in Time: How User Experience in Human-Robot Interaction Changes over Time". In: *ICSR2013: Proceedings of the 5th International Conference on Social Robotics*. Springer International Publishing. Jan. 1, 2013, pp. 138–147.

[Bus+11] Martin Buss, Daniel Carton, Barbara Gonsior, Kolja Kuehnlenz, Christian Landsiedel, Nikos Mitsou, Roderick de Nijs, Jakub Zlotowski, Stefan Sosnowski, Ewald Strasser, et al. "Towards proactive human-robot interaction in human environments". In: *2011 2nd International Conference on Cognitive Infocommunications (CogInfoCom)*. IEEE. 2011, pp. 1–6.

[Fis+20] Kerstin Fischer, Johanna Seibt, Raffaele Rodogno, Maike Kirkegård Rasmussen, Astrid **Weiss**, Leon Bodenhagen, William Kristian Juel, and Norbert Krüger. "Integrative Social Robotics Hands-on". In: *Interaction Studies* 21.1 (Jan. 1, 2020), pp. 145–185. ISSN: 1572-0373. DOI: https://doi.org/10.1075/is.18058.fis.

[Fis+14] David Fischinger, Peter Einramhof, Konstantinos Papoutsakis, Walter Wohlkinger, Peter Mayer, Paul Panek, Stefan Hofmann, Tobias Koertner, Astrid **Weiss**, Antonis Argyros, and et al. "Hobbit, a Care Robot Supporting Independent Living at Home". In: *Robot. Auton. Syst.* 75.PA (June 1, 2014), pp. 60–78. ISSN: 0921-8890. DOI: 10.1016/j.robot.2014.09.029.

[FWV15] David Fischinger, Astrid **Weiss**, and Markus Vincze. "Learning grasps with topographic features". In: *The International Journal of Robotics Research* 34.9 (Jan. 1, 2015), pp. 1167–1194. DOI: 10.1177/0278364915577105.

[Hub+14] Andreas Huber, Lara Lammer, Astrid **Weiss**, and Markus Vincze. "Designing Adaptive Roles for Socially Assistive Robots: A New Method to Reduce Technological Determinism and Role Stereotypes". In: *J. Hum.-Robot Interact.* 3.2 (Jan. 1, 2014), pp. 100–115. DOI: 10.5898/JHRI.3.2.Huber.

[HW17] Andreas Huber and Astrid **Weiss**. "Developing Human-Robot Interaction for an Industry 4.0 Robot: How Industry Workers Helped to Improve Remote-HRI to Physical-HRI". In: *Proceedings of the Companion of the 2017 ACM/IEEE International Conference on Human-Robot Interaction*. HRI '17. Vienna, Austria: Association for Computing Machinery, Jan. 1, 2017, pp. 137–138. DOI: 10.1145/3029798.3038346.

[HWR16] Andreas Huber, Astrid **Weiss**, and Marjo Rauhala. "The ethical risk of attachment how to identify, investigate and predict potential ethical risks in the development of social companion robots". In: *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. Mar. 1, 2016, pp. 367–374. DOI: 10.1109/HRI.2016.7451774.

[Lam+14] Lara Lammer, Andreas Huber, Astrid **Weiss**, and Markus Vincze. "Mutual Care: How older adults react when they should help their care robot". In: *AISB2014: Proceedings of the 3rd international symposium on New Frontiers in Human-Robot interaction*. Jan. 1, 2014.

[LWV15] Lara Lammer, Astrid **Weiss**, and Markus Vincze. "The 5-Step Plan". In: *Human-Computer Interaction – INTERACT 2015*. Ed. by Julio Abascal, Simone Barbosa, Mirko Fetter, Tom Gross, Philippe Palanque, and Marco Winckler. Cham: Springer International Publishing, Jan. 1, 2015, pp. 557–564.

[LWH16] Tamara Lorenz, Astrid **Weiss**, and Sandra Hirche. "Synchrony and Reciprocity: Key Mechanisms for Social Companion Robots in Therapy and Care". In: *International Journal of Social Robotics* 8.1 (Jan. 1, 2016), pp. 125–143. ISSN: 1875-4805. DOI: 10.1007/s12369-015-0325-8.

[Mir+12] Nicole Mirnig, Barbara Gonsior, Stefan Sosnowski, Christian Landsiedel, Dirk Wollherr, Astrid **Weiss**, and Manfred Tscheligi. "Feedback guidelines for multimodal human-robot interaction: How should a robot give feedback when asking for directions?" In: *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. Sept. 1, 2012, pp. 533–538. DOI: 10.1109/ROMAN.2012.6343806.

[Mir+15] Nicole Mirnig, Ewald Strasser, Astrid **Weiss**, Barbara Kühnlenz, Dirk Wollherr, and Manfred Tscheligi. "Can you read my face?" In: *International Journal of Social Robotics* 7.1 (Jan. 1, 2015), pp. 63–76. DOI: 10.1007/s12369-014-0261-z.

[MWT11] Nicole Mirnig, Astrid **Weiss**, and Manfred Tscheligi. "A communication structure for human-robot itinerary requests". In: *2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. Mar. 1, 2011, pp. 205–206. DOI: 10.1145/1957656.1957733.

[Pri+16] Jürgen Pripfl, Tobias Körtner, Daliah Batko-Klein, Denise Hebesberger, Markus Weninger, Christoph Gisinger, Susanne Frennert, Hakan Eftring, Margarita Antona, Ilia Adami, Astrid **Weiss**, Markus Bajones, and Markus Vincze. "Results of a Real World Trial with a Mobile Social Service Robot for Older Adults". In: *The Eleventh ACM/IEEE International Conference on Human Robot Interaction*. HRI '16. Christchurch, New Zealand: IEEE Press, 2016, pp. 497–498.

[PW15] Astrid Rosenthal-von der Pütten and Astrid **Weiss**. "The uncanny valley phenomenon does it affect all of us?" In: *Interaction Studies* 16.2 (Jan. 1, 2015), pp. 206–214. ISSN: 1572-0373. DOI: https://doi.org/10.1075/is.16.2.07ros.

[Tol+20] Suzanne Tolmeijer, Astrid **Weiss**, Marc Hanheide, Felix Lindner, Thomas M Powers, Clare Dixon, and Myrthe L Tielman. "Taxonomy of Trust-Relevant Failures and Mitigation Strategies". In: *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*. HRI '20. Cambridge, United Kingdom: Association for Computing Machinery, Jan. 1, 2020, pp. 3–12. DOI: 10.1145/3319502.3374793.

[TPW20] Christiana Tsiourti, Anna Pillinger, and Astrid **Weiss**. "Was Vector a Companion during Shutdown? Insights from an Ethnographic Study in Austria". In: *Proceedings of the 8th International Conference on Human-Agent Interaction*. HAI '20. Virtual Event, USA: Association for Computing Machinery, 2020, pp. 269–271. DOI: 10.1145/3406499.3418767.

[Tsi+19] Christiana Tsiourti, Astrid **Weiss**, Katarzyna Wac, and Markus Vincze. "Multimodal Integration of Emotional Signals from Voice, Body, and Context: Effects of (In)Congruence on Emotion Recognition and Attitudes Towards Robots". In: *International Journal of Social Robotics* 11.4 (Aug. 1, 2019), pp. 555–573. ISSN: 1875-4805. DOI: 10.1007/s12369-019-00524-z.

[Wei12] Astrid **Weiss**. "HRI research: The interdisciplinary challenge or the dawning of the discipline?" In: *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. Mar. 1, 2012, pp. 271–271. DOI: 10.1145/2157689.2157789.

[Wei20] Astrid **Weiss**. "Wenn Menschen und Roboter in Interaktion treten — Beispielhafte Problemstellungen und Anwendungsgebiete eines transdisziplinären Forschungsfeldes". In: *Mensch und Maschine*. Blätter für Technikgeschichte (BfTG), 2020, pp. 31–56.

[WB15] Astrid **Weiss** and Christoph Bartneck. "Meta analysis of the usage of the Godspeed Questionnaire Series". In: *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. Aug. 1, 2015, pp. 381–388. DOI: 10.1109/ROMAN.2015.7333568.

[WBT11] Astrid **Weiss**, Regina Bernhaupt, and Manfred Tscheligi. "The USUS evaluation framework for user-centered HRI". In: *New Frontiers in Human–Robot Interaction*. John Benjamins, Jan. 1, 2011, pp. 89–110.

[WH18] Astrid **Weiss** and Glenda Hannibal. "What Makes People Accept or Reject Companion Robots? A Research Agenda". In: *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference*. PETRA '18.

Corfu, Greece: Association for Computing Machinery, Jan. 1, 2018, pp. 397–404. DOI: 10.1145/3197768.3203177.

[WH16] Astrid **Weiss** and Andreas Huber. "User Experience of a Smart Factory Robot: Assembly Line Workers Demand Adaptive Robots". In: vol. abs/1606.03846. Jan. 1, 2016.

[Wei+16] Astrid **Weiss**, Andreas Huber, Jürgen Minichberger, and Markus Ikeda. "First Application of Robot Teaching in an Existing Industry 4.0 Environment: Does It Really Work?" In: *Societies* 6.3 (Jan. 1, 2016), p. 20.

[Wei+15] Astrid **Weiss**, Nicole Mirnig, Ulrike Bruckenberger, Ewald Strasser, Manfred Tscheligi, Barbara Kühnlenz, Dirk Wollherr, and Bartlomiej Stanczyk. "The interactive urban robot: user-centered development and final field trial of a direction requesting robot". In: *Paladyn, Journal of Behavioral Robotics* 6.1 (Jan. 1, 2015). DOI: https://doi.org/10.1515/pjbr-2015-0005.

[Wei+11] Astrid **Weiss**, Nicole Mirnig, Roland Buchner, Florian Förster, and Manfred Tscheligi. "Transferring human-human interaction studies to HRI scenarios in public space". In: *INTERACT2011: Proceedings of the 13th IFIP TC13 Conference on Human-Computer Interaction*. Springer Berlin Heidelberg. Jan. 1, 2011, pp. 230–247.

[Wei+20] Astrid **Weiss**, Anna Pillinger, Katta Spiel, and Sabine Zauchner-Studnicka. "Inconsequential Appearances: An Analysis of Anthropomorphic Language in Voice Assistant Forums". In: *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*. CHI EA '20. Honolulu, HI, USA: Association for Computing Machinery, Jan. 1, 2020, pp. 1–7. DOI: 10.1145/3334480.3382793.

[WS21] Astrid **Weiss** and Katta Spiel. "Robots beyond Science Fiction: Mutual Learning in Human-Robot Interaction on the Way to Participatory Approaches". In: *Accepted for Publication in the Special Issue on Critical Robotics in AI & SOCIETY* (2021).

[ZWT11] Jakub Złotowski, Astrid **Weiss**, and Manfred Tscheligi. "Interaction scenarios for HRI in public space". In: *ICSR2011: Proceedings of the 3rd International Conference on Social Robotics*. Springer Berlin Heidelberg. Jan. 1, 2011, pp. 1–10.