Does a Robot Lawnmower Know Anything About Grass?

A Phenomenological Gaze on Human–Robot Interactions

Thesis submitted for the degree of Philosophiae Doctor

Department of Informatics
The Faculty of Mathematics and Natural Sciences
University of Oslo

By Rebekka Soma.

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 2773

ISSN 1501-7710

Cover: UiO.
Print production: Graphic center, University of Oslo.

Abstract

This thesis contributes four theoretically elucidated concepts to the field of Human–Robot Interaction (HRI). The topic of this thesis originates from a genuine sense of wonder at an interactional phenomenon, the experience of perceiving robotic artefacts as animate or agentive, or even experiencing a certain sociality by their presence. The thesis focuses specifically on understanding aspects of mutual intelligibility between humans and robotic artefacts during interaction in light of this phenomenon. Thus, this work demonstrates what it means to hold what I call a *phenomenological gaze* on human–robot interactions. It does so by widening the explorative space of this study, focusing on *what* and *how* of an experience, rather than why.

I examine different aspects of this topic through three theories. First, to address what might be described as a mismatch between the "apparent" and "actual" capabilities of robotic artefacts, I apply biologist Jakob von Uexküll's Umwelt theory. Through this examination, I outline artificial Umwelt, an analytical tool that can be used to determine how well the design of a robotic artefact currently fits its operational space. Next, I examine the structure of ecological psychologist James J. Gibson's affordances and compare them to Donald Norman's design-specific affordances. Rather than being there to facilitate communication between designers and users, affordances support users in making meaningful evaluations during interactions with computational and robotic artefacts. Lastly, I elucidate dancer and philosopher Maxine Sheets-Johnstone's phenomenology of movement. Through her concept of animate form, she demonstrates that all living beings think primarily in movement, and that movement itself constitutes both experience and consciousness. Using concepts from her phenomenology, I explore the intertwined nature of animacy, agency, and intentionality in light of kinesthetic intercorporeality, which I characterise as a presence-based sociality.

Finally, I discuss the concepts introduced above (*artificial Umwelt*, *kinesthetic intercorporeality*) in combination with other concepts introduced in the four articles included in this thesis (*datanomy*, *movement acts*). At the end of this discussion, I outline *kinetic affordances* as a way of addressing movement itself as the meaningful relation between humans and robotic artefacts. These concepts all elucidate aspects of the intelligibility between humans and robotic artefacts as interactional partners.

Acknowledgements

As a child, I knew at least two things. I was going to be a scientist (preferably an astrophysicist), and I strongly disliked being in a sauna. Though I as a child did not understand what being a scientist entailed, I understood that a scientist found out new things, uncovering currently unknown facts about the world (or universe). Naturally, I have since found out that discovering new things is a demanding task, but with the support of others, it becomes a little easier.

I therefore wish to begin by expressing my gratitude toward the University of Oslo, and the Department of Informatics for offering the study program Informatics: Design, Use and Interaction. I would've never thought so at the beginning of it, but I now look back on over ten years at this institution. I am grateful for having had the opportunity to spend my time here. To those responsible for the content of the education I received, thank you for showing me that you can be a scientist without necessarily uncovering hard facts. On that note, I wish to extend a special thanks to, Martine Eklund, for demonstrating to me all those years ago what it means to be a student, and not just enrolled in a study program. I am honestly worried I might have changed my study program if it wasn't for you. Through our student projects and report writing, I discovered that I enjoy expressing myself with precision under your guidance. And of course, our friendship is one of those I hold dear.

When I embarked on the long journey that has been my PhD, I was given the freedom to do whatever I wanted, a freedom which has proven both terrifying and confusing, but all the more rewarding. Thank you, Tone Bratteteig, for entrusting me with this freedom. I express my deepest gratitude for your support and encouragement. Without it, I would have never gotten through this. Thank you for seeing in me what I still can't quite see in myself. I would further like to thank my co-supervisor, Jo Herstad, for countless interesting discussions, intermittent reminders of the moon-phases, and for giving me enough book recommendations to keep me busy for the rest of my life. Not least, thank you for encouraging me to pursue my research interests.

I would also like to thank all of my colleagues in the Design of Information Systems (DESIGN group): This includes both our permanent staff and everyone I have shared the

PhD-existence with. Thank you all for being great colleagues. I further would like to extend thanks to everyone I met and participated with in the MECS project. Not least must I thank Jim Tørresen, head of the Robotics and Intelligent Systems group (ROBIN) at the Department of Informatics, for welcoming me into the MECS-project activities.

I wish to thank Guri Birgitte Verne for always being available for methodological consultation, and for sharing your auto-ethnographical perspectives on Roberto with me. Additionally, I wish to express my gratitude to both Guri and Ivar Frønes for the feedback offered at my 3rd semester upgrade. I must also commend Astrid Twenebowa Ansa-Quarshie Larssen for throwing yourself around and finding time to proof this manuscript on very short notice. To Svein Anton Hovde, I would like to express my heartfelt gratitude for sharing your academic perspectives, which have also been invaluable to me. So has your warm and inclusive nature, and for being someone to share my appreciation of cats with. To Diana Saplacan, Trenton Schulz, and Alice Franz Schneider, who began at the same time as me, thank you for all the knowledge and inspiration you have provided. A special thank you to Trenton for being patient with me during paper collaborations. Being your colleague has been truly inspiring. Finally, we have all made it through!

Johanne Svanes Oskarsen, I am grateful for our friendship, and not least for our shared sense of fate. Thank you for tagging along to the gym, for all our hikes to Vettakollen and for the countless hours of Zoom-support (emotional, not technical).

To round off, I must naturally thank all my friends and family for your support, and dedicated curiosity toward the state of things, even when I most of the time hurried to change the subject. Thank you, Harald, for always being my safe harbour.

To Joakim, whom I cannot live without. We have enjoyed all phases of our higher education together, albeit in different fields. Thank you for all the joy, comfort, and security you bring to my life. Nurtured by your love, I have been able to unfold in ways I never thought possible. Being with you has led me to discover new and surprising things about myself, such as my love for exercise, walking, and that sauna is so much more than just a warm room. Finally, I must of course thank my lovely Sylvester. They say that there is never a good time to have children, but your coming into the world could never have happened at a better time. While it has definitely been challenging to not be fully in charge of my own schedule in the final months of writing, it has been my true privilege to witness you discover yourself in movement—a discovery of yourself through movement that there is for me no doubt about starting already in the womb. Now, I enjoy every moment I get to spend with you, observing with fascination as you figure out your corporeal powers. I am grateful for getting to spend my time with you both.

Table of contents

List of papers

Paper 1: Turning Away from an Anthropocentric View on Robotics

Soma, R., & Herstad, J. (2018). Turning Away from an Anthropocentric View on Robotics. Frontiers in Artificial Intelligence and Applications, 53–62. https://doi.org/10.3233/978-1-61499-931-7-53

Paper 2: Engaging in Deep Wonder at the Experience of Encountering a Lawnmower Robot

Soma, R. (2020). Engaging in Deep Wonder at the Experience of Encountering a Lawnmower Robot. In M. Nørskov, J. Seibt, & O. S. Quick (Eds.), Frontiers in Artificial Intelligence and Applications. IOS Press. https://doi.org/10.3233/FAIA200905

Paper 3: Movement Acts in Breakdown Situations: How a Robot's Recovery Procedure Affects Participants' Opinions

Schulz, T., Soma, R., & Holthaus, P. (2021). Movement Acts in Breakdown Situations: How a Robot's Recovery Procedure Affects Participants' Opinions. Paladyn, Journal of Behavioral Robotics, 12(1), 336–355. https://doi.org/10.1515/pjbr-2021-0027

Paper 4: Strengthening Human Autonomy. In the Era of Autonomous Technology

Soma, Rebekka; Bratteteig, Tone; Saplacan, Diana; Schimmer, Robyn; Campano, Erik; and Verne, Guri B. (2022) "Strengthening Human Autonomy. In the era of autonomous technology," Scandinavian Journal of Information Systems: Vol. 34: Iss. 2, Article 5.

Preface

Ever since childhood, a feeling of wonder at *what it would be like to be someone else* has persisted in my thoughts. My first concrete memory of this question occurring was probably when I was around ten years old. In my earliest memory, my curiosity was turned toward a particular friend. I cannot answer why she in particular. Perhaps it was circumstantial, and she was the person I was around when this question started appearing, or that I started asking this question because she, in particular, intrigued me. This I will never know.

Though it might appear so at first glance, I didn't as much have a wish to be someone else. Rather, I had what I would describe as a starving curiosity of wanting to know *what it would be like to be someone else*. The main reason I describe it as a 'starving sensation' was because I also always knew that I could never know: Because I didn't wonder what it would be like to be "me" inside "the body" of someone else. I wanted to know how the experiences of others differed from my own, and what the world was like from *another point of view*. In the case of my friend, I knew I could never experience *being* her through her, but I could to an extent imagine some aspects—even though I was not interested in my imaginings as they could never do justice to the realness of true experience. I wondered not merely *what* she was thinking, but *how*—how did she view the world, how did she experience it? I wondered *what it would be like to be her*. I remember wanting to experience the world through another's first-person view, out of sheer curiosity of what it would mean not to be me but someone else. Not because I didn't like being me, but more because I knew my perspective lacked an understanding of what the world was like for others.

Later, I would for periods nearly daily have these reflections and imaginings when interacting with my cat. In the case of the cat, the curiosity was even stronger, but also more frustrating. I didn't just wonder "what is he thinking?" but "what *is it* to be him?" My thoughts on the matter can be better articulated as a "conglomerate" around questions such as "what is it like to be born a cat and have a cat's life? How do you understand the world if you have only experienced it as a cat? What is it like to walk on four legs? What does the world look like when you are cat-sized? What is it like to not understand language but also not care about it because you are a cat? What is it like to have a tail, see in the dark, be covered in fur and have all fuzzy ears?"

My curiosity would of course never be satiated; my imaginings could never give a proper account of his true experience of being a cat. Not just any cat, but this particular cat, which I know, would have a different way of being than all other cats. "What does this cat think of me?", "how does this cat experience his life in our family's care? What does he feel when we are together, and how does it differ from my experience of being with him? What does this cat do when it is alone?" While these thoughts no longer occur on a daily basis, this way of thinking has stuck with me, and is the basis for my curiosity of being a living thing in the world, and of all the experiences that can be had as such.

In writing this, I have taken a step back to reflect on a recurring experience. It has been an exercise in articulating something I have never before articulated. It has also been an exercise in articulating the particular *feeling* of these thoughts. I find the curiosity I am describing in this articulation, different from the two questions "what is the other thinking" and "I wish I were someone else." Instead, it is an acknowledgement that no one can only truly know what another is experiencing in their particular way of being. I know I will never get answers to questions such as these, at least not in any way which could satisfy my curiosity. Still, this way of wondering never quite stopped.

1 Introduction

> When we let ourselves begin wondering […], we give in to our longings to understand something about human life, something we do not understand, something that begs us to pay attention to it, something with which we feel we have to come to terms.
> — *Maxine Sheets-Johnstone, The primacy of Movement*

1.1 THIS STUDY AND ITS MOTIVATION

Over the course of my formal education in interaction design, especially as I underwent the journey of writing my Master's thesis, I became increasingly interested in questions having to do with the fundamentals of interactions. Particularly one approach to designing interactive technology, called *embodied interaction* caught my attention. Paul Dourish's (Dourish 2001) detailed explanation of its phenomenological underpinnings and why we should care about such philosophical foundations resonated strongly with me. The ideas on which embodied interaction is built, Dourish (2001) explains, "exploit our familiarity and facility with the everyday world—whether it is a world of social interaction or physical artefacts" (p. 17). I was fascinated.

When reading these explanations of how computing moved from "inside screens" and "into the real world." It made me wonder. This kind of dichotomous separation, between the "physical" and "digital" world, puzzled me. Weren't screens, after all, a part of the real world? In turn, it sparked a new interest, concerned with the importance of precise terminology—and the realisation of how difficult this is to achieve. As I set out on the journey that has become this study, I did so with a phenomenological outlook inspired by what Dourish (2001) had taught me.

As my project was in its infancy, another, larger research project, MECS (Multimodal Elderly Care Systems), was starting up in the department, and I was invited to partake in their research activities. MECS (2016-2020)[1] was a collaboration between the

Department of Informatics research group for Robotics and Intelligent Systems (ROBIN) and the research group for the Design of Information Systems (DESIGN) where I was a member. Its overarching purpose was exploring how new technological solutions might ease the current load on the welfare state, a result of the increasing number of retired citizens compared to the number of working citizens—vernacularly referred to as "the senior wave". Modern technology can contribute to lightening this burden if we find the right ways to use it to support older adults in living independently in their own homes. To this end, the MECS project's research focus has been on performance, privacy and enhancing fall prediction. Even if the intention is to facilitate independent living for older adults, the thought of installing sensors for monitoring health inside the very walls of their home will for many be very invasive.

Therefore, MECS set out to explore if robots could not only work as a way of collecting health data but also be a physical representation of the health data collected. By giving the data-collecting technology a physical manifestation, the users of this welfare technology would remain more autonomous in their encounters and interactions with this technology. For instance, a robot can be sent away to a different area of the home (Schulz, Torresen, and Herstad 2018). In addition, it was suggested that a robot—as opposed to sensors installed in walls—could also serve the role of a *companion* for the older adults who would be living with this technology in their homes. This specific suggestion caught my attention. Accordingly, the earliest conception of my project had *robot companionship* as its main focus. As I began to familiarise myself with the Human–Robot Interaction field, I was increasingly puzzled by diverging terminology. I found it challenging to navigate the myriad of categories: social robots, sociable robots, companion robots, therapeutic robots, toy robots, household robots, and so on. That categories overlap is not surprising in itself. Nevertheless, it was not clear to me where, or with whom, the power sits to define what kind of robot a specific robotic artefact is.

When the other researchers and I were meeting with the older adults at Kampen Omsorg+ they seemed not only uninterested in, but directly opposed to a robot companion. However, when describing the robot vacuum cleaner they were borrowing from MECS, they were not shy of describing it using a "sociable" vocabulary. I began *wondering*. I wondered about the nature of companionship, the nature of robot technology, and how these could go together. Was a companion robot a companion whenever users felt companionship, or should it be designed and labelled specifically for achieving this purpose in use? Was a companion robot first and foremost a social robot, or was it first and foremost a companion robot? Are these categories necessary if the user feels no companionship with a companion robot, but does so with a household robot? It was my understanding that these were timely questions within the research field Human–Robot Interaction (HRI), Social Robotics (SR), as well as other related fields dealing with questions about robots and society.

Eventually, it was the movements of robotic artefacts that captured my unparalleled interest. Especially those designed with only a functional purpose, yet somehow eliciting social reactions from those interacting with the robot. It intrigued me that robots seemed to be attributed social characteristics, even when the robots were not intended to socially engage their users. The central issue of this thesis concerns what I have come to refer to as a *dissonance between the apparent capabilities of robotic artefacts and their actual capabilities during interaction*.

While the study described in this thesis cannot be called a phenomenological study, it sprung out of a curiosity—a genuine sense of wonder—about a human experiential phenomenon (van Manen 2014). Therefore, I present it as a study that maintains a *phenomenological gaze* on human–robot interaction. Rather than a case study of robotic artefacts, it is better understood as an *intrinsic case* study (Myers 1997; Stake 2005; Walsham 2006) of a specific kind of interaction that takes place between humans and robotic artefacts. The study is thus relevant beyond the superficial considerations of the specific robotic artefacts occasionally discussed, such as robot lawnmowers or robot vacuum cleaners. The work and contributions of this thesis is almost exclusively theoretical, yet I see the matters discussed in this thesis as highly relevant to those aiming to create successful interactions between humans and robots.

1.2 THE FIELD OF HUMAN–ROBOT INTERACTION

Human–Robot Interaction (HRI) is a multidisciplinary field considered to have emerged during the mid-1990s and early 2000s. This field is "dedicated to understanding, designing, and evaluating robotic systems for use by or with humans" (Goodrich & Schultz 2007, p. 204), indicating that research and application are both of equal significance. The interdisciplinarity of HRI spans not only purely technical domains within robotics and artificial intelligence, but also intersects with research on human factors, such as psychology, cognitive science, and linguistics (Dautenhahn 2007; Goodrich and Schultz 2007). In other words, researchers of HRI, need to understand their research within a broad context.

HRI, including the domain of Social Robotics (SR), has been criticised for its scattered methodologies (Veling and McGinn 2021). Its relatively young age is often considered a culprit, as "the approaches, standards and methods are still in the process of negotiation" (Veling and McGinn 2021, p. 1689). Others have pointed to the difficulty in defining common metrics as "the incredibly diverse range of human-robot applications" (Steinfeld et al. 2006, p. 33).

In the extension of this, one of the main challenges faced by HRI is the transient nature of its very foundation: The concept of *robot* is a "moving target" (Dautenhahn 2014). As such, there is no straightforward way of rigidly categorising the various types of robotic artefacts that exist. Some have attempted to classify robots according to what application area they are designed for (e.g., Thrun 2004). Others hold that the way a person perceives the role of the robot they collaborate with has "important ramifications for how they interact with [it]" (Goodrich and Schultz 2007, p. 234).

Therefore, Goodrich and Schultz (2007) build on the taxonomy put forth by Scholtz (2003) categorising robots according to the role they take in a collaborative interaction (*supervisor*, *operator*, *mechanic*, *peer*, *bystander, mentor* and *information consumer*). Thus, Goodrich and Schultz's (2007) categorisation has an inherent view of sociality: It builds on the idea that it would be beneficial for interaction if "the human and the robot engage in dialogue to exchange ideas, to ask questions, and to resolve differences" (Fong, Thorpe, and Baur 2001, p. 1).

The overall research interest of HRI is well summarised by Veling and McGinn (2021) as being "concerned with understanding, designing, and evaluating robots for use by, or with, humans, often in uncontrolled, or 'real-world' settings" (p.1689). Dautenhahn explains that studies on people's attitudes toward robots and their interactions with them "typically entail large-scale evaluations trying to find statistically significant results" (2014). Further, she remarks that it is unfortunate should HRI end up being equated with such a narrow niche of studies. This is in line with the observations made by Veling and McGinn (2021), who point toward "a tendency in HRI to aim for precision, characterized by the use of quantitative metrics and clearly defined hypotheses" (p. 1690). Quantitative studies can provide knowledge concerning general trends within a large sample size, studying a phenomenon independent of context. Still, as Veling and McGinn (2021) demonstrate, there are many outstanding qualitative HRI studies. The use of qualitative methods is more suited when one is interested in understanding "social contexts, human perspectives, the nature of interaction, and to generate new understandings and explanations" (p. 1689). The social aspects of human–robot interactions will be explicated in Chapter 2.

As Veling and McGinn (2021) emphasise, the HRI field would likely benefit from "a foundation of complementary approaches and methodologies" (p. 1689). In a similar vein, Sheets-Johnstone (2011) states that while phenomenology is by no means a substitute for objective scientific methodology, it serves as a valuable complementary approach. Indeed, phenomenology "illuminates the very ground of objective science, and this is because any objective science necessarily begins with experience" (Sheets-Johnstone 2011, p. 501). Phenomenology, and its influence on the work presented in this thesis, is further described in Chapter 3.

Seibt (2016) points to a common trait throughout HRI of describing interactions between humans and robotic artefacts in HRI through using a so-called *intentionalist vocabulary* (Seibt 2016). She characterises this as the "description problem of social robotics" (p. 107). Her assessment is that the HRI and SR fields do not have a suitable vocabulary to "describe human-robot interactions properly" (2016, p. 108). Therefore, it leans on metaphorical extensions of the mentalist idioms used to literally describe human capacities. A common strategy to avoid this is by treating these interactions as fictional. The problem with this is, however, that there is no such thing as a fictional social interaction. Any form of social interaction must make either an implicit or explicit commitment. "If one behaves as if making a commitment, a commitment is made, and just as we cannot fictionalize our commitments, we cannot fictionalize social roles" (Seibt 2016, p. 107).

One way to address this particular issue could be through the use of some common notation system for robot movement. Inspired by dance notation, and its relation to expressive gestures, Bianchini et al. (2016) attempt to integrate qualitative, relational, and behavioural aspects of movements. They take a rather large and concrete step toward the aim of advancing the articulation of the complex phenomenon of movement, through an approach that highlights the intertwined relationship between the contextuality of behaviour, motion in space, and semantics of movement. This can be seen as an effort toward what Korcsok and Korondi (2023) point to as a growing interest within HRI and SR in making detailed descriptions of robot and human behaviour during interaction. For this purpose, Korcsok and Korondi (2023) suggest HRI adapt the ethogram—a method used in ethology for analysing the behaviour of animals—as it supports making detailed descriptions of robot and human behaviour during interaction. They assert that the ethogram as a method could be used to increase the comparability of HRI studies.

In line with Seibt's (2016) critique, it is my understanding that intentionalist vocabulary and functionalist strategies obscure, rather than simplify, how human–robot interaction is understood. This vocabulary sustains a mismatch between a robotic artefact's *apparent* capabilities and its *actual* capabilities. The overall ambition of successful interactions is not exactly assisted ignoring the social asymmetry inherent in human–robot interactions. Rather than supporting mutual intelligibility, it undermines it. Through a thorough exploration of relevant theories described in Chapters 4, 5, and 6, this thesis offers a new terminology with concepts that are useful for clarifying what transpires during human–robot interaction. A synthesis of these theories is presented in Chapter 7, resulting in a system of concepts for accurately describing elements of robotic artefacts and robot movement that are significant for our experience of them as interactional partners.

1.3 RESEARCH CONTRIBUTIONS

This thesis consists of a summary and four papers, all of which make theoretical contributions. The summary itself contributes with a synthesis of three theories in the form of *kinetic affordances*. In this section, I provide a short summary of the papers included in this doctoral thesis, highlighting what I consider to be their main contributions.

Paper 1: *Turning Away from an Anthropocentric View on Robotics*

The paper explains some selected aspects of *Umwelt theory*, pertaining to the theory offered by biologist Jakob von Uexküll, and applies these to the case of a robot lawnmower. Uexküll was concerned with how other organisms were fitted to their way of life. Through textual descriptions and drawings, Uexküll illustrated what it *might be like* to perceive the world as another creature. This method is called *participant observation*. In using this method, one seeks insight into not only a creature's perspective of the environment but also how its perspective matters for its particular lifestyle. By using knowledge about a creature's physiological and behavioural traits and focusing on the relationship between them, the researcher attempts to identify the *meaningful objects* of its Umwelt. The meaningful objects and functional structures of behaviour make it possible to discern how the organism might perceive its surroundings and portray these so we may see the environment not from a human perspective, but from the possible perspective of the organism under consideration.

Building on an auto-ethnographical study conducted by my colleague of herself and her household lawnmower robot *Roberto* (a Husqvarna auto mower 308)[2], as well as my observations, study of documentation of the robot model, and discussions with Roberto's owner about its period of deployment in her garden, the paper introduces Uexküll's participant observation to the field of robotics research and adapts it to describe a lawnmower robot's metaphorical Umwelt. The contribution of this paper is a novel way of analysing what I in this thesis term the *artificial Umwelt* of this robotic artefact. The paper thus also demonstrates that this technology is simpler than it appears during use. For instance, the analysis reveals that the robot does not have any concept of "grass" even if its main task, in the view of a user, is to cut it. Thus, both the content and contribution of the paper is predominantly a theoretical endeavour, but with the potential for a methodological contribution. Chapter 4 in this thesis addresses this topic and delves further into it. Because Umwelt is a method for describing the subjective view of the world, based on biological life, I specify that when using it to describe the perspective of a robotic artefact, it is an *artificial Umwelt*.

As the first author of this paper, I am responsible for the production of the text. The contents were thoroughly and often discussed with co-author Jo Herstad.

Paper 2: *Engaging in Deep Wonder at the Experience of Encountering a Lawnmower Robot*

This paper builds on the Uexküllian concepts presented and made known by Soma and Herstad (2018). Still using Roberto the robot lawnmower as an example, the paper presents the lifestyle of *Genus Paramecium* and compares its Umwelt to the *artificial Umwelt* of Roberto. I illustrate the notion of Umwelt by comparing a single-celled protozoan and a robot lawnmower. I point at fundamental differences in the purpose underlying their movements, hence showing that the nature of movement in living organisms is different from the movement of robots. This analytical study introduces the notion of holding a *phenomenological gaze* on human–robot interaction, by examining an experience that sparked a genuine sense of wonder in the author. I describe and examine the experience of perceiving a robotic artefact as animate, or appearing somehow lifelike, even when I am deeply aware that it is in fact inanimate and lifeless. Even if my study is not designed as a phenomenological study it is initiated from wondering about *what* gives itself and *how* that something gives itself (van Manen 2014, p. 27), which is why I label it a phenomenological gaze.

The main contribution of this paper is therefore methodological. The subject matter of this paper originates in the author being "swept up in a spell of wonder" (van Manen 2014, p. 4) at the experience of encountering a lawnmower robot—and subsequently the following encounters with other similar robotic artefacts. As such, the paper exemplifies how a genuine sense of wonder can elicit a phenomenological gaze and lead to a deeper understanding of the phenomenon studied. The phenomenological gaze as a method for studying technology is further explored in chapter 3 of this thesis.

Paper 3: *Movement Acts in Breakdown Situations: How a Robot's Recovery Procedure Affects Participants' Opinions*

This paper presents a case study of a robot recovery procedure. It builds on a series of experiments with a Fetch robot that examined how a mobile robot moved. The experimental data was gathered in the MECS-project (I did not partake in its collection). The case study sprung out the serendipity of a breakdown situation taking place during some of the experiments when the robot would unexpectedly pause or rotate itself to recover from a navigation problem and a curiosity of the participant's understanding of the situation that arose. The overall contribution of this paper is its examination of how future study designs could consider breakdowns better and look at suggestions for better robot behaviours in such situations.

Through manual annotation and emergent coding of the data, we arrived at three themes. All comments in all experiments were examined and used to induce the themes, even if the great majority of participants did not experience any breakdown, delay, or other unplanned events. We found that both movement and stillness came across as communicative acts, even at times when there was no intention behind them, and that the "messages" of movements with communicative intentions went by unnoticed. As responsible for the qualitative analysis presented in this paper my main contribution was the concept of *movement acts*, a concept rooted in the axioms of communication by Watzlawick, Bavelas, and Jackson (1967) and Searle's (1969) *speech acts*. The concept is an analytical term that can be used to make sense of what the movement of a robotic artefact communicates explicitly, and it can be used to explore between intended and unintended interpretations of movements (or lack of movement). In splitting a robotic artefact's movement into movement acts designers may focus on the implicit and explicit communication and the act and thereby create better communication.

As the second author of this paper, I contributed with text throughout it. Primarily, I authored the background on semiosis and the description of our analytical procedure. The analysis itself was conducted by both the first and second authors. My main contribution to this paper was arriving at, describing, and discussing *movement acts*.

Paper 4: *Strengthening Human Autonomy. In the Era of Autonomous Technology*

In this theoretical paper, we discuss the concept of *autonomous technology* and compare it to the notion of *human autonomy*. In the paper, we use two empirical studies as illustrations to highlight inherent differences between the two. In this manner, we examine what autonomous technology can do, and perhaps more importantly, what it cannot do. Through this elucidation, the paper also touches on ontological differences between the living and self-governed, and the artificial and data-governed. This constitutes one of the main contributions of this paper, as well as my main contribution to the paper, the term *datanomous*. This term is more precise than *autonomous* for characterising these technologies, as it accurately describes what the technology is capable of. The term is intended to make obvious that it is *data* that governs these technologies.

In this manner, the paper also concerns the importance of relations and the qualities of human activity. We find that *human activity* and not autonomous (datanomous) technology is what strengthens or weakens human autonomy. Rather, knowledge of the autonomous (datanomous) technology and its *data* can be more important for human autonomy than direct control over the technology itself. Finally, we conclude this exploration by suggesting that human autonomy can be strengthened by datanomous technologies, but only if they support the human space for action. It is the purpose of human activity that determines if technology strengthens or weakens human autonomy.

As first author, I was responsible for the overall writing and progression of the paper. I arrived at the term *datanomous*, which is the central theoretical and conceptual contribution offered by this paper.

1.4 STRUCTURE OF THE THESIS

The contents of this thesis form three distinct themes. Its structure is as follows:

Chapters 2 and 3 constitute the background for this thesis. In Chapter 2, I position my work within Human–Computer Interaction (HCI) and Human–Robot Interaction (HRI). I present my view on interaction and interactivity and highlight the specific aspects of interaction with *robotic artefacts* that will receive focus in this thesis. This chapter also gives a brief explanation of the perception of animacy, a phenomenon that is highly relevant to understanding interactions with robotic artefacts. In Chapter 3 I give a brief introduction to the origins of phenomenology. The chapter explores the dynamic between phenomenology as a practice and phenomenology as a school of thought. I outline how phenomenology is often applied in HCI, as well as explain how this thesis maintains a *phenomenological gaze* on human–robot interactions.

Chapters 4, 5, and 6 each give an in-depth explanation of a theoretical perspective. In Chapter 4, I build on the theoretical discussions in *Papers 1* and *2* included in this thesis. I explain Umwelt theory and demonstrate how the notion of the *artificial Umwelt* can be used to illustrate how the environment might appear for a robotic lawnmower. This analytical tool is rooted in trying to discern the actual capabilities of the robotic artefact. Chapter 5 gives a theoretical discussion on the nature of affordances. Here, I compare Gibsonian affordances with those of Norman with respect to design. Further, I explore the way people skilfully evaluate the available opportunities for behaviour. It becomes clear that what a person chooses to do in a given situation is not just contingent on the affordances specified by ecological or sociomaterial information alone. Chapter 6 focuses exclusively on Maxine Sheets-Johnstone's phenomenology of movement. First, I give a thorough explanation of some of the central concepts of this theory, centred around the main concept *animate form*. As such, this chapter explores the significance of movement to living beings, revealing that the linkage between animation and agency is in no way trivial. Sheets-Johnstone's work on pronouncing the interwoven nature of *being* and *movement* sheds new light on our current understanding of human–robot interactions by emphasising aspects of our manner of being as animate forms.

The last theme contains only Chapter 7. Here, I discuss the three main conceptual contributions from the papers included in the thesis. I combine the terms *artificial Umwelt* and *datanomous technology* to explore the robotic perspective of its surrounding environment and characterise some of the ways this matters for human–robot interaction. I further explain how *movement acts* underscore presence in itself as meaningful, also in human–robot interactions. Finally, I outline *kinetic affordances* as a way of addressing movement itself as the meaningful relation between humans and robotic artefacts.

These concepts presented in this thesis more precisely describe what and how the robot is in its interaction with humans. Rooted in the phenomenological tradition, the thesis explores how humans—as living subjects—relate to robotic artefacts, seeking to investigate these relations. The contribution of this thesis is theoretical concepts intended to give a precise vocabulary for describing human interaction with robotic artefacts.

2 Background

> Millions of parents have bought computer toys hoping they will encourage their children to practice spelling, arithmetic, and hand eye coordination. But in the hands of the child, they do something else as well: they become the occasion for theorizing, for fantasizing, for thinking through metaphysically charged questions to which childhood searches for a response.
> — *Sherry Turkle, The Second Self*

2.1 SETTING THE STAGE

During the 1980s, the field of Human–Computer Interaction (HCI) emerged through the coming together of multiple, somewhat already related, fields: computer graphics, operating systems, human factors, ergonomics, industrial engineering, cognitive psychology, and the systems part of computer science (Hewett et al. 1992). In its early days, HCI favoured problems that were concrete and simple performance metrics (Duarte and Baranauskas 2016). Since then, the field's focus has changed in accordance with technological advancements. From the personal computer, through to mobile devices, and perhaps currently the connectivity between the digital devices that surrounds us (Internet of Things). The field has advanced not only through technological but also paradigmatic, shifts. As new technologies become available in new contexts, what is considered important must necessarily follow (Bødker 2015; Frauenberger 2019; Harrison, Sengers, and Tatar 2011).

We see a similar parallel with the emergence of Human–Robot Interaction (HRI) during the during the mid-1990s and early 2000s. Similar to how HCI shifted focus toward usability when computers entered people's homes, HRI arose out of a demand for increased attention on good interactions with robotic artefacts when this technology also moved from industrial environments into homes and other everyday settings.

Following this development, the average users of robotic technologies are now non-experts in the functional properties of this technology (Steinfeld 2004). While both fields share many similarities, Dautenhahn (2007) emphasise that HRI distinguishes itself from HCI because the interactions with robot technologies are often marked by their *spatial modality*. She refers to this as an *embodied nature*, the essence of which is described as the "need to coordinate their activities in time and space in real-time, often 'face-to-face'" (Dautenhahn 2007, p. 103). Similarly, Young et al. (2011, p. 54) points to robots' "well-defined physical manifestations" and their exhibition of physical movements that allow them to "autonomously interact within peoples' personal spaces" are the properties that set them apart from the technological artefacts traditionally considered in HCI. They emphasise that it is exactly this *tangible nature* that gives rise to their "unique effect on the social structures surrounding interaction" (Young et al. 2011, p. 54).

2.1.1 Interaction and interactivity

Interaction is undoubtedly central to both HCI and HRI. The word interaction itself, however, encompasses a multitude of meanings depending on its area of application which is naturally not limited to these two fields. Several standard dictionary definitions are quite broad, such as the one provided by Merriam-Webster online dictionary, where *an interaction* is defined as a "mutual or reciprocal action or influence." While this definition covers what an interaction is, it somehow does not quite manage to capture what it actually means to interact. Perhaps this is because its connotative meaning will change depending on the interactants associated with various kinds of interactions, and in turn how the mutuality or reciprocal influence unfolds. Luckily, when it comes to both HCI and HRI, the interactants generally considered to be involved are revealed in the very name; the interactions taking place are between humans and computers, or humans and robots, respectively. No surprises there. Nonetheless, the nature of the mutuality remains inconclusive.

One perspective on interactions that has had a major influence on HCI is that of Suchman (1987). Through the understanding of interaction from her own field, social anthropology, she suggested that "rather than just *using* machines, we *interact* with them" (Suchman 1987, p. 3 [emphasis added]). By taking seriously the computer as an interactional partner, she explained how the issue of human–computer interaction was essentially one of *mutual intelligibility*. This perspective on interaction focuses on what the interactional partners can understand about each other's actions. It is the "*relation* between observable behaviour, and the processes—not available to direct observation—that make behaviour meaningful" (Suchman 1987, p. 3 [emphasis added]). What she classified as *computational artefacts* are artefacts functionally capable of engaging with humans interactionally, rather than just being subject to use.

Building on this perspective, Bratteteig (2021) describes interaction as an interplay between a human and an artefact where both parties are active and alternate between—

or "take turns" in—acting and re-acting to the other. Central to this conception of interaction is the notion of a series of *exchanges* between humans and artefacts taking place. In contrast to interactions in physical sciences, the exchanges must involve more than merely "mechanical" responses to the world. She exemplifies this distinction by describing the activity of writing with a pencil. When being put against a piece of paper, the pencil makes physical changes in the world. These changes directly result from a combination of her manipulation of the pencil and the pencil's mechanical properties. The pencil makes a mark on the paper adhering to the path of manipulation through the physical process where graphite flakes off the pencil and onto the paper. This process is mechanical rather than reciprocal, resembling neither actions nor reactions.

Through the explanation of interaction offered here, a few things have been specified. Firstly, interaction with computational artefacts can be differentiated from use in general. Secondly, each action of the involved interactants influences the other part or parts to *counter* any action with an appropriate *re*-action. In other words, the interactants "answer" each other through a series of *exchanges*. Lastly, for these exchanges to succeed as a series of actions and reactions that answer each other in an appropriate manner, we must assume that they are all *meaningful*—mutually intelligible—to the involved parts. Based on this, this thesis defines interactions as *a series of meaningful exchanges between humans and computational artefacts*. Further, I adopt Suchman's (1987) *computational artefact* as an umbrella term and place *robotic artefacts* under it.

2.2 INTERACTING WITH ROBOTIC ARTEFACTS

While robot technologies are a fairly new innovation (Goodrich and Schultz 2007), the notion upon which they are based is not. Envisioned as *artificial beings* "self-regulating in ways we commonly associate with living, or animate beings" (Suchman 1987, p. 8), the idea of constructing artificial servants or mechanical creatures that are "endowed with life" appears in a variety of cultures across the world and throughout history (Goodrich & Schulz 2007). It is only recently—during the last century—that the word *robot* was popularised to replace the idea of *automata* or humanoid *androids*. In the 1920 Czech play R.U.R. (Rossum's Universal Robots) by Čapek (1920), *roboti* appear as simple and factory mass-produced synthetic human-like workers. Yet, the word *robot* represents more than just a kind of artefact. Not insignificantly, its etymology is directly related to the notion of carrying out labour (whereof its closest relative *robota* means "forced labour" in Czech). Woven into the very idea of various portrayals of artificial beings are questions concerning the nature of agency, dreams of human liberation from labour, and an awareness that any such creations would potentially shake with our very understanding of the mutuality of companionship. There is also an air of wariness concerning what might come to pass should any such man-made devices in fact gain real independence from their creators. While most people nowadays have some conception of what a robot is, there is still no unequivocal definition or formal requirements for what should or should not be regarded as one. The current conception of it encompasses a tremendous variety of forms and functions such as lumpy and mechanical machines or synthetically produced humanoids composed by some kind of synthetic biology. Nonetheless, it is the conceptual thematic found in the balance between human control and machine independence that dramaturgically intrigues.

"Every human tool relies upon, and reifies in material form, some underlying conception of the activity that it is designed to support" (Suchman 1987, p. 3). With precision, Suchman's observation pinpoints that the essence of robotic artefacts as tools is the same today as it was millennia ago. As I see it, the capability of "acting upon the world independently of real-time human control" (Soltanzadeh 2019, p. 1) constitutes the activity robotic artefacts, as tools, are designed to support. The technological means to this end is referred to as *machine autonomy*. Whether it is as mechanical creatures, artificial beings, or animated substances, and whichever way any such man-made creations might have been thought to gain their characteristic of lifelikeness, their essence lies in a capability for *agency* (albeit artificial). Indeed, "the word *agency* itself refers to the capacity to act and carries the notion of intentionality" (Dewey 1980; Young et al. 2011, p. 54).

The function of machine autonomy in the design of robotic artefacts is not considered an end in itself, but "a means to supporting productive interaction" (Goodrich and Schultz 2007, p. 217). Its implementation as a technological feature is not binary: Some describe it as coming in levels (e.g., Sheridan, Verplank, and Brooks 1978), degrees (Formosa 2021), or as instances of a variety of different phenomena (Bradshaw et al. 2013). The type and degree of autonomy vary with the specific task and environment a robotic artefact is intended to operate (Thrun 2004). Any apparent autonomy will not necessarily reflect its actual autonomy as technically constraining a system's operational space can make it appear more autonomous, "even the simplest machine can seem to function 'autonomously' if the task and context are sufficiently constrained" (Bradshaw et al. 2013, p. 57). Machine autonomy is just that, a technical feature, not an intrinsic quality. Unlike the autonomy of autopoietic systems (Maturana and Varela 1980), machine autonomy does not and should not carry with it the "metaphysical assumptions about the internal properties of the systems in question" (Soltanzadeh 2021, p. 2).

As it turns out, the self-regulation we associate with living beings is extremely hard to synthesise, even if it is relatively easy to mimic. Still, genuine self-organisation might not be strictly necessary for every form in which agency may appear, on account of how some agentive qualities, such as self-sufficiency and self-directedness, can be exerted within limited operational space (Bradshaw et al. 2013). While artificial agency will indeed lack intrinsic intentionality, some form of agency will nonetheless manifest by the functional movements accompanied by its manipulation of objects in physical, and inevitably, social space. In turn, the actual reach of a robotic artefact's capabilities might be difficult to ascertain. As Young et al. (2011) state, "agency contributes to the development of expectations of the robot's abilities (such as learning ability) or can create the expectation that the robot will be an active social agent, all in a much more prominent way than with more traditional technologies" (pp. 54-5).

The issue at hand is closely connected to the philosophical "problem of other minds" (Brincker 2016), which refers to the puzzle of how subjective and intrinsic qualities of entities—such as agency—can be discernible from an outside perspective. While this is sometimes called *attribution* (Bianchini et al. 2016; Malle 2011), it usually goes under *anthropomorphisation* in HRI. Anthropomorphism may also refer to deliberate design choices made to purposefully give a robotic artefact human trait, for instance through visual appearance or speech. Similarly, zoomorphism refers to robotic artefacts' that have forms similar to, or give associations of, animal morphology. Seibt, Vestergaard, and Damholdt (2020) suggest that the term *sociomorphing* constitutes a better description of the phenomenon of attribution, as it highlights how humans attribute not only the capacities of human social agents specifically but also the capacities of social agents in general. Sociomorophing is "premised on a diversification of the notion of sociality, taking 'sociality' to denote both what is expressed in social interactions and as well as what is experienced in such interactions" (Seibt et al. 2020, p. 55). In HRI this is established as how "people perceive robots to make autonomous, intelligent decisions based on a series of cognitive actions" (Young et al. 2011, p. 54). Generally, people "apply a variety of mental models relating to animacy, sociality, affect, and consciousness to explain their experiences and emerging relationships with robots" (Mutlu, Roy, and Šabanović 2016, p. 1909). This is seen as an important mechanism during interaction with robotic artefacts. It is also generally accepted that such inference is more often directed at robotic artefacts than to other technologies (Young et al. 2011). Because attribution is considered to be a mechanism that assists people in building and adjusting their expectations in the face of robotic artefacts, it should be considered of utmost importance for our understanding of interaction with them.

The connection between sociomorphing and the spatial modality of robotic artefacts seems to be nothing short of significant. Its physical presence in social contexts contributes to a kind of interactivity that indeed is dissimilar to interactions with other computational artefacts. As Bianchini et al. (2016) remark, humans "collect cues from the robot's movements and various transformations, and infers psychological properties about its perceptive skills (how much of the environment the robot is aware of), its ability to plan an action, to reason, to make decisions adapted to specific circumstances, etc." (p. 5). Bianchini et al. (2016) further hold that the expressivity of movement is independent of robotic morphology. Even when a robotic artefact has a completely non-anthropomorphic, non-zoomorphic form—such as an ottoman (Sirkin et al. 2015) or a toaster (Burneleit, Hemmert, and Wettach 2009)—the way these artefacts move will express *behavioural cues*. Such robotic artefacts have been referred to as *behavioural objects* (Levillain and Zibetti 2017). Bianchini et al. (2016) consider movements in themselves to be expressive and that they can be "extracted" from any morphological cues of a robot's physical design. To this end, they have systemised these cues into a model of *three levels of interpretation* for describing behavioural patterns, based exactly on this connection between movement and attribution: *animacy*, *agency*, and *social agency*. These categories signal the importance of sociomorphing to interaction and emphasise that attribution concerning agency—and inevitably capabilities—are primarily sourced in the *expression* of robotic movement.

In this thesis, I first and foremost approach robotic artefacts as a kind of tool, with a pronounced focus on the capability of movement as their central functionality. My interest lies specifically within the intersection between the functionality and interactivity of robotic movements as functional movements are a source of expressivity as well. I take there to be two major categories of robotic artefacts, differentiated by what their capabilities for movement are predominantly intended for. *Functional robots* move to enable independent execution of physical tasks, and thus their movements are primarily considered to be functional rather than interactional. *Social robots* move to engage human users socially. In other words, interaction is the primary function of the robot's movements.

2.2.1 Social functionality

The distinct group of robotic artefacts "designed to interact with people in a natural, interpersonal manner – often to achieve positive outcomes in diverse applications such as education, health, quality of life, entertainment, communication, and tasks requiring collaborative teamwork" (Breazeal, Dautenhahn, and Kanda 2016, p. 1936) are generally considered to be *social* robots. These artefacts are specifically equipped with *social-communicative functions*, designed with the primary task of reacting to humans as "social and communicative beings and refer particularly to these essential facets of human existence" (Gasser 2021, p. 329). The purpose for which social robotic artefacts are designed is "to interact with people in human-centric terms and to operate in human environments alongside people" (Breazeal et al. 2016, p. 1935). *Social robots* as a term is mostly used as a collective, often delineated as an umbrella underneath which you can place all robotic artefacts whose primary functionality is intended to be social interaction with humans (e.g., Fong, Nourbakhsh, and Dautenhahn 2003). The social character of a specific robotic artefact is judged by how well it engages people in social activity. Their sociability is considered to exist as a result of the human inclination for projection or anthropomorphising (Damiano and Dumouchel 2020).

One of the earliest definitions of social robots makes a split definition based on whether one takes on the perspective of a human observer and to which extent the robot design supports and validates this perspective (Breazeal 2003). Robotic artefacts that fall under the *social* umbrella may be characterised as a class of autonomous systems to which people apply a "social model" when interacting. Similarly, relational artefacts are a specific kind of robotic artefacts that are purposefully designed to present themselves as having "states of mind" (Turkle et al. 2006). It is held by some (e.g., Breazeal 2003) that the application of such social models enriches human encounters with them. Therefore, the idea that a system's intentions can be deliberately communicated through verbal, nonverbal, or affective modalities (Breazeal et al. 2016) is central to the development of sociable robotic artefacts. This way of approaching human–robot interaction which focuses on the design of social functionality, aims to facilitate socially meaningful interactions between humans and robots.

Other perspectives on robot sociality are sourced in the social structures that emerge or change through everyday interaction with robotic artefacts. These are robotic artefacts with functional purposes, deployed for instance in homes or workplaces. Forlizzi (2007) describes how a robotic artefact that works with people fosters social relationships, and that a Roomba in fact changed the cleaning activities of its users. Søraa and Fostervold (2021) describe how a robotic artefact designed and developed for the transportation of goods in a hospital became domesticated and the centre of several social activities through a process they call "social domestication" of robot technology, in which unintended social interaction is key. The robotic artefacts studied in these contexts are usually not covered by the "social" umbrella, in that their sociability has not been provoked through purposefully giving them an anthropomorphic or zoomorphic look. The perspectives of interest in research concern how human social structures are seemingly sensitive to the physical presence of an artificial agent. They are not concerned with the perceptual experiences of interaction with robotic artefacts, but instead on structure of work and the role of the technology in human environments. Robot sociality such as the social characters of robotic artefacts is conceived as a distributed property that exists in the relation between users and robots (Damiano and Dumouchel 2020).

Young et al. (2011) present three perspectives on social interaction with robotic artefacts that can be used to explore and understand them. These perspectives are not sensitive to whether the robotic artefacts are intended to elicit social responses from people during interaction, but to the actual responses. First, *visceral factors of interaction* concern more or less immediate and automatic responses such as fear or joy at seeing and interacting with a robotic artefact. These are impulses that can be difficult to control. Second, *social mechanisms* refer to the way people tend to interpret the interactional exchanges of robotic artefacts as socially meaningful and in turn respond through human social communication. Lastly, they point to how *social structures* may undergo changes during prolonged interactions with robotic artefacts.

These studies show that also robotic artefacts that are not designed with social interaction as their primary functionality can be sociomorphed. This witness to the myriad of ways sociality may be experienced during interaction with robotic artefacts, from outright directly copying the structure of human–human social interaction (such as the humanoid Sophia by Hanson robotics[1]) to a more presence-based sociality (such as 360° Presence[2]). Seibt et al. (2020) argue that each such experience can be categorised as a "Type of Experienced Sociality" (TES). TES highlights that humans do not necessarily attribute robotic artefacts to a *human* social agency, but a social agency in general. The founding level is a "feeling of co-presence" with the robotic artefact, while at the other end of the scale, the robotic artefacts will be experienced as skilful in their social interactions, capable of cooperation and teamwork. "In order to intuitively rehearse what might be meant by a TES and the phenomenological differences involved, the reader might imagine what it feels like to be with agents with different responsive capacities—e.g. being-with a lizard, vs. being-with a cow, vs. being-with a cat, vs. being-with a dog, vs. being-with a human" (Seibt et al. 2020, p. 59). It is important to note, however, that the robotic artefacts will not actually "qualify as social agents on any of these accounts, even though they may simulate criterial capacities at certain levels so well that they are perceptually indistinguishable" (Seibt et al. 2020, p. 56). Indeed, Seibt (2016, 2018) has problematised the recurring use of fictionalist strategies and intentionalist vocabulary used in HRI literature for describing human–robot social interactions. These descriptions side-step the social asymmetry inherent to them.

> The conceptual norms that govern the semantics of the verbs highlighted—recognizing, engaging in social interactions, perceiving, interpreting, communicating, learning, following a norm—require that the subject of these verbs is aware, has intentionality or the capacity of symbolic representation, and understands what a norm is. Since robots—currently at least—do not possess such capacities—at least not how they are defined relative to our current conceptual norms—such characterizations are strictly speaking false. At best, they are metaphorical extensions of the—in philosophical terminology—'mentalist' or 'intentionalist' idioms we use literally to describe human capacities. The current literature in social robotics is replete with such metaphorical descriptions of robots "guessing," "smiling," "greeting," "responding," etc. (Seibt 2016, p. 106).

Agents that are within social ontology considered to be *real* social agents are living beings, and their actions are thus considered to have *real* intentionality. Robotic artefacts, on the other hand, can only *simulate* such actions, as they do not possess intentionality. Time and again are robotic artefacts referred to as "social," when in fact they are not. Seibt (2016) highlights the absence of an adequate vocabulary for describing the social character of human–robot interactions. "Metaphorical extensions of mentalist vocabulary are not scientific, and we cannot somehow 'apply yet bracket' our common terms for social interactions, since social actions cannot be fictionalized" (Seibt 2016, p. 107). Ignoring the social asymmetry of human–robot interactions merely contributes to obscuring a robotic artefact's actual capabilities by focusing only on apparent capabilities. As long as this mismatch is perpetuated, the interaction between humans and robots will not be mutually intelligible. When acknowledging this asymmetry, the lack of capabilities of robotic artefacts becomes visible. Even if a robotic artefact is purposefully designed to trigger all the right interactional points to evoke a social response from a person, they are not social agents. Being a social agent is an ontogenetic and intrinsic quality that cannot be added as a feature.

2.2.2 Meaningful functionality

When designing a robotic artefact for independent or collaborative task performance, the need for machine independence must be balanced with people's experience and the need for control over the artefact during interactional situations. By this, I mean not to invoke any doomsday scenario in which we are posed with a genuine disability to "pull the plug" of some rouge autonomous computer or robotic system. In this thesis, I will specifically address the ways functionally necessary movements of robotic artefacts *can* or *may* be meaningful to humans as an interactive partner, and what it means for these exchanges to be mutually intelligible. Functional movements become interesting from an interactional perspective in that people will, in some way or another, be attentive to the movements and feel the need to respond or adjust themselves in accordance. For instance, a situation will turn interactional if a person who is "just watching" a robotic artefact in fact also adjusts their behaviour—even if only ever so slightly. In such a scenario, the functional movements double as exchanges in interactions, from the perspective of the human. This is recognised by Goodrich and Schultz (2007), who acknowledge that interactions between humans and robotic artefacts will be inherently present in all robotics, regardless of the level of autonomy. They present a perspective on interaction as *requiring* communication between humans and robots. The interactions are desired to be *beneficial* in some sense, and thus take on a specific mission, as "*the process of working together to accomplish a goal*" (p. 217 [emphasis in original]). In my reading of this definition, the robotic artefact is torn between being a tool and a supposed partner.

Distinct from the law-abiding interactions studied within the natural sciences or the socio-culturally determined rules (laws and norms) that govern social interactions, the most unique characteristic of the interactions of computational artefacts is the exchanges are designed by someone to *transpire* in a specific manner. This is a non-trivial aspect of understanding the nature of the mutuality central to these interactions. Design in its most pronounced form involves the decision-making by a few people on behalf of many. This has of course some moral and ethical ramifications. However, the issue of meaningfulness in interaction is ultimately a contextual one and an issue of how the system should sort between what is relevant and irrelevant to it. As Dourish (2004) points out, "the determination of relevance—or of contextuality—is not one that can be made a priori. It is an emergent feature of the interaction, determined in the moment and in the doing" (p. 23). Thus, designers are required to strike a compromise between intended and actual use. Indeed, the design of computational artefacts revolves around not only determining the intended meaning behind each exchange but analysing and considering *possible* meanings as well. In other words, it involves the prediction of how the exchanges will be understood during use, and whether they will be meaningful to the involved parts throughout the intended sequence of actions and re-actions. As subordinate to computational artefacts, this applies to the design process of robotic artefacts as well.

It is not to get away from that a view on interaction as communication or information exchange, such as that described by Goodrich and Schultz (2007), confuses the multifaceted nature of meaning for multimodality. Rather than focusing on the manner multiple meanings may arise during the exchanges, the exchanges are considered the sending of messages through the various modalities as different channels for communication. The exchanges may only be meaningful in those cases where the "message" is correctly interpreted by a human interactant and vice versa. I find that the way Bianchini et al. (2016) refers to the movements of robotic artefacts as a source of *expressivity* gives a more accurate description of the multifaceted nature of their spatial modality.

When we consider the task of designing interactions that enable mutual intelligibility, I maintain that a focus on manipulating the dynamic expression of a robot's movement acknowledges that *meaning* cannot be provoked in any interactional exchanges. When the designer is designated as the primary sense-maker, the process of sense-making happens long before the interaction itself takes place. The mutual intelligibility may quickly turn akin to a *non sequitur*. Like with Chinese whispers, the "message" becomes more and more distorted with each "failed" exchange. Viewing interaction between humans and robotic artefacts first and foremost as the exchange of messages leapfrogs the central role meaningful exchanges fulfil in mutually intelligible interactions. The mission or goal of HRI as outlined by Goodrich and Schultz (2007), may easily fail.

2.3 PERCEPTUAL ANIMACY

Whether something is alive or not—whether it is *animate* or *inanimate*—appears to be a concept humans grasp from a very young age. As a linguistic principle, it appears to be one of our most fundamental, occurring in languages across the world. The state of a thing's *animacy*—derived from Latin *anima*, meaning "breath" or "soul"—is a grammatical and semantic element, representing an ontological categorisation where humans, animates, or inanimates are the entities (de Swart and de Hoop 2018). It appears evident that the criteria in this categorisation would be coupled with the presence of qualities we associate with living beings. Yet, it is not necessarily dependent on biological criteria alone. Different philosophical perspectives on the principles of vital forces are significant in contributing to what a culture, and in turn a language, consider to be the qualities of living beings. State or level of animacy is not expressed in the exact same ways across languages but it is generally agreed upon that there are no sharp lines separating these categories in any languages. Instead, they operate on a hierarchical continuum (Santazilia 2020).

Gelman (1990) points out that there is "a clear distinction between what defines a category […] and what commonly determines which objects are assigned to which categories" (p. 103). By this, she means that even when there are causal principles that define the categories, these principles do not specify how the various entities in these categories are to be recognised in everyday contexts. She highlights a mediating process between perceptual mechanisms informed by causal principles and accumulated knowledge. When it comes to the categorisation of *moveable entities* as either animate or inanimate, Gelman, Durgin, and Kaufman (1996) accentuate that the categorisation will not be made based on the dynamic characteristics of trajectories alone. These may be influential but remain just as ambiguous as static characteristics in isolation. The authors use the example of the characteristics of tomatoes to illustrate: Tomatoes are red and round, but so are many apples. In addition, tomatoes can also be orange and ovals, but they are tomatoes nonetheless. When it comes to the everyday categorisation of movable entities, it will be based on the dynamic qualities of the movements in addition to their static qualities. When the dynamic and static qualities of living beings repeatedly co-occur, certain static qualities will become associated with the animate category. Gelman (1990) emphasises that we will recognise animate entities from inanimate even when they are not moving because we form ideas about the material composition of objects in the different categories. Gelman et al. (1996) argue that humans become aware of material types that correspond to the entities belonging to either category. Thus, the static qualities that reveal material makeup will be just as important as movement when discerning animate from the inanimate.

This would seem to contradict the phenomenon of *perceptual animacy*, which refers to the attribution of animacy to an entity—that would not otherwise be categorised as animate—based on its movements alone. This phenomenon was formally outlined half a century ago by Heider and Simmel (1944) who demonstrated that their participants interpreted movements of geometrical shapes as intentional actions. When describing the movements, personality traits and motivations that correspond with human social behaviour were attributed to the geometrical shapes. The experiments are seen as a clear illustration of how humans tend to attribute agency and social meaning to abstract or inanimate objects. Another pioneering body of work on perceptual causality, the experimental phenomenology of Albert Michotte, examined how people perceive the functional relations between objects (Thinès, Costall, and Butterworth 1991). Michotte believed that causality could be immediately experienced through vision alone and was convinced that we are able to "perceive actions performed by objects or animate beings ("agents") on one another in the same way we can see simple kinetic movements" (Wagemans, van Lier, and Scholl 2006, p. 3). Thus, Michotte held that *seeing* causality differs fundamentally from believing or inferring its existence (Michotte 1963, p. 220). It seems likely that there exists a perceptual "grammar" of causality and animacy where the recognition of certain dynamic trajectories as animate—likely directionality, discontinuity, and environmental responsiveness—is immediate and irresistible (Scholl and Tremoulet 2000). For instance, objects that continuously follow the same fixed path (like a ceiling fan), or objects whose movements appear random (like tree branches swaying in the wind), are unlikely to be viewed as animate (Luo, Kaufman, and Baillargeon 2009).

However, instead of looking at these perspectives as contradicting each other, perceptual animacy and the ordinary categorisation of animacy are two sides of the same experiential phenomenon. No matter what the exact mechanism responsible for the experience of this phenomenon is, the basic nature of perceptual animacy seems to be rooted in an apparent mismatch between the ontological categorisation of an object's animacy, and its current perceptual appearance. As a perception of animacy unfolds in someone's experience, at least some of the current perceptual qualities of an object change its usual ontological status from inanimate to animate. The human ability to categorise is flexible and dynamic—after all, they are human constructions, existing to aid us in our everyday sense-making of the world around us. The concepts are grounded in our experiences and will thus remain relevant to our everyday lives. Gelman et al. (1996) point to machines and robotic artefacts as particularly interesting. "Although they appear to move on their own, they are quintessentially ambiguous. They are made of inanimate material and do not exhibit biomechanical motions, nor do they adjust very well to local environmental problems" (Gelman et al. 1996, p. 181). Gelman et al. (1996) describe the ontological categorisation of moving entities as a complicated interplay between form, material, and dynamic trajectories. It is likely that robotic artefacts will remain troublesome in a binary classification system of animacy. Instead, they will be placed in a hybrid ontological category for machines. What such a hybrid category entails is not yet known (Kahn et al. 2011).

3 The phenomenological gaze

> But, the further we delve into the phenomenological literature, the clearer it should become that phenomenological method cannot be fitted to a rule book, an interpretive schema, a set of steps, or a systematic set of procedures.
> — *Max van Manen, Phenomenology of Practice*

3.1 A BRIEF INTRODUCTION TO PHENOMENOLOGY

There is no standard methodological way of doing phenomenology, and—as I attempt to explain in this chapter—nor should there be. Instead, one can "think of the basic method of phenomenology as taking up a certain attitude and practising a certain attentive awareness to the things of the world as we live in them rather than as we conceptualise or theorise them, and as we take them for granted" (van Manen 2014, p. 41). The concern of phenomenology is, in some very broad strokes, with the structure and nature of *how we experience*, from a subjective and first-person perspective. Structures of experience deal with matters that typically involve intentionality (directedness of experience) toward the things in the world, encompassing concepts like perception, thought, bodily awareness, or social and linguistic activity. In other words, phenomenology concerns the meaning of "things" (for instance objects, people, places, events) we meet in our everyday lives have in our experience. Summarised into one simple sentence, the subject matter of phenomenology is the study of *experience* (Smith 2018).

Käufer and Chemero (2015) describe phenomenology as two different *strains*, where one is concerned with "the structures that make a shared, objective world intelligible" and the other "gives a description of subjective experiences, especially experiences that are unusual and hard to explain" (2015, p. 2). Herein lies a discrepancy that long eluded me. It has been my experience that because phenomenology may be considered both a philosophical tradition as well as a research paradigm, it can be difficult to pin down exactly what phenomenology is. I now understand this distinction as trying to understand *what experience is* and attempting to describe *what a specific experience is like*. I categorise them as philosophy and practice, respectively.

I have found it helpful to distinguish the nuances represented by the role *experience* assumes in a study. In this chapter, I therefore briefly present phenomenology as philosophy and as practice. Thereafter, I give a few examples of how some of the forms phenomenology has taken in HCI and HRI. Finally, I explain how I understand phenomenology to have guided this study in the form of a *phenomenological gaze*.

3.1.1 Phenomenology as a philosophy

Quite literally, phenomenology translates to "the study of phenomena" (Smith 2018). However, van Manen remarks that even if phenomenology appears to be formed similarly to "bio-logy" and "psycho-logy"—where the first part refers to a subject, and the second part (from the Greek word *logos* meaning "word" or "study") refers to the science or study of this subject, "phenomenology does not have a subject matter or subject domain in the same sense; a phenomenon is not a subject" (van Manen 2014, p. 27). The history of the phenomenological tradition is deeply connected with the use and etymology of its root word, *phenomenon*, meaning "that which appears." This meaning can be traced back to philosophy's infancy when Plato made the distinction of phenomena as "how something appears" as opposed to "reality". Throughout the history of philosophy, both what is considered to appear, and how it is appearing—how we understand phenomena—will vary and be dependent on epistemological convictions (such as empiricism or realism). From being understood as the grounds for building knowledge, perhaps especially in science, *phenomena* took a slight change of meaning throughout the 19th century as psychology emerged.

One of the major steps in the history of philosophy which contributed to the establishment of the phenomenological tradition was Kant's envisioning of consciousness, a radical view at the time of writing. Kant's two-stem theory of cognition put forth that sense impressions and concepts are not the same kind of mental content, followed by a recognition that subjective structures are specific to experience. Right around the time when psychology became established as a science separate from philosophy of mind, Brentano, influenced by Kant, conceptualised a distinction between mental and physical phenomena. Brentano's conceptualisation of mental phenomena are acts of consciousness (or their contents), and physical phenomena are objects of external perception. Further, mental phenomena are directed toward some object—physical phenomena—that exist "intentionally" in any act of consciousness. This conception of phenomena became the basis for phenomenology. That is an intentional directedness as the cornerstone of Brentano's "descriptive psychology" for which he also used the word "phenomenology" (Käufer and Chemero 2015).

Even so, Edmund Husserl is usually considered the "founder" of phenomenology and started using the word to denote the specific approach to philosophy in his work around 1890-1900. As Brentano's student, Husserl built on his ideas of descriptive psychology. Husserl further integrated these with ideas from Bolzano's distinguishing of subjective and objective ideas or representation. This became the basis for phenomenology as we know it today. In the spirit of Husserl's original conception of phenomenology, we may say that it is the "study of consciousness—that is, conscious experience of various types—as experienced from the first-person point of view" (Smith 2018). Indeed, it was defined by Husserl in *Ideas I* (1913, ¤¤33ff) as "the science of the essence of consciousness" and centred on the defining trait of intentionality, approaches explicitly "in the first-person" (as cited in Smith 2018).

Käufer and Chemero (2015) explain how Husserl hoped that by defining the methods and basic concepts of phenomenology, his followers could (and would) adopt to practice phenomenology in various domains, so it would flourish as a research field (or phenomenological "school"). While phenomenology did not become a scientific philosophy, Husserl's efforts still sparked a phenomenological movement "followed by a flurry of phenomenological writing in the first half of the 20th century" (Smith 2018), even if many of his followers disagreed on aspects of his thinking (Käufer & Chemero, 2015). Through the influence of Brentano, three major ideas from Kant have directly influenced the classical phenomenologists (usually considered to be Husserl, Heidegger, Satre, and Merleau-Ponty), even if they in their ways disagreed with certain aspects of his thought. These three ideas concern (1) the subjective structures that constitute the objects of experience, (2) the temporal structure of synthesis, and (3) subject–object identity (Käufer and Chemero 2015).

Especially the last idea, which is also closely related to the more familiar mind–body problem, is fundamental to how phenomenology establishes itself as a philosophical tradition that breaks with dualistic thinking. Phenomenology as a school of thought cannot be severed from the history of philosophy of mind, for it is deeply rooted in it through a set of philosophical movements made in reaction to the prevailing epistemological and ontological attitudes. Its emergence marks a move from the dualist and cognitivist conception of the mind and body (and perception and ontological and epistemological questions), toward both the philosophical tradition and the research paradigm we now refer to when speaking about phenomenology. The phenomenological movement within philosophy is not easily summarised, but if I were to draw forth one specific aspect, it would be the conception of the subject as *being-in-the-world*. The structure of thoughts is not of some substance different from physical matter but exists in the same realm as the objects of the world, and they are inseparable from it.

Thus, even amongst those considered to have founded phenomenology, we find different conceptions of phenomenology, different methods, and different ideas of its results (Käufer and Chemero 2015). Fällman (2003) characterises phenomenology as being "made up of a number of combined contributions, which originate from several different and not always mutually supportive sources" who all "have individual interpretations and ideas about what phenomenology *is*; what it *does*; and what and whom it is *for*" (p. 16).

3.1.2 Phenomenology as a practice

A notion that I find is next to never communicated about phenomenology is that it is "not just the name of a philosophical perspective, it is also the source of questioning life meanings as we live it and the nature of responsibility for personal actions and decisions" (van Manen 2023, p. xv). Therefore, what is easily misunderstood about the phenomenological method is that when one raises a phenomenological question (rather than a research question), one asks for the essence of a lived experience, and not an explanation of the phenomenon.

The basic things in everyday life (the lifeworld) are *ineffable*, that is, pre-verbal and hard to describe. The scientific result of a good phenomenological study is a description that constitutes the essence of something so that the lived experience is revealed to us. Such a description should grasp the nature and the significance of an experience in a way unseen until now. The essence of the craft is to put into words things that have never been fully articulated before, capturing the variety and possibility of human experience in condensed and transcended form. This makes phenomenological research a scientific approach for questioning the essential nature of a lived experience, aiming to describe it as correctly as possible *how* and *what* unfolds, but without theorising about *why* it does. Phenomenology concerns all things that contribute to meaning in our ways of being in the world—including sociocultural and historical traditions. A phenomenological result can only be achieved through an attentive practice of thoughtfulness and a search for what it means to be human (van Manen 2014).

When *doing* phenomenology, one attempts to take on the *phenomenological attitude*. This involves practising what Husserl called *bracketing*, and attempt to, as best one can, suspend one's *natural attitude*.

> Husserl reserved the notion of 'natural attitude' not just to point at the taken-for-grantedness of everyday thinking and acting. For him, the natural attitude is manifested in our natural inclination to believe that the world exists out there, independent of our personal human existence. The challenge for phenomenology is not to deny the external existence of the world, but to substitute the phenomenological attitude for the natural attitude in order to be able to return to the beginnings, to the things themselves as they give themselves in lived through experience—not as externally real or eternally existent, but as an openness that invites us to see them as if for the first time (van Manen 2014, p. 43)

Assuming the phenomenological attitude does not involve denying or opposing the scientific method or the existence of an external world. It means that we, when attempting to investigate the structure of an experience or a phenomenon, should avoid as best we can to bring with us our *common sense*. The natural attitude is how we meet the world with the knowledge we take for granted. This is a bodily and cultural knowledge we have accumulated throughout a person's entire life. And while this knowledge naturally cannot be completely suspended, once we no longer take the knowledge for *granted*, we may see the phenomenon at hand anew. When we bracket out the natural attitude and step into the phenomenological attitude, we are free to inquire about the structure of the *experience* in a manner traditional scientific methods cannot.

Indeed, the rationale of phenomenology is that human life can be made intelligible, even if there is always an element of the ineffable (van Manen 2016, p. 16). It may be that phenomenological research appears as a scattered methodology due to its absence of an objective quest. While those with other metaphysical convictions will find this a weakness, I consider this a strength because the researcher finds themselves free to decide how best to get to the essence of an experience. Yet, its inherent lack of an objective enterprise likely contributes to an occlusion of what a phenomenological study might offer. For this reason, the phenomenological ambition toward making the ineffable effable, as it were, need not include answering *why* the experience took place or *what* the naturalistic meaning of a phenomenon may be. Instead, describing the *actual* lived experience is in focus.

3.2 PHENOMENOLOGICAL PERSPECTIVES

Phenomenology is, as Fällman (2003) points out, "just as much about positioning oneself towards aspects of earlier work in phenomenology [...] as it is to accumulate on and take that work further" (p. 15). Phenomenology, as a method of inquiry, should not be reduced to a set of standard strategies and techniques (van Manen 2023). As a method for questioning rather than answering (van Manen 2014, p. 26), a phenomenology is a well suited for widening the explorative space of a study. It allows the researcher to approach a phenomenon of interest with curiosity, rather than with a determination achieving a specific scientific result. Phenomenology operates under a different rationale from natural sciences and distinguishes itself from other social sciences in its explicit focus on meaning (van Manen 2016, p. 11). Thus, it is not that phenomenological research is incompatible with the themes of HCI and HRI, as Fällman (2003) is proof of, but the idea of what constitutes knowledge interesting to technology research differs.

The distinction I attempt to highlight is that phenomenological terminology has proven useful for uncovering new perspectives on human interactions with computer systems and other interactive artefacts. However, these concepts will not, in themselves, give a thorough and in-depth description of the actual experience of interacting with computational and robotic artefacts.

3.2.1 In HCI

Within the last few decades, phenomenology has gained a foothold in HCI. Some early examples include Winograd and Flores' (1986) use of Heidegger's *readiness-to-hand* to shed new light on breakdown situations in interaction with computers. Ehn (1988) also took on a Heideggerian perspective to understand the notion of skill and Wittgenstein's *language-games* to understand design and use of computational artefacts. Robertson (1997) presents a taxonomy of embodied actions developed on the basis of Merleau-Ponty's phenomenology, with a special focus on the reciprocity of perceiving and being perceived. Svanæs (2000) used perspectives from Merleau-Ponty's phenomenology to understand the holistic nature of interactivity. Another notable contribution from Svanæs (2013) gives an in-depth account of phenomenology in HCI. In a similar vein, Loke and Robertson (2013) provide an instructive overview of phenomenological frameworks and approaches within the HCI field. They also describe a design methodology for movement-based interactions with technology that focuses on a first-person perspective. Fällman (2003) gives a comprehensive description of the *phenomenological attitude* and demonstrates its relevance to understanding and designing mobile information technologies.

Phenomenology in the field of HCI is perhaps most known through Paul Dourish's book *Where the Action Is* (2001). This book is not just instructive in its offer of—what was at least at the time—a fresh and somewhat unknown perspective on HCI. It also established a collective term for the new approaches that were emerging, labelling them as *embodied interactions*. It seems that especially the introduction of whole-body interaction has demanded a shift toward learning from phenomenological perspectives. Dourish explicitly addresses the ongoing shift from the traditional model of the world in rationalistic terms of plans, procedures, tasks, and goals toward digital artefacts that exploit our familiarity and facility with the everyday world. Because software depends on our ideas of representation and reality, they inadvertently reflect certain philosophical commitments. Thus, in bringing attention to the relevance of philosophy in technology studies for developing new technology, and subsequently new ways of interacting with them, Dourish gives an instructive introduction of phenomenology in HCI for a wide audience. In his words, phenomenology is in an intimate relationship between "our inner experience and the mundane world that we occupy" (2001, p. 127), where the world is already filled with meaning.

Dourish (2001) holds that knowledge about philosophical heritages, for instance, *embodiment*—which carries with it practical action in the world being the foundation for conscious experience—is useful also in practical fields because such an awareness allows the researcher to understand the contributions and opportunities from new forms of technological practice. A similar notion is held by Svanæs (2000) who points out that "philosophy can be used as a resource and inspiration without having to become a philosopher" (p. 11). Notably, van Manen (2014, p. 23) asserts that one does not have to be a philosopher to do phenomenology.

Dourish (2001) further emphasises that even if technology and engineering may view philosophy as irrelevant, argumentation from philosophy is deeply relevant for understanding the limits of what can and cannot be achieved in the venture for better and more effective interactions between people and computational artefacts. The phenomenological perspective allows the researcher to understand how meaning emerges relative to our needs and actions, and that it is not only structured around the way the world is physically organised. Social and historical structures of meaning are equally important. Thus, as Dourish (2001) highlights, maintaining a perspective of *meaning* in the use of technology users become freer to create and communicate the meaning of the actions they perform, rather than having to struggle with meanings that are rigidly encoded into the technology itself. The major lesson Dourish draws from phenomenology to the field of HCI is that its perspectives begin "to illuminate not just how we act *on* technology, but how we act *through* it" (2001, p. 154).

What Dourish (2001) demonstrates is that the phenomenological perspective provides insight into how humans orient themselves toward or through technology in a way existing perspectives could not. For me, his endeavour constitutes an example of the strain of phenomenology that recognises skills and bodies are foundational to the intelligibility of the world. In this case, phenomenological perspectives are used to illuminate a phenomenon in a new way. By this, I mean that phenomenology is to a larger extent used to understand the use of technology, and to a lesser extent used to describe the actual experience of using technology. Another example is the work resulting in Hornecker and Buur's (2006) highly influential framework for Tangible Interaction is quite clearly based on work with a strong phenomenological anchoring, yet *subjectivity* is never mentioned. Conversely, Fällman's (2003) constitutes an example of a study attentive to the actual experience of interaction with mobile technologies. By being fully committed to the phenomenological attitude, he manages to capture how one might go about conducting a phenomenological study with technology in the context of HCI and use the results to inform design.

That is not to say that this specific way of applying phenomenology does not take actual experience seriously, but that it aims to achieve something different from phenomenological research into actual experience. I bring this up not because one way of doing phenomenology is better or more correct than the other. Rather, I wish to share an observation that I, at least, have never come across being explicated. In my current understanding of phenomenology, the way it is often used in HCI is by aligning phenomenological perspectives with the research in a way that opens it up to findings that pay attention to the nature of human experience of the interaction. I characterise this approach as *phenomenologically aligned* research, which I believe to be different from distilling and describing actual experience in that it adheres to the phenomenological focus on the first-person perspective, yet is not *primarily* interested in collecting specific subjective experiences for the purpose of describing that specific experiential phenomenon.

Indeed, the researcher is “confronted by the challenge to focus on those parts of the phenomenological literature that would be of relevance to his or her research interest” (van Manen 2014, p. 30). By reflecting on what a specific study aims to achieve and considering the explicit role *experience* will have toward this aim, the relevant phenomenological grounding can be appropriately customised.

3.2.2 In HRI

Traces of phenomenological thinking are, at least in my eyes, less visible in HRI. Still, there are some examples, such as Dautenhahn’s (1997) description of the phenomenological dimension of social understanding in the context of developing socially intelligent autonomous agents. More recently, inspired by the relations between humans and animals, Coeckelbergh (2011) gives attention to how robotic artefacts appear to humans using phenomenology[1] characterising they as *alterity relations* (from Ihde 1990).

One significant work springing out of phenomenological thinking stretches back all the way a time before the conception of both HCI and HRI. Dreyfus’ (1965, 1986) critique of contemporary AI research argued why the assumptions prevailing within the AI field could never be fruitful to the end AI researchers believed they were. Once the notions he presented were slowly accepted as accurate, Dreyfus’ perspectives became embedded in the fields by virtue of laying down the stones for a new direction, influencing several of the fields that eventually formed HCI and HRI (such as cognitive science and AI). The phenomenological influence of Dreyfus’ thinking is more evident within HCI through the use of his take on equipment and skilful action. For HRI, his ideas have been worked into the bearings of present robotics through the fundamental change in thinking taking place in the wake of his (Dreyfus 1965, 1986) commentary (Käufer and Chemero 2015). This exemplifies that phenomenology has had a significant influence on a field without a standardised methodological formula.

The focus on using standardised metrics for comparing results across studies can be traced to the large variety of robot designs and application areas. “To be able to compare the performance of one robot to another, particularly if these robots are developed at different labs, well-defined tasks and context are necessary in order for results to be comparable. Slight variations in the environments in which the robots operate might already influence the performance of the robots” (Bartneck 2023) (p. 1). One of the most cited papers in HRI is Bartneck et al.’s (2009) description of the Godspeed Questionnaire Series. This series is a set of five consistent questionnaires linked to five key concepts within HRI: *Anthropomorphism*, *animacy*, *likeability*, *perceived intelligence*, and *perceived safety*. It is commonly used to make participants evaluate their interactions with a robotic artefact.

That this paper is one of the most cited within HRI (Bartneck 2023), as well as the emergence of new scales (e.g., Carpinella et al. 2017; Devin et al. 2018; Ho and MacDorman 2010; Schaefer 2016; Spatola, Kühnlenz, and Cheng 2021) testifies to this being a widely used data gathering method. While this is certainly not the case for all of HRI—as is explicated by Dautenhahn (2014)—a certain preoccupation with "measuring" perceptions is one research trend within the field. This specific manner of approaching experience is fundamentally different from a phenomenological perspective. It is not that HRI is generally disinterested in human experience. Rather, it is merely interested in another aspect of experience than phenomenological research is. In the Godspeed Questionnaire Series, for instance, all the items a participant is asked to put on a scale are tied to an actual experience of the phenomenon of interacting with a specific robotic system. Take, for instance, the concept of anthropomorphism. When using the questionnaire to examine the participants' experience of the interaction, one is primarily interested in the degree or level to which the robotic artefact was anthropomorphised. However, this grading cannot explain or describe what the experience of anthropomorphising *is*. Neither can it answer how it differs from, say, zoomorphing or sociomorphing. This perspective takes for granted the experience of anthropomorphism as a universal experience that does not need explaining or elucidating.

3.3 HOLDING A PHENOMENOLOGICAL GAZE ON HUMAN–ROBOT INTERACTIONS

The study described in this thesis is not a phenomenological one in the sense phenomenological research is outlined by van Manen (2014). Van Manen describes phenomenological research as a human science research approach driven by a pathos "to discern the primordial secrets of the living meanings of the human world" (2014, p. 17). While I identify my work with this sentiment, the essential subject matter of phenomenology is always lived experience—of which there is very little in this thesis. While I consider the structure of the meanings that emerge during peoples' interactions with robotic artefacts as imperative, my work cannot excavate their essential nature in a manner satisfactory to the phenomenological method. Phenomenological research does not seek to explain the causes of events or behaviours of people. It seeks to understand what a lived experience *is like* for the person or people who live through it.

That HCI and HRI find themselves at the intersection of natural science and human sciences, is reflected in the original conception of this thesis, and this very important distinction between *why* and *how* or *what* took quite a while for me to grasp. The work in this thesis has to a large extent been driven by a wish to explain *why* because I for a long time did not realise that explaining *how* was more than good enough. A phenomenological study springs out from a place of *wonder* at *what* gives itself and *how* that something gives itself (van Manen 2014, p. 27). In my preoccupation with finding the ways in which phenomenology could be theoretically applied to explain *why*, it did not occur to me that it could have been practised for describing *what*. I do not find that I am necessarily successful in my attempts of bracketing—to set aside my natural attitude and my inclination to answer why. Nonetheless, I find that my sense of *wonder* at the phenomenon at hand comes from a place of genuine curiosity toward exactly *what* and *how* the movement of robotic artefacts gives itself during interactions.

> So, the practice of "doing phenomenology" is thinking and seeing our world phenomenologically. And to think phenomenologically is to be swept up in a spell of wonder about the originary [sic] meaningfulness of this or that phenomenon or event as they appear, show, present, or give themselves to us in an experience or consciousness. In the experiential encounter with things and events of the world, phenomenology assists us by directing our gaze toward the regions where understandings, emotions, meanings, and feelings originate, well up, percolate through the porous membranes of past existential sedimentations—then infuse, permeate, infect, touch, stir us, and exercise a formative and affective effect on our being and becoming (van Manen 2023, p 4).

I choose to characterise my study as a mixture of using phenomenological concepts of what experience is toward the aim of understanding what a specific experience is like. The study described in this thesis was not designed as a phenomenological study in the sense described by van Manen (2014), it does not uncover the essential nature of the phenomenon by which I was swept away. Instead, I have delved into phenomenological literature to look for theoretical accounts relevant to the structure of the specific experience I wanted to investigate. Therefore, I would argue that my work carries with it, in a significant manner, the *phenomenological gaze*.

> Phenomenology orients to the meanings that arise in experiences. Any and every possible human experience (event, happening, incident, occurrence, object, relation, situation, thought, feeling, and so on) may become a topic for phenomenological inquiry. What makes phenomenology so fascinating is that any ordinary experience tends to become quite extraordinary when we lift it up from our daily existence and hold it with our *phenomenological gaze*. Wondering about the meaning of a certain moment of our lived life may turn into a phenomenological question: we may then wonder and ask, what is this experience like? (van Manen 2014, p. 38 [emphasis added]).

The phrase describes the manner phenomenology assists in directing the gaze of a researcher who has been *swept up in a spell of wonder* at some experiential encounter. Sheets-Johnstone (2011) describes wonder as "a spontaneous feeling variably weighted with fear and longing" (p. 284). I take this to be a fear of the familiar turning somehow profoundly unfamiliar, and in turn a longing for understanding or grasping the phenomenon one has discovered to be unknown. "When we let ourselves begin wondering […] we give in to our longing to understand something about human life, something we do not understand, something that begs us to pay attention to it, something with which we feel we have to come to terms […]" (Sheets-Johnstone 2011, pp. 288-9). A genuine pursuit of wonder is generally bypassed by the compelling authority of scientific world models. "The result is that the complex *experiential* realities of our everyday lives and of the everyday world are jettisoned in favour of experimental findings and laboratory statistics, computer imaginings and modellings of brains, a bean-bag genetics of traits and behaviours, and so on" (Sheets-Johnstone 2011,

p. 289 [emphasis in original]). Sheets-Johnstone (2011, p. 291) emphasise the crucial distinction between wondering *at* a phenomenon from that of being *moved to wonder*. Indeed, Sheets-Johnstone (2009) affirms the need to verify scientific results "by one's own experience" (p. 509).

I understand the phenomenological gaze to be a part of the phenomenological attitude, but I do not think this means it is limited to it. The hallmark, I should think, of the phenomenological gaze, is being evoked by wonder, making the two intimately intertwined. Van Manen (2014) describes it as our gaze being drawn toward the previously familiar—but now profoundly unfamiliar—phenomenon or experience. In drawing our gaze toward it, we find "something" that gazes back at us. The experience will not maintain this unfamiliar character forever, but by maintaining and holding a phenomenological gaze, one holds this specific experience central to the investigations one undertakes. Being evoked by a profound sense of wonder is a first-person experience through and through. However, this particular experience is not directly subjected to study, it is merely what sparked its pursuit. The phenomenological gaze thus also holds first-person experience central, even if it does not explicate one specific first-person experience in particular.

What sparked a sense of genuine wonder in me—what impelled me "to explore, to investigate, to ponder" (Sheets-Johnstone 2011, p. 291)—was the experience of *seeing* a robot lawnmower as animate, while simultaneously *knowing* it to be inanimate. I gauge elucidating this specific first-person experience in Soma (2020). In this thesis, I thus maintain a phenomenological gaze on this experiential phenomenon as I conduct my theoretical investigation.

4 The prospect of an artificial Umwelt

> I want to know what it is like for a bat to be a bat. Yet if I try to imagine this, I am restricted to the resources of my own mind, and those re-sources are inadequate to the task.
> —*Thomas Nagel, What is it Like to be a Bat?*

4.1 A BIT OF HISTORICAL CONTEXT

Jakob von Uexküll (1864-1944) was an Estonian-born biologist who started his academic career as a student of zoology before further specialising in the muscular physiology of animal locomotion. Through his university education, he was schooled in the mechanistic view that dominated biology and the natural sciences in general. Even so, throughout his university years, Uexküll underwent a change of perspective from a mechanistic view toward a more vitalistic one, involving an increased curiosity toward the connection between physiology and perception, and in turn experience. This change was in part influenced by his reading of Kant, whose philosophy led to an interest in "species-specific" subjective worlds. His ambition as a biologist would be directed toward gaining an "understanding the cognitive modalities that shape animals' species perceived environment" (Brentari 2015, p. 23). Uexküll is of course best known for creating the theory of *Umwelt*. The Umwelt is a *phenomenal reconstruction* of the external environment. As each and every subject constructs their own Umwelt, there must exist as many Umwelten as there are subjects. The construction is a dynamic process where the organism perceives, interprets, and constructs its Umwelt based on continuous feedback actions and perceptions.

The German word *Umwelt*—which is not originally German but introduced into German through a German translation of a Danish poem—is usually translated as "environment" (Sutrop 2001). Through the meaning of the word held in this poem, the vernacular use of the word has established Umwelt as a German version of the French word *milieu*. The Uexküllian conception of this word, however, contains more than this. For Uexküll's conception of what *Umwelt* is, a translation generally considered as more precise is "environment-world."[1] To Uexküll the French *milieu* came with deterministic and law-like implications that something acts upon and shapes the subject because it referred only to the physical context in which a species live (Brentari 2015). In contrast, he himself held that there are no such one-way causalities: "Nobody is the product of their milieu—each is the master of his Umwelt" (von Uexküll 1923; cited in Winthrop-Young 2010, p. 216). In other words, no living being is like a machine that merely responds to the stimuli of their environment. The environment is more than just its physical presence and more than just a place where living beings happen to be. It is "*the intertwining of vital relations with other living beings*" (Brentari, p. 79).

Uexküll introduced the idea of subjectivity into biology by saying—as perhaps the only one of his time—that not only humans but also other animals actually have experiences. At his time, biology attempted to explain "everything in terms of local cause and effect, stimulus and response, the material interaction of connected parts" (Sagan 2010, p. 12). The research being conducted all complied with an anthropocentric worldview, rigorously denying any kind of animal subjectivity. Non-human organisms have been viewed as "aimlessly running machines" that experienced some random effects and in turn sent out "random responses," rather than being subjects who have experiences. Consequently, it posited one objective environment for all life forms and, subsequently, proceeded to analyse animals from the outside in. Researchers were after specifically selected and isolated stimuli, that were measured in mechanical responses (Winthrop-Young 2010, p. 231). In other words, animals were viewed as objects rather than individuals who perceive and act according to their own subjectivity. Beings—as precisely living *beings*, subjectively acting and perceiving—had "been left out in the rush to explain living 'things' (as we sometimes say) as effectively and scientifically as Newton had explained celestial motions by mechanics" (Sagan 2010, p. 12). For Uexküll, such a mechanistic investigation belonged to physiology only. Interactions between living organisms are unlike those between objects, as objects interact with each other only according to physical laws. Biology, he held, needed to integrate the phenomenal experience of animals. The prevailing indifference toward the environment should be replaced by attention to what is important to the animal rather than to what is significant for humans (Winthrop-Young 2010, p. 231). In other words, biology needed to be holistic.

These anti-mechanistic views are also underlying reasons why Uexküll was so fundamentally opposed to Darwinism and his approach to evolutionary theory (Brentari 2015), though it needs to be specified that his anti-Darwinian should not be confused with being anti-evolutionary (Emmeche 2001). Uexküll believed that the "law of natural selection" could not explain the *inner* or *phenomenal* world of animals. As Uexküll sees it, the Darwinian account of evolution tries to systemise nature into law-like structures similar to Newton's Laws of Physics. Consequently, Uexküll believed Darwin did more harm to biology than good. For this reason, Uexküll held Darwinism as unfit to describe a biology where everything was perfectly fitted into each other through a plan of Nature. The accumulating evidence for the existence of DNA that was gradually verified throughout his working career was not viewed by Uexküll as supporting evidence for Darwinian evolution, but rather as bringing about a "new wave of mechanical understanding of living things" (Sagan 2010, p. 17).

Uexküll also underwent another major change throughout the course of his academic career. Due to an unfortunate mix of circumstances, his work gradually transformed from purely empirical to purely theoretical. The theoretical perspectives provided by the Uexküllian Umwelt theory emerged as a synthesis of his previous empirical work and philosophical convictions based on Kantian transcendentalism (Brentari 2015). At its core, whether empirical or theoretical, the research of Uexküll focused on the *phenomenal worlds* of animals, or the worlds around animals as they perceive them and how this perception determines their behaviour. While embedded within the realm of biology, Umwelt theory, as a theory of meaning, expands well beyond the limits we have set for the field of biology today. Umwelt is a theory of *meaningful signs*, and in virtue of this primarily a semiotic theory. However, because Uexküll knew of neither Pierce nor Saussure, his work cannot easily be placed within either school of semiotics. Uexküll's work on understanding the behaviour of animals is considered a central contribution to the foundation of ethology (the science of behavioural psychology), and his theory of meaning as the predecessor of biosemiotics (T. von Uexküll 1982).

4.2 A TERMINOLOGICAL TAPESTRY

The semiotic terminology Uexküll created to describe the relations between an organism and its environment is intricate. One can imagine each concept as threads that, when weaved together, make up an elaborate tapestry. The tapestry impressively illustrates the manner in which organisms are not passive recipients of stimuli, but active participants in their environment; they seek meaningful signs and act on these. To keep the complexity at a bare minimum, I only include and use a small handful of them in this chapter. There are subtle but important nuances that can only be appreciated when considered in unity. However, I will not attempt to explain these concepts and their distinctions here, in part because I find myself ill-equipped to do so in a satisfying manner[2]. Another complicating factor is that Uexküll, like most German-speaking thinkers, invented quite a large number of new words carrying semantic and connotative significances that are not easily and unambiguously translated. Most importantly for the English-speaking reader, various translators will disagree on how to best capture the ideas in their respective translations. "The differences in terminology [caused by different translations], however, should not be seen only as creating difficulties — they also may prove helpful in throwing light upon areas where the various semiotic theories diverge" (T. von Uexküll 1982)[3].

One of the most challenging distinctions in the Uexküllian terminology is the difference between *Umwelt* and *Innenwelt*. Emmeche (2001) explain that when Umwelt and Innenwelt are conceptually distinguished, Umwelt gains a narrower meaning as "the species significant surround" while Innenwelt refers to "an individual organism's actual version of that surround" (pp. 683-4). The Umwelt is an organism's subjective and individual reconstruction of the external environment which is based on its physiology and its possibility for perception and interpretation of meaningful signs. Innenwelt is more like an inner drive—thoughts, feelings, or intentionality if you like—of an organism. The inner world engages the organism to partake in different behaviours. I shall not dwell too much on this distinction as it is rarely used in the literature—presumably because there is no English translation of *Umwelt und Innenwelt der Tiere* (von Uexküll 1921) where this term seems to primarily feature. This makes an already obscure distinction even harder to demarcate.

Making matters no less confusing, each subjective Umwelt has two aspects, or sides, that represent different perspectives on the organism's meaningful interactions with the world. Both sides are equipped with their own fleet of similar, yet carefully thought-out and nuanced terms that are differentiated by a prefix pertaining to their functional role in the behaviour of an organism. The *Merkwelt* is the side of perception—with its *Merkmal* (perceptual marks or cues) and *Merkzeichen* (perceptual sign)—and is shaped by the subjective experience and interpretation of the environment. *Wirkwelt* concerns the actions and behaviours of the organism—with its *Wirkmal* (effector marks or cues) and *Wirkzeichen* (effector sign)—and emerges through an organism's interaction with the environment according to its ability to make changes to or manipulate it (observable behavioural responses).

Merkwelt and Wirkwelt—perceptions and actions, respectively—reciprocally influence each other in a continuous and cyclical manner, what Uexküll calls *functional cycles*. The behaviour of an organism is guided and driven by a process of signs. It involves perceiving cues and meaningful signs, engaging in a behaviour which leaves some kind of effect on the environment, which can again be meaningfully perceived. Functional cycles can be understood as a schematic description of the meaningful relationship between an organism and the environment. The reciprocal influence of Merkwelt and Wirkwelt are the two main components of a functional cycle. All functional cycles of an organism add up to make a full description of an organism's behaviours.

When studying one functional cycle at a time, the focus is on the relationship between a subject and its current object of interest. Objects in the Umwelt might be described as appearing in various kinds of modalities depending not primarily on the properties of the objects, but more on the properties of what we call the sensory organs of organisms. Within the Umwelt, the senses distinctively present themselves as *sensory spheres*. I am fond of Uexküll's illustrative analogy of sensory spheres as garments that "wrap themselves" around the subject who is always perfectly centred within them. The reach of each subject's Umwelt is whatever is "touched" by the garment of the sensory sphere. Umwelt contains at all times everything between the subject and the horizon of each sensory sphere. From substances in direct contact with the skin, tongue, and nose to noises around the corner and, literally, as far as the eye can see.

> This island of the senses, that wraps every man like a garment, we call his Umwelt. It separates into distinct sensory spheres, that become manifest one after the other at the approach of an object. For man, all distant objects are sight objects only, when they come closer they become hearing-objects, then smell-objects and finally touch-objects as well. Finally, objects can be taken into the mouth and be made taste-objects. (von Uexküll 2001, p. 107)

This means that every object in a subject's perceptual world is meaningfully represented in the distinct modality of each sphere for which it is available. For instance, a flame is a see-object and a smell-object—and perhaps even a hear-object—but it is *not* a touch or taste-object.

4.2.1 The procedure of participant observation

By use of examples rich in detail, Uexküll conveyed what it *might be like* to perceive the world as another creature. It is imperative to note that it should never be presumed as an attempt to access the real experience of what it is like to *be* this creature under consideration.

> An organism has only primary access to its own Umwelt, and only humans (and some rather clever 'mind-reading' animals, such as certain predators interpreting the mind of their prey) may by inferences have indirect access to the Umwelt of other species. However, this 'indirect access' is never the same thing as the real Umwelt of the species in question—e.g., our scientific understanding of the sonar system of a bat gives us an indirect and functional picture of the bat's Umwelt, but we cannot enter into that Umwelt itself; all we have is a model in our (linguistic, cognitive, and perceptual) Umwelt of the bat's Umwelt. Science attempts to build a model-based 'view from nowhere' (Nagel 1986) but can only do so mediated by our species-specific Umwelt, our subjective point-of-view from which we collectively construct a shared human sphere of public knowledge. (Emmeche 2001, p. 656)

At its core, the idea is rather that we might, by imagining its perception of the environment based on its physiology, gain a better understanding of the perspective of the creature. When we can identify what the *meaningful objects* in a species-specific environment are, it becomes possible to construct an accurate description of its lifestyle on the premises of the organism, and not from a human perspective. The magic, so to speak, of Uexküll's observations, is how they eloquently illustrate how the specific organism has adapted their Umwelt to serve them well in their way of life. The most famous may be that of the tick, where an impoverished Umwelt is rightfully revealed as a strength rather than a weakness, as would be the human prejudice. Exactly this ability is what Sagan (2015) refers to when he characterises Uexküll's descriptions as shamanic. The procedure is called *participant observation*. The following paragraph is a recount of Uexküll's participant observation of the fully developed and copulated female tick. He describes in detail how "she" will climb onto a branch and stay there—waiting for up to eighteen years—until a warm-blooded, mammalian animal passes beneath her:

> The eyeless tick is directed to this watchtower by a general photosensitivity of her skin. The approaching prey is revealed to the blind and deaf highway woman by her sense of smell. The odour of butyric acid, that emanates from the skin glands of all mammals, acts on the tick as a signal to leave her watchtower and hurl herself downwards (von Uexküll 1992, p. 321).

If the tick fails to land on something warm, she knows she missed her prey and will again climb upwards into her self-fashioned "watchtower." If she feels warmth as she lands, she knows she succeeded and will burrow through fur until she finds a membrane she can bite through. She now eats what will be her last meal; when she is finished, she will fall to the ground to lie eggs and die. This is the life cycle of the tick, made possible by what is essentially the three perceptual cues and three effector cues that make up her entire Umwelt. The first perceptual cue lets her become aware of a mammalian presence, which "coincidently" is the only molecule all mammals have in common: butyric acid. The second cue tells her what she needs to know about the temperature of her landing spot, a cue that informs her whether she hit her target or not—her desired destination being on a surface that keeps the temperature of a mammalian animal. The third and final cue is also tied to temperature: instead of a sense of taste, she is sensitive to the temperature of the liquid she starts consuming as her hypostome penetrates the membrane of the skin. Her physiology restricts the available perceptual cues. She does not need a sense of taste; in nature, the fulfilment of all these three cues means that she is eating mammalian blood.

Uexküll captures the significance of participant observation by sharing his insights such as this: "out of the vast world which surrounds the tick, three stimuli shine forth from the dark like beacons and serve unerringly to her goal" (von Uexküll 1992, p. 325). In Uexküll's description of the Umwelt of the tick, it becomes apparent that it is the poverty of her world that ensures her success in surviving and reproducing because security is more important than wealth. In this way, her physiology is perfectly fitted to her environment through her functional cycles. Using participant observation to analyse the Umwelt of an organism is intended to provide insight into its lifestyle and subsequently behaviour. When equipped with knowledge about the behaviour and physiology of an organism—giving insight into its Merkmal and Wirkmal—it becomes possible to describe the functional cycles. Step by step, the significance each meaningful object has for a behavioural trait will construct a complete illustration of the Umwelt (at least as complete as it can be for the participant observer).

4.3 WHAT A ROBOTIC ARTEFACT KNOWS ABOUT THE WORLD

Umwelt theory will at first sight—in being primarily a biological theory of how living beings are inseparable from their environments—seem far removed from HCI and HRI. Uexküll's primary contribution is the exact opposite of portraying machines as having a subjective point of view. It emphasises the fact that organisms are not machines or machine-like in their way of responding to the world they live in. Given this, it might appear counterintuitive to apply these subject-oriented, anti-mechanistic, biological theories for the analysis of robotic artefacts (Emmeche 2001). After all, Uexküllian theories arose out of a disagreement of the prevailing mechanistic view on biology and the entire concept of Umwelt revolves around both the individuality and subjectivity of organisms.

Not surprisingly, I did not come up with the thought of applying Umwelt theory to robotic artefacts myself[4]. Umwelt theory is well established in some research communities with a particular interest in the nature of symbol grounding in artificial systems and in turn the prospect of autonomy of such artificial systems. This reveals that Umwelt theory—despite its anti-mechanistic mantra—is compatible with research on robotic artefacts and scientific fields related to both technical and conceptual aspects of robotic artefacts. Indeed, it has proven useful to highlight issues of symbol grounding and situatedness in regard to the autonomy of the system (e.g., Ziemke and Sharkey 2001). It is not the symbol handling that gives the Umwelt its phenomenal character, but that organisms are "constituted as an active subject with some agency." (Emmeche 2001, p. 678). The manner in which I suggest using Umwelt theory for gaining insight into a robotic perspective, through what I propose to call *artificial Umwelt*, is compatible with this point. I do not suggest that robotic artefacts have a phenomenal experience of their environment, but that the technique of Uexküll's participant observation can be used to provide a better understanding of the actual capabilities of the system. The way in which the objects in the world appear (naturally, not *phenomenally*) will matter for interaction.

Aspects from Umwelt theory were, to my knowledge, first used in the context of robotics by Brooks (1991). While Umwelt is never mentioned, Brooks is confident in his assertion that different kinds of robotic artefacts will have their own distinctive Merkwelten. The sensors of the robotic artefacts will determine the modality objects in the world will take in their Merkwelt. Brooks does not express any opinion on what the structure of this Merkwelt will be like except for making the observation that the robotic Merkwelt will most likely be strikingly different from our human Merkwelt. He argues against the then prevailing approach in AI research, saying that the superior approach is letting the system sort out the categories of the environment themselves. Because the Merkwelt of the robotic artefact most likely will be highly dissimilar from the human Merkwelt, it is counterproductive to impose the categories that are meaningful in our Merkwelt on a system that will most likely not find any meaning in these.

The question of whether or not robotic artefacts—or other artificial systems—can or cannot *actually* have an Umwelt in the same manner as living beings inevitably arises when thinking along these lines. The subject is of course fascinating, but I wish to avoid an extensive discussion on the topic in this chapter. A thorough examination of this subject can be found in Emmeche (2001). While it does not provide any definitive conclusion, the author argues convincingly toward its unlikeliness, seeing as artificially constructed systems are unable to realise genuine semiosis. I refrain from having any strong opinion on the topic except from admitting to finding that Emmeche's (2001) argumentation resonates well with me[5]. Regardless of what the answer might "truly" be, Emmeche (2001) highlights an essential detail in that one can acknowledge a robot's instantiations of functional circles without these having to be "a *true* instance of a functional circle" (p. 678 [*emphasis added*]).

According to Uexküll, when faced with the challenge of understanding animal environments, especially the most elementary, the researcher must adopt a research method laid out in two phases. First, for every species studied, one must identify the objects which have a role in its environment-world; in order to do so he must start with the environment that is most accessible to man (a large part of which overlaps with animal environments – a point we will look into in greater detail) and proceed with the elimination, the "cancelling" of all objects and object properties that have no relevance for the species being studied. Second, one must identify the sensorial stimuli that, after having originated from the remaining significant objects, reach the nervous system of the animal being studied. Upon completion of these two passages, the map of that species-specific environment should emerge.

4.3.1 The artificial Umwelt of a robot lawnmower

Acknowledging the robotic artefact's metaphorical functional cycles is exactly what we do in Soma and Herstad (2018) when we identify a robot lawnmower's meaningful objects and use this to describe its behavioural traits (also metaphorically, of course). The text in this section is a recitation of the analysis made in this paper, reworked to fit better with a revived understanding of Umwelt theory, and the use of the term *artificial Umwelt*.

Just as every animal will be surrounded by different things—a dog by dog things, a human by human-things, and a tick is surrounded by tick-things—a robot will be surrounded by robot-things. More specifically, every variety of robot models will be surrounded by model-specific things. So, when we apply Umwelt theory for the purpose of analysing the artificial Umwelt of a robotic artefact, we must raise the question of what the model-specific robot-things surrounding the robotic artefact in question are. In the example used in Soma and Herstad (2018), we exemplify this procedure through the robot lawnmower Husqvarna auto mower 308, named Roberto by its owner. We start by making a scheme of Roberto's sensory spheres.

In order not to destroy itself, flowers, or otherwise cutting objects that are not lawn, the setup of the robot artefact involves installing its boundary wire. This electronic "fence" is, to a human, nothing more than a closed-circuit cable laid around the edges of the lawn from the charging station. When active, the robot roams freely and blindly in a randomized, irregular pattern within the boundary wire. It does not stop until it meets its fence, or unless its crash-sensor is triggered, informing it about an obstacle. About itself, Roberto knows when it is low on power, and when it is fully charged. It has been equipped with a clock and can thus be given a schedule for when it should be moving across the lawn and when it should be docked. Leading from the charging station and outwards in a straight line is a cable that makes out the robot's guide wire, which the robot uses to navigate to and from its charging station. If it is lifted, the knives will immediately stop spinning, and will not start again until the robot puts the right side down, the knives safely facing the ground.

After outlining Roberto's basic functionality, we can start to imagine the Roberto-specific objects of its artificial Umwelt and describe the significance of its task performance. While the guide wire is always physically present, it will be ignored unless Roberto is low on power, or its schedule indicates that work is done for the day. In both cases, it requires returning to its charging station. Thus, there is a guide wire-object only when the robot is on its way to or from the charging station. Similarly, the robot is not aware of the boundary wire when working and roaming. Only when it approaches the boundary wire does there appear a boundary wire-object in Roberto's Umwelt, and it can cross the boundary with a few centimetres before turning around. This makes the perceptual cue of the boundary wire different from that of a crash. Whenever the robot meets an obstacle, it must physically crash into it to detect it. Once it does, it stops immediately, turns, and goes in another direction. Roberto does not plan a trajectory to avoid physical objects because, to it, there are no such things. In a way, we can then say that every object in the lawnmower's Umwelt is a crash-object, and unless it crashes, the Umwelt is devoid of any objects. Roberto's sensory spheres are confined to the boundaries of its plastic shell.

The Umwelt of the robot lawnmower is simple; until it crashes, finds itself close to its boundaries, or is low on power, it works in complete isolation from the environment. This illustration is seen in Figure 4.1 below. The functional cycle initiated by the perceptual cue of a crash-object has only one effector mark. It stops immediately, before turning and going another way. The environment as one could imagine it might appear to Roberto consists of three meaningful objects and five functional cycles.

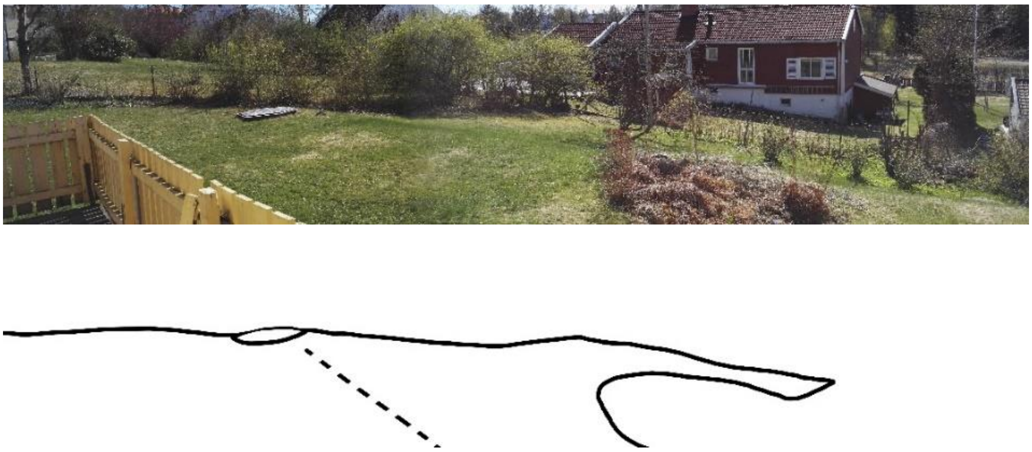

Figure 4.1: A garden according to the human Umwelt (top), the same garden according to Roberto's Umwelt (bottom) where the thin line represents the boundary wire and the dotted line the guide wire (Soma and Herstad 2018, p. 60).

Unless Roberto's crash sensor is triggered and the world is devoid of anything, the robot will continue to mow, even if it runs over objects on the ground. This can cause trouble. During late summer, apples from the garden's apple tree fall to the ground. In Roberto's artificial Umwelt, these apples do not exist because they cannot be detected by its sensors. Instead of avoiding the apples, Roberto runs over them. Its knives become stuck so it can no longer move and turn off. To get out of this situation and back to work, the lawnmower requires human assistance. Further, the garden has a small hill that is too steep for Roberto on rainy days, or when there is a lot of morning dew. Seeing that Roberto's artificial Umwelt does not contain grass-objects, it can naturally enough not distinguish between a dry or wet lawn. Should the robot attempt to climb this small hill when the grass is wet, its motors are not powerful enough, and the lawnmower becomes stuck. Again, only human assistance can get the lawnmower back to work.

One could be tempted to think that, like in the example of the tick, "security" is more important than wealth, and that the poverty of Roberto's artificial Umwelt ensures it will not become stuck, destroy things, or hurt anyone. Herein lies the key distinction between the Umwelt of living beings and the artificial Umwelt. Unlike living beings, lawnmower robots such as Roberto are not fitted into their area of deployment—its environment—with the equal completeness that animals are fitted into their environments. The artificial Umwelt will not necessarily contain all the meaningful objects needed for the robot's independent task performance. In turn, it might be presumed that the robotic artefact will face difficulties in handling the ever-changing and dynamic environment.

5 An ecological perspective on meaning

> No one ever had a simple sensation by itself.
> — *William James, The Principles of Psychology*

5.1 THE ECOLOGICAL APPROACH

In its total rejection of the stimulus-response formula, ecological psychology became a wholly new approach to the field of psychology. The ecological approach to visual perception, and in turn ecological psychology as a scientific movement, arose out of American psychologist James J. Gibson's dissatisfaction with existing theories of perception, stemming from frustration over how perceiving was considered separate from acting when he saw them as being inseparable.

> "I used to suppose that the way to understand [vision] was to learn what is accepted as true about the physics of light and the retinal image, to master the anatomy and physiology of the eye and the brain, and then to put it together into a theory of perception that could be tested with experiments. But the more I learned about physics, optics, anatomy, and visual physiology, the deeper the puzzles got. [...] Physics, optics, anatomy, and physiology describe facts, but not facts at a level appropriate for the study of perception" (Gibson [1979] 2015, p. xi).

The ecological approach to visual perception was a response to conventional psychology. Traditionally, research on perception had a sole focus on pure stimuli and isolated sensations. In Gibson's opinion, this was the wrong way to go about properly understanding perception, doubting that "vision is simplest when the experimenter tries to make the eye work as if it were a photographic camera" (Gibson [1979] 2015, p. xiv). Thus, the inheritance of the Jamesian functionalist tradition can be clearly seen in Gibson's ecological psychology: "We are told that vision depends on the eye, which is connected to the brain. I shall suggest that natural vision depends on the eyes in the head on a body supported by the ground, the brain being only the central organ of a complete visual system" (Gibson [1979] 2015, p. xiii). Like functionalism, Gibson's ecological psychology stands in contrast to the view of perception as simple sensations, where the meaning of the world is only made available through cognitive effort. The functionalist founder, William James, doubted there could even be such a thing as a simple sensation, standing alone and separated from everything: "What we call simple sensations are results of discriminative attention, pushed often to a very high degree" (James 1890, p. 224, as cited in Käufer and Chemero 2015, p. 146).

The functionalist movement emerged at the turn of the 20th century (Käufer & Chemero 2015, p. 145), around the time psychology became a field of science separated from philosophy but had yet to be fully established—more or less coinciding with the time Wundt established his approach to psychology and science. James was not convinced that isolating the senses was the best approach to understanding them. James was heavily influenced by Darwin and evolutionary theory. He believed aspects of our mental life were best understood in terms of *what they were for*—their function—and that like physiological adaptations of "the body" to the physical environment, aspects of the mind could also be understood in evolutionary terms, as *adaptations* to the environment (Käufer & Chemero 2015, p. 146-7). In turn, James' work and thinking combined the disciplines of physiology, psychology, and philosophy, while also containing elements of pragmatism and phenomenology (Goodman 2017).

James rejected the Kantian distinction between the empirical world that we experience and the world-in-itself. Likewise, Gibson was convinced that the world must be experienced *directly*. For Gibson, perception and action should not be viewed as two separate processes of a body and a mind, joined somehow by the brain. Perceptions are activities with the function of an intentional animal for the purpose of detecting information specific to itself and to the environment. Perceiving is an activity, not a set of passive bodily sensations. For instance, seeing is not best described as photons hitting the retina, which triggers sensory cells in a certain pattern which is sent to the brain for processing. Seeing is not merely "sensing light" and "cognising signals" but a *perceptual act* using the eyes-in-the-head-on-the-body-resting-on-the-ground (Gibson 2015 p. 195). Likewise, hearing is equally active as one uses ears-on-the-head-on-the-body-on-the-ground. Similarly, the animal, as an active observer, obtains the information by seeking it and "picking it up" using all of their perceptual systems simultaneously. In other words, the entire animal is active in the pursuit of information.

To complement his unique ideas about the nature of perception as direct, Gibson also presented a novel understanding of the environment to goes hand in hand with it. In fact, nearly half of *The Ecological Approach to Visual Perception* (originally published in 1979) is explicitly devoted to outlining how the environment is structured, and how the information about this structure is both directly available and meaningful to organisms (Bruineberg, Chemero, and Rietveld 2019, p. 5235). When Gibson uses the word information, it means something quite different from the traditional understanding of "knowledge communicated to a receiver" (Gibson [1979] 2015, p. 231).

Communicated information is encoded and limited to the channels used for that specific case of communication. In contrast, *ecological information* is abundant, and it is everywhere. The entire environment is fully saturated with optical, mechanical, or chemical information. Ecological information is not encoded but *structured* by the properties of the environment and made available for direct perception as *opportunities for behaviour*. These opportunities for behaviour are what Gibson called *affordances*. Affordances "point in two ways, to the environment and to the observer" (Gibson [1979] 2015, p. 132). In so doing, they exist as meaning conjoining a subject and their environment. Affordances eradicate the need for mental representation and mental construction of meaning in theories of perception because meaning instead exists as the relationship between organisms and their environment in what Gibson called *animal—environment mutuality.*

Gibson was—like his contemporary Merleau-Pontly—influenced by Gestalt psychology where the contents of the world are thought to be perceivable as wholes and not the sum of many different stimuli. Gibson's concept of *affordance* is according to himself ([1979] 2015, p. 130) directly derived from the valences and invitations of Gestalt theory, but with a substantially different apprehension of how meaning emerges between organisms and the environment. Unlike Gestalt psychology, Gibson held that meaning is neither physical nor phenomenal. He describes his conception of meaning as "a meaning of a new sort" ([1979] 2015, p. 130) where value is not bestowed upon objects by the need of an observer. Affordances are objective in that their existence does not depend on value, meaning, or interpretation. Yet they are subjective in that an organism is needed as a frame of reference. The organism does not find the affordance by means of a set of values that allow it to perceive them. Yet affordances—as some entity of meaning—can only exist within the frame of reference provided by the capabilities of the subject. In being a property belonging to neither of the two alone, as the coupling in the animal—environment mutuality, affordances traverse the barrier between subject and object. In Gibson's view, the ecological approach is not just an attempt at, but successful in overcoming the age-old body—mind dichotomy. Even if Gibson's ideas did not follow directly from a phenomenological tradition, his "new framework for understanding perception from the ground up" (Käufer and Chermero 2015, p. 145) shed new light on questions concerning subject—object identity.

5.2 THE ROLE OF AFFORDANCES IN DESIGN

Affordances are without doubt the most influential concept from Gibson's work outside of ecological psychology. The term has gained an especially central role in HCI. As one can presume everybody oriented in the field of HCI is familiar with, the concept of affordances was introduced by Don Norman in 1988 through *Psychology* (later *Design*) *of Everyday Things*, which quickly became wildly popular. However, due to some "inherent ambiguities" in Norman's definition of affordances, the usage of the term in

HCI literature is "widely varying" (McGrenere and Ho 2000). Even though these ambiguities have later been commented on and sorted out by Norman, the most critical discrepancy between the original source and the adaption remains. The role of affordances in ecological psychology enables Gibson's theory of perception to fully disassociate itself from mental representations. Norman, being a representationalist following the Wundtian tradition, sees affordances interpretations of an object's perceived features. In the original print of the book, Norman comments on this topic in a footnote:

> I believe that affordances result from the mental interpretation of things, based on our past knowledge and experience applied to our perception of the things about us. My view is somewhat in conflict with the views of many Gibsonian psychologists, but this initial debate within modern psychology is of little relevance here (Norman 1988, p. 219).

While I agree that the debate might be out of scope for Norman's book, it is hardly irrelevant to it. For HCI, the question of the ontological nature of affordances will perhaps seem trivial, as the field does not primarily deal with answering questions about the nature of perception or philosophy of mind. However, affordances, as they are conceptualised by Gibson, are completely incompatible with a representationalist view on perception. In turn, what is conceptualised as being meaningful during interaction will be fundamentally different.

I will approach this topic by comparing Gibsonian affordances with Norman's HCI-specific adaption of them. To do this, I lean on the summary of McGrenere and Ho (2000, *from Table 1*, p. 3). The summary clearly shows that these differences are a result of distinct views on the ontological nature of affordances.

Gibsonian affordances	*Norman's affordances*
(G1) Offerings or action possibilities in the environment in relation to the action capabilities of an actor. **(G2)** Independent of the actor's experience, knowledge, culture, or ability to perceive. **(G3)** Existence is binary – an affordance exists, or it does not exist.	**(N1)** Perceived properties that may or may not actually exist. **(N2)** Suggestions or clues as to how to use the properties. **(N3)** Can be dependent on the experience, knowledge, or culture of the actor. **(N4)** Can make an action difficult or easy.

Table 5.1: *The contents are retrieved from* Table 1 *in McGrenere and Ho (2000, p. 3). The formatting has been slightly adapted to aid the discussion below by adding a lettered numbering.*

When comparing (G3) and (N1) from the table above, the diverging views on the nature of affordances seem pretty straightforward. However, there is more to it than just a difference of opinion on whether affordances are of a binary or fluctuant existence. In fact, there is some disagreement even amongst Gibson's followers on how exactly affordances are to be understood (Anderson & Chemero 2003). Primarily, the difference of opinion is entrenched in some unresolved metaphysics of affordances. And primarily, it is an issue for those interested in resolving the challenges this poses to the validity of ecological psychology as a complete theory of perception. Yet, whether the nature of affordances is binary or fluctuant is not irrelevant to HCI and design because the answer gives us direction when we during design take into consideration what people bring with them into an interaction; how people perceive affordances and subsequently how they make sense of the available opportunities for behaviour.

An important aspect of Norman's affordances is their visibility, and when it comes to design, one should consider how past experience, culture, and context will affect the "perceivability" of an affordance—what is *considered* to be possible. As (N1) points to, affordances are perceived properties that may or may not exist. Norman specifies that "an affordance is jointly determined by the qualities of the object and the abilities of the agent that is interacting" (Norman 2013, p. 27), and does indeed stress the relational nature of affordances is central to the concept. Whether or not a behaviour is supported is almost completely contingent on what is brought into the interaction by the subject in terms of experience and culture. Even if Norman's affordances are relational, they become more like a specification of the properties of objects where meaning *can* be discerned. Gibsonian affordances form the direct connection between an organism and the environment. They are the already meaningfully structured ecological information that is available for being "picked up" by the animal. Norman's affordances arise in or *by* the act of perception. In fact, they must, because they cannot be structured by other means than representations inside the mind whereupon they are given their meaning. This view on affordances entails the perception of properties which can be constructed inside the mind and interpreted *as* opportunities for behaviour, rather than being opportunities for behaviour specified *by* affordances.

Gibsonian affordances are present no matter what, and they are not present for some specific purpose: "The affordances of the environment are what it offers the animal, what it provides or furnishes, either for *good* or *ill*" (Gibson [1979] 2015, p. 119). A body of water with the right properties, in relation to the properties of the organism, will afford both drinking from and drowning in, even if the organism is not thirsty when it sees it, or whether or not it realises drowning is a possibility. In contrast, a central aspect of Norman's affordances is as structures dependent on the experience and culture of the perceiver (N3). And, as revealed by (N4), they gain a different role during interactions in determining what the experience of use will be. Gibsonian affordances can exist independently of perception because they are not, as (G2) states, *value laden*. But, as McGreenere and Ho (2000) identify, these value-free affordances do not fit well with respect to design.

> Recall the example of a stair being climbable or non-climbable by a particular individual. Reality obviously isn't this black and white; a gray area exists that is meaningful to the stair

> climber. For a particular individual one stair may be climbable with great difficulty whereas a different stair may be climbable with ease. Gibson doesn't address this range; they are both climbable and thus they both qualify as affordances. From a design perspective, an affordance that is extremely difficult to undertake versus one that is undertaken with ease can hardly be put in the same category. In the design of everyday things, the goal should be to design information that uniquely specifies an affordance and also to design useful affordances that can be undertaken with ease. (McGreenere and Ho 2000, pp. 3-4)

This quote illustrates well that there are indeed more nuances to meaningful behaviour than what one physically can or cannot do. According to (G3), the binary nature of Gibsonian affordances provides no such nuance. Previous experience and knowledge of the subject may not affect the *existence* of Gibsonian affordances. However, it is important not to confuse value with meaning. That affordances form the meaningful connection of the animal—environment mutuality, must mean that meaningfulness unfolds dynamically in all situations involving a subject. "Even when we are ready to act on an affordance, we are prepared for something that we could do, but what we could do is not yet done, so in a sense, something is not yet there. There is no light bouncing off the future. This implies that there is something necessarily anticipatory or future-oriented in the perception of an affordance" (Bruineberg et al. 2019, p. 5244). Perceiving affordances necessarily come with some conception of the outcome of acting upon them. This conception does not require any form of mental models or representations to exist but exists in the subjectively felt knowledge about oneself in the world. In concert with intentionality, or directedness, meaningful gradients emerge continuously as a result of the constant evaluation made by subjects during their interactions with their environments.

As McGreenere and Ho (2000) emphasise, we should keep in mind that Gibsonian and Norman's affordances are meant to achieve different things both theoretically and practically. Gibson builds his entire explanation of perception in animals around affordances, making them by far the most important and central concept to the ecological approach and its theory of direct perception. The role of affordances in Norman's terminology is far less central. Though it gained a lot of popularity in the design communities, it is only one concept of many presented. Affordances are not central to, merely part of, a larger conceptual apparatus for clarifying the significance certain properties or features of the designed artefact will have during an interaction. Affordances do not discriminate based on origin. Even if an object arose out of an elaborate design process, its affordances invite and constrain use in ways that do not necessarily correspond to its intended function (Costall 1995). I argue that because affordances are not *features* of artefacts, one cannot *design* affordances. One can, however, design an artefact so that its features specify and allow certain opportunities for behaviour. Because affordances *are* the meaningful connection between subjects and objects, their presence in designed artefacts cannot function as *instructing* use. The designer will never be in full control over which affordances will be specified making unintended affordances just as likely to be perceived as those intended.

5.3 MEANINGFUL EVALUATIONS

Gibsonian affordances are, like the information that specifies them, abundant. As a natural consequence, most affordances are never actualised. This gives rise to a conundrum concerning the myriad of unrealised opportunities for behaviour. As Stoffregen (2003) points out, “the number of actions that are available to a given animal in a given situation is unlimited (this does not mean that all actions are possible, only that the number of possible actions is uncountably large)” (p. 119). But as he also points out, the mere availability of an affordance “does not lead to the involuntary actualization of the action afforded. Affordances are what one *can* do, not what one *must* do” (Stoffregen 2003, p. 119). Any object can be used in limitless ways (Costall 1995). By this follows the matter of course that it would not be physically possible for any individual to engage with all affordances at once. Some affordances are mutually exclusive, others go by unnoticed, while a great majority are simply irrelevant. As a relationship between organism and environment, affordances are not just meaningful when a supported behaviour is actualised. Meaningfulness also unfolds as an organism’s evaluation of the situation and whether or not a supported behaviour is worth the trouble should it for instance appear possible, but difficult or dangerous, to actualise. In this regard, *not* doing something is also experienced as meaningful.

An example of one such evaluation provided by Gibson himself is the visual cliff. In the experiment designed and conducted by his wife Elanor J. Gibson (Gibson and Walk 1960), infants were encouraged by their mothers to crawl over a glass-covered ledge. The experiment of the visual cliff was not designed to determine whether or not infants were fooled by fake ledges, but to investigate the relationship between learning to crawl and the development of visual depth perception. Nearly all of the infant participants refused to cross the glass, even if they had patted the glass surface. The visual appearance of a drop—affording “falling”—was too scary, even if the tactile affordance of “support” was available for those exploring the apparent drop with their hands. Whether the infants are subject to misinformation or simply fail to pick up all the available information due to their immaturity as perceivers is not for me to say. Even if the infants showed clear signs of wishing to reach their mothers, most of them indeed determined that crossing the glass was not safe. This manner of determining what constitutes the best option on how to proceed—even with an immature skillset—examples a *meaningful evaluation*.

Just like learning to safely navigate the physical environment, learning to navigate a socio-cultural landscape is an acquired skill. More importantly, the two are not separated. In the visual cliff experiments, the children do not just evaluate their physical safety. The intimidating visual drop (affording falling and getting hurt) is balanced against the need to be close to their mother. “Many of the infants crawled away from the mother when she called to them from the cliff side; others cried when she stood there because they could not come to her without crossing an apparent chasm” (p. 67). Indeed, the infants are subjected to a peculiar social interaction with their mothers during this experiment, being encouraged to venture outside the ledge. Children rely heavily on the guidance of adults in learning to master the skill of evaluating the available opportunities for behaviour. As Costall (1995) points out, a child “is not simply left to ‘discover’ the function of a cup or a spoon; rather, the learning situation requires careful instruction by the parent” (p. 472). We are introduced to the functionality of artefacts and objects, and these are again experienced in relation to the community one is currently situated in. “Yet, learning about affordances does not simply concern the uses of an object *happens* to afford, but what it is *meant* to afford. Objects can have their proper or ‘preferred’ affordances. They *can* be used in other ways, but even when these alternative uses occur to us, there may be sanctions against such deviation.” (Costall 1995, p. 472).

When Gibson outlined his theories of affordances, he deliberately avoided mentioning sociality which has led to some still-remaining unclarities about the nature of their existence (Costall 1995; Michaels 2003; Sanders 1997; Stoffregen 2000, 2003). However, because human behaviour cannot be severed from human culture, the notion that affordances are only specified by law-based (ecological) information comes to short. Indeed, humans find the socially layered environment meaningful despite social information not being a lawful environmental feature. There is therefore a general consensus within the ecological psychology community that affordances should be able to handle both sociality and our socio-material environment (Bruineberg et al. 2019; Costall 1995). For instance, you cannot eat the food in a grocery store before you have purchased it. The food’s ecological affordances support the behaviour of eating, but the information specified by their sociomaterial affordances offers buying (or stealing). In other words, both ecological and sociomaterial affordances are taken into consideration when humans make meaningful evaluations. Indeed, as Bruinberg et al. (2019) explain, the “overwhelming majority of affordances in human social relations are not lawfully specified by the energy in the environment but are determined in part by socio-cultural practices, such as conventions and customs, or other regularities” (p. 5236). Accordingly, they demonstrate that even though Gibsionan affordances are value-free, they are still very much applicable to the human social environment. In fact, according to them, ecological psychology is able to deal with “the full range of human social activities” (Briuneberg et al., 2019, p. 5232). Affordances are compatible with sociomaterial activities like creativity, long-term planning and imagination, and not just simple sensorimotor coordination and scale: “It is a matter of embodiment, skill, and power” (Costall 1995, p. 476). Thus, Briuneberg et al., (2019) put forth a new and wider interpretation of Gibsonian information that makes room for prior knowledge such as skills or culture, also for Gibsonian affordances.

> The individual's skills (most of which are acquired via a process of education of attention in sociocultural practices) provide access to the regularities of the world; some skills are primarily sensorimotor, such as grasping a cup, others are typically characterized as more abstract skills (e.g. imagination, but also appropriately grasping your own coffee cup rather than someone else's from those on the table in front of you). Part of being skilled is knowing how to attune to the relevant pieces of information; i.e. coordinate with the relevant aspects of the environment. (Bruineberg et al. 2019, pp. 5244-5)

Rather than specifying which opportunities for behaviour that are or are not available, *the skilled intentionality framework* emphasises the evaluation an active observer undertakes as they *consider* which opportunities to be both physically and socially appropriate. Thus, skilled intentionality lies at the base of how we conduct our meaningful evaluations. This ongoing process involves both our understanding of our own capacities, as well as a recognition of the resources available in the current situation. Throughout our lives, we become experts at sorting between the matters in need of attention, and those we can allow ourselves to ignore. Using our skilled intentionality, we constantly evaluate what we find meaningful to do in a situation, navigating between both what is physically possible and socially appropriate. Through our skilled intentionality, we direct ourselves and our attention toward the relevant affordances. Possible, but culturally unacceptable behaviours are unconsciously filtered out, in the same way that ecological affordances are constantly ignored when they do not align with your intentions. For instance, the more or less constantly available *jump*-affordance offered by the ground is seemingly ignored by most people, most of the time.

Our skilled intentionality also enables us to discern affordances that are not directly available in the environment but will be sequentially revealed over time. Previous knowledge about behaviours and their outcomes makes affordances easier to navigate. What (Gaver 1991) terms *sequential* and *nested affordances* denote the unfolding structure of complex behaviours; doing one thing leads to new opportunities for behaviour. "Sequential affordances explain how affordances can be revealed over time nested affordances describe affordances that are grouped in space" (Gaver 1991, p. 82). What becomes clear about our skilled intentionality in the light of sequential affordances is how it enables us actively to seek out affordances that will *become* available through preceding actions. They also relate to what Bruineberg et al. (2019) refers to as coinciding aspects of the environment. These are things and events in the environment that often occur together, and therefore usually imply each other.

> General ecological information pertains to the ways in which aspects of the environment tend to occur together, like smoke and fire, an object and a shadow, or a pub and beer. […] This example of the bird and its shadow also shows that the case of such general ecological information—due to the regularities in our ecological niche—is such that an aspect of the environment constrains (but does not necessarily specify lawfully) another aspect of the environment. (Bruineberg et al. 2019, p. 5237).

For instance, a refrigerator is a cold storage with a door and an attached handle. Because refrigerator doors are rarely transparent, the initiation of such a sequence relies on the person's skilled intentionality in addition to sensorimotor capabilities. The visual and tactile information of the handle is first and foremost the ecologically available information that offers the behaviour of holding. The door itself offers being opened. In other words, the door and handle together are nested into the affordance of the opening. Once the refrigerator is opened, a new set of affordances appears. All items inside will likely afford to be lifted and taken out of the cold storage. Some items are inedible without cooking and upon seeing these food items, a specific set of sequential affordances opens up. Other items afford to be consumed immediately. If the refrigerator is shared by multiple people, some items will probably be socially unacceptable to consume for anyone else but their owner. The nutrients themselves—the food and drink—are often stored inside containers. Therefore, what you perceive when you see *milk*, is usually a carton of milk. The behaviour of drinking milk is not directly afforded by the carton. The carton only affords *storing* milk (which is not something the consumer is usually responsible for). Yet, upon seeing a carton of milk, it somehow affords *drinking* milk. It does so because we, through our skilled intentionality, are able to imagine acting on the sequential affordances that lead to the milk being consumed, even if it is a process of intermediate steps.

6 Animate movement

> Our engineers have built spacecraft that can land on comets, and our computers have beaten grand masters at chess, but we have yet to see a robot whose movements come even close to the elegance, ease, and flexibility of human walking and running.
> — *Matt Wilkinson, Restless Creatures*

6.1 ON THE SIGNIFICANCE OF MOVEMENT

As the title predicts, the primordial role of movement for life takes centre stage throughout the collected essays in *The Primacy of Movement* (2011)[1]. The aim of the book is therefore best summarised as giving a phenomenological account of movement as the experiential dimension of being. In particular, it is the significance of animation—the movement of living beings—that receives the most attention. Central to the collected essays in *The Primacy of Movement*, but also Sheets-Johnstone's phenomenology in general, is a fundamentally holistic view of the manner in which the experience of *being* is deeply rooted in natural history. With a background in dance, philosopher Maxine Sheets-Johnstone focuses her attention on what she holds to be the central role of movement to life itself.

Having also studied evolutionary biology, her philosophical endeavours are interdisciplinary, described by Gunn (1998) as a "phenomenological-biological-evolutionary" approach. According to her theories, *all* aspects of what it means to *be* (*alive*) are inherently grounded in evolutionary history. In other words, what it means for me, or you (or any other organism) to be alive right here and now—the full extent of being a subject, and specifically being the subject one is—cannot be severed from our natural history. In an assured manner, she tirelessly put forward arguments that support her principal premise that movement, or *animation*, is the primary mode of being for all living things: the phenomenon of movement appeared in living beings long before any mental phenomena. She is fundamentally dedicated to imparting the phenomenological elucidation of how we[2], living beings *are* bodies, rather than the prevalent proclamation that we *have* bodies. These elucidations are consistently made in a manner carefully heeding a vocabulary that focuses on the holistic nature of being, achieved through the use of meticulously precise terminology.

Being a phenomenologist, Sheets-Johnstone is naturally concerned with the nature of experience, or what she often refers to as "actual experience." Therefore, the core of what Sheets-Johnstone sets out to elucidate in her phenomenology are inquiries into the very nature of life and being itself. Her writing revolves around *the structure of consciousness*, *conscious being*, *the nature of experience*, the question of *the relationship between body and mind*, and how everything living is grounded in the *dynamic nature of movement*. As part of her evolutionary perspective, she regards consciousness and thinking to be grounded in the creaturely form of living beings, rather than being centred "inside" the brain. Her thinking is in direct opposition to dualistic thinking and objectification of movement, especially the phenomenal experience of *self*-movement. The insight brought to us by *The Primacy of Movement* is that it was in response to a dynamic world that perception, and in turn consciousness, evolved, not the other way around.

6.1.1 The corporeal consciousness of animate forms

Sheets-Johnstone consistently uses the phrase *animate form* to denote the entire being of a living creature. The concept is thus grounded on the pillar of her phenomenology, the *primal* role of movement, which is there throughout the life of any organism. Indeed, as everything living is animated (Sheets-Johnstone 2011, p. 543), the term animate form must naturally encompass *all* forms of life. Moreover, animate form encompasses all aspects of *being alive*, of being an organic organism. It encapsulates the nature of being alive, portraying movement as nothing less than life itself.

> We do well to begin our investigations of life by acknowledging […] animation as the foundational ground of life itself. Animation encapsulates what is fundamental to life, the vibrant and spirited way living creatures come onto the world and the vibrant and spirited way that is gone when they die; it engenders dynamics, the essence of life in all its varied and vital kinetic contours; it articulates in an exacting linguistic sense the living wholeness of animate forms and is thus properly descriptive of life itself. What is fundamental is that we are indeed animate forms of life, and as such, are necessarily and from the beginning subjects of a world, an Umwelt (Sheet-Johnstone 2011, p. 453).

Animate form is to a great extent the continuation and further development of the Husserlian concept of *animate organism*. Sheets-Johnstone explains that "animate organism" was used by Husserl to call attention to a *species-specific sense of animation* that highlight animateness as foundational to both the internal and external perceptual world of living beings, "to living creatures in the full sense of their livingness" and emphasise the "significance of movement to creaturely life" (Sheets-Johnstone 2011, pp.

115-6). This foundational insight provided by Husserl, calls attention to the dependence of movement for the possibility of perception. “Not only is our own perception of the world everywhere and always animated, but our movement is everywhere and always kinesthetically informed” (Sheets-Johnstone 2011, p. 113). Knowledge of ourselves and our own movements arise from our own *felt* sense of it. Thus, the “evolutionary understanding of consciousness on the basis of animate forms” as it is described by Sheets-Johnstone herself (2011, p. 69), focuses on *animation* as the foundation for consciousness. The concept of *animate form* is also heavily influenced by Aristotle’s understanding of motion as any kind of change. In Sheets-Johnstone’s (2011, p. xvii) reading of Aristotle, motion is the fundamental principle of nature and is thus of great significance to his understanding of *anima*—the soul. It is found in life as various kinds of change: growth, locomotion, and alteration—though not necessarily all present in all forms of life. Growth is a quantitative change, locomotion is a dimension of the soul, and alteration is a qualitative change (Sheets-Johnstone 2011, p. 233). In this inherently dynamic world, movement is a dimension of being, rather than a “capability” or “feature”[3].

Perception is traditionally categorised into five “classical” senses. Even Gibson (Gibson 1966, 1979), whom Sheets-Johnstone acknowledges as giving an account of perception that is both insightful and radically novel in an otherwise cognitivist-oriented time and domain, considers neither proprioception nor kinesthesia[4] as perceptual systems in their own right. Knowledge about self-movement is picked up by means of the other perceptual systems. This, Sheets-Johnstone holds, does not give either proprioception or kinesthesia their due. “Proprioception in general and kinesthesia in particular advert to a knowing subject, a subject that, at minimum, knows when its moving and knows when it is not” (Sheets-Johnstone 2011, p. 69). Rather than our sense of proprioception and kinesthesia springing out of the other five classical senses—or perceptual systems if you will—these five senses are grounded in proprioception and kinesthesia. Therefore, Sheets-Johnstone highlights them as a vital aspect of not only how we know ourselves, but also how we know ourselves as subjects in the world. Proprioception and kinesthesia constitute our *sense* of self-moving; we experience ourselves as animate forms.

Proprioception is a general sense of movement and position, in the form of a continuous awareness of our own bodily condition—such as kinetic potential, postural conformation, and spatio-temporal dynamics of one’s own movement. Sheets-Johnstone (2011) describes it as an internally structured and felt consciousness, a *corporeal consciousness*, founding a corporeal sense of “here and now”. This corporeal here and now adapts to the growth and change of the creaturely form and always supports contextually appropriate behaviour to the situation at hand. It is a sensitivity to oneself that constantly informs about the posture and potential for movement, “an awareness of movement and position through tactility as well as kinesthesia, that is, through surface as well as internal events, including also a sense of gravitational orientation through vestibular sensory organs. Kinesthesia refers specifically to a sense of movement through muscular effort" (Sheets-Johnstone 2011, p. 73). Kinesthesia is the sensitivity animate forms have to their own self-movement. Through this sensitivity, we have an awareness of the qualia that arise through movement, which can be forceful, smooth, weak, clumsy, and so on. In other words, qualia are created by movement itself. Any animate form that is adapted to a locomotive lifestyle can, through their kinesthesia, distinguish between the particular qualitative dynamics of the movements they are engaging in. Thus, qualia are not pure "mental phenomena" but emerge from the movement of the organism itself as *bodily-felt* distinctions of movement qualities (Sheets-Johnstone 2011, p. 51). For instance, we effortlessly distinguish running from walking and walking from rest. We need no linguistic concepts to understand when our movements are clumsy or forceful or smooth or weak. These qualities are first and foremost *felt*. In this manner, kinesthesia is "fundamental not only to our knowledge of 'which thing in the world we are'; it is fundamental both to our ability to make our way in the world — to move knowably in it — and to our knowledge of the world itself" (Sheets-Johnstone 2011, p. 52). The qualities created by our own movement and felt by kinesthesia are foundational to our sense of agency (Sheets-Johnstone 2011, p. xxii).

This perspective on movement and consciousness tells me that my movements are not just the wilful self-movement engaged in activities. The entireness of my being is engaged in all activity. I breathe while I walk, sleep, run, rest, jump, eat. I look around and tap my foot while I chew and eat, I stop breathing as I swallow food and resume again when I am done. I tap my foot and smile and breathe, as I lift the next spoonful of food to my mouth. All the while I hold my posture and keep vital organs up and running. This is so worked into my own felt sense of being that most of it passes me by in the moment. I turn and respond to the world that surrounds me in the here and now, and spontaneously act knowledgeably and appropriately.

6.2 LEXICAL BAND-AIDS (DOESN'T CURE DUALISM)

A pronounced element throughout Sheets-Johnstone's phenomenology is her commitment to arguing against Cartesian dualism. She does this through addressing the consequences of what she likes to call "the 350 year-old wound inflicted by the Cartesian split of mind and body" (Sheets-Johnstone 2011, p. 453). Accordingly, *The Primacy of Movement* is pervaded with numerous and varied articulations of "what it is to be the bodies we are" and a meticulous spelling out the ramifications of linguistically disregarding the fact of animate being as holistic. Sheets-Johnstone maintains that, as a direct aftereffect of Cartesian dualism, society regards "the body as drone to an all-powerful, rational mind" (Sheets-Johnstone 1992, p. 2). These views have become deeply rooted in everyday attitudes: We are "indoctrinated into thinking we are minds, rather than acknowledging that we are first and foremost bodies" (Sheets-Johnstone 2011, p. 304). *The body* is viewed as a material possession of *the self* rather than the actual self itself. While it may be a seemingly trivial nuance, Sheets-Johnstone contends that there is a significant difference between "the body that is me and the body that is merely mine" (Danto 1999, p. 197; cited in Sheets-Johnstone 2011, p. 498).

Sheets-Johnstone demonstrates how the English vocabulary (amongst many) is ill-suited for describing movement in a phenomenological manner. She points to the main culprit as the underlying dualism in the English language. This dualism has the form of lexical distinctions that foregoes the inherent animacy of living beings, words that "package" the mind into the body or place boundaries on the dynamic phenomenon of movement. This is what she refers to as *lexical band-aids*. The most prominent example of this packaging is the term *embodiment* as it packages a "something" (whatever that may be) *into* an organism, instead of recognising that an organism *is the body it is*. Lexical band-aids like embodiment perpetuate the body–mind split and "testifies to the fact that we have not yet fathomed what it is to be the bodies we are" (Sheets-Johnstone 2011, p. 300). The notion of lexical band-aids may seem unnecessarily quarrelsome. However, Sheets-Johnstone does raise a good question when taking the term into scrutiny and asking "*what* [exactly] we believe to be embodied—a mind, a soul, a spirit, a self, our organism, or whatever—is embodied by the body" (Sheets-Johnstone 2011, p. 312). In so doing, Sheets-Johnstone calls attention to how humans, as we grow into adults, forget to pay attention to the *kinetic knowledge* we acquire from birth and throughout our infancy and childhood.

When we, in the natural attitude, think of movement in forms of life it is usually as animals moving their limbs (heads or legs) or moving from one point to another within the environment. However, this understanding of movement constitutes what Sheets-Johnstone calls a *mathematisation* of movement—to which she is of course strictly opposed. "Factual" or "mathematised" conceptions of movement are not consistent with a *phenomenological account* of the kinetic experience of moving. For instance, while it might seem like merely a trivial linguistic nuance, there is, in terms of animation, a significant difference between "laying still" and "being still." From a mathematized perspective, they are the same as they neither involve any shift in position in the landscape nor a change of bodily position. However, from a phenomenological, first-person perspective "laying still" is a bodily-felt dynamic while "being still" is synonymous with being dead. Rather than "the standard dictionary definition of movement as a *change of position*" (Sheets-Johnstone 2011, p. 202 [emphasis added]), movement in animate forms unfolds in what Sheets-Johnstone describe a *kinetical dynamic*—what we might call *self*-movement in a more familiar vocabulary. Contrary to how we usually think about the movement of living beings, it is not "a thing that I do, an action that I take, or a behaviour in which I engage" (Sheets-Johnstone 2011, p. 424). Movement is, as an *unfolding dynamic*, impertinent and intangible, passing as moments, with no clear beginnings or endings. This is how, Sheets-Johnstone explains, self-movement is in fact *experience itself*, dynamically unfolding in the here and now. Experience is momentary and fleeting rather than bounded into separate states of various doings or beings. As such, self-movement is not an *object* of consciousness as much as it is *consciousness itself*, making consciousness one of many dimensions of living (Sheets-Johnstone 2011, p. 52) and inseparable from movement. This lays the foundation for the *corporeal consciousness* of all animate forms. From before we are born, we start forming our corporeal—not mental—concepts.

Sheets-Johnstone accentuates that as long as we keep using words that perpetuate the Cartesian divide, it will remain part of our thinking. More importantly, we remain ignorant of our own experience of the world as taking place *through* moving and *in* movement. Even though our kinetic knowledge is wholly integrated into our being-in-the world (and thus fundamental to how we are able to make our way in it), it is increasingly disregarded when we transition into adulthood and our focus is turned toward linguistic concepts. This does not mean that infancy is a "primitive" and proto-linguistic state.

6.2.1 Thinking in movement (not words)

Sheets-Johnstone's main thesis that movement is the foundation for consciousness and thinking does, according to her, stand in stark opposition to the common conception of the brain as its foundation. The notion of *thinking in movement* as Sheets-Johnstone (2011, p. 426) presents it, challenges the widespread assumption that thinking, language, and rationality are the root of all human cognition—which in turn includes human experience. This assumption is built on two other assumptions about thinking: 1) Thinking is tied to language and takes place only via language, and 2) rationality is exclusively tied to thinking and language. These assumptions are accompanied by a parallel assumption rooted in the Cartesian split of mind and body, namely that *thinking is something only a mind does and doing or moving is something only a body does*. This is, according to Sheets-Johnstone, akin to claiming that the thoughts of animals are successively transcribed or transliterated into movement, a claim she naturally finds absurd. To her, these assumptions imply that "when the mind formulates a thought, for example, the tongue and lips move to express it; when the mind thinks of going to the store, the body compiles by walking or driving there" (Sheets-Johnstone 2011, p. 428). Indeed, her concept of *thinking in movement* directly challenges the assumption that thought in one's head must always come prior to its corporeal expression. Movement should not be conceived as the medium of a body's transaction with the world. Rather, the movement of living beings is qualitative and the natural mode of being a body. Animate forms inhabit movement in the literal sense of living in it.

The sense of oneself as an animate form begins already as a foetus in utero. Our animation starts, or continues, at conception. Understanding movement as a fundamental Aristotelian principle of life entails seeing it as always dynamically unfolding. Thus, the merging germline cells are never not in movement in the Aristotelian sense of change: Indeed, they are always moving, and each generation forms the next link in an unbroken chain of movement. It is in this manner movement is primal. Movement is our primary mode of being. Not only do we find ourselves in a continuous state of moving; we were literally *born in movement*. As Sheets-Johnstone (2011) point out on several occasions, life is not, and can never be, *still*: "We are indeed either movement-born or still-born" (p. 200); "Newborn creatures *move*. Stillborns are precisely stillborn" (p. 347 [emphasis in original]); "In the beginning, after all, we do not *try* to move, *think* about movement possibilities, or put ourselves to *the task* of moving. We come straightaway moving into the world; we are precisely not *stillborn*" (p. 117 [emphasis in original]). Stillness is the *absence* of animate movement. Movement is always already there, prior to any linguistic concepts. Through moving, not abstract thinking, we learn the ways in which we can move ourselves.

> Moreover, in the beginning, we are not surprised by our movements, disappointed by them, or wish that they were different. In the beginning, we are simply infused with movement—not merely with a *propensity* to move, but with the real thing. This primal animateness, this original spontaneity that infuses our being and defines our aliveness, is our point of departure for living in the world and making sense of it. It is the epistemological foundation of our learning to move ourselves with respect to objects, and thus the foundation of developing a repertoire of "I cans" with respect to both the natural and artificial array of objects that happen to surround us as individuals in our particular worlds. It is in effect the foundation of our sense of ourselves as agents within a surrounding world. But it is even more basically the epistemological foundation of our sense of who and what we are. *We literally discover ourselves in movement*. We grow kinetically into our bodies. In particular, we grow into those distinctive ways of moving that come with our being the bodies we are. In our spontaneity of movement, we discover arms that extend, spines that bend, knees that flex, mouths that shut, and so on. We make sense of ourselves in the course of moving. We discover ourselves as animate organisms (Sheets-Johnstone 2011, p. 117 [emphasis in original]).

Through moving during infancy and early childhood, we discover our *corporeal powers*. We experience directly how the world responds to these movements, "we discover a realm of sheer kinetic 'I cans': I can stretch, I can twist, I can reach, I can turn over, and so on" (Sheets-Johnstone 2011, p. 117). The particular dynamic qualities of our movements are structured into non-linguistic *corporeal concepts*—what we might call "activities", "emotions" or other homeostatic states—such as hunger and tiredness, eating and resting. Thus, thinking in movement is a *kinetic intelligence*—an intelligence that has been with us since before we were born—that gives us an understanding of the dynamic environment. It is our kinetic intelligence, what Sheets-Johnstone also call our *kinetic bodily logos*, that enables movement to be spontaneous and contextually appropriate.

Consider hunger. I have since birth learned to know and understand the feeling of being hungry, and the experience of it gradually appearing between meals followed by a fading away after consuming nutrients (the corporeal concepts we call "eating" or "drinking"). When I am hungry, I seek out food. In my corporeal conception of what it means to be hungry, I know what to do to make that feeling go away. As it is gradually replaced by a new feeling, that of fullness, I will eventually stop and begin engaging in some other activity. Equally, when I am thirsty, I drink water. I stop drinking water when I am no longer feeling thirsty. And I know that if I do not, I will eventually start feeling unwell. None of these feelings or activities are bounded states, but gliding transitions between being hungry to becoming less hungry, to feeling full—and even unwell if I do not heed the rising feeling of fullness. Hunger or thirst does not appear suddenly from one moment to the other. While the transition between states (from "not hungry" to "hungry") can be slow or swift, it never "skips" from one state to another. Animate activities are not bounded and neatly stacked into a set of successive states.

When we experience these transitions within ourselves, we are witnessing our subjective animate movement as the *change that we live through*. Although humans have linguistic labels for their corporeal concepts (for example hunger, resting, or eating), "there is nothing basically linguistic about them at all" (Sheets-Johnstone 2011, p. 438). Corporeal concepts are grounded in the bodily form of an organism. Even for the human animate form, it is our movements, and thus our thinking in movement that is consistent with the actual here-and-now experience of ourselves in the world. We do indeed have words to encapsulate or describe corporeal concepts like, for instance, hunger and thirst. But the words are not the feeling or experience of hunger or thirst themselves, neither do thinking or expressing them summon the feeling. The feeling is always prior to the linguistic symbol. For instance, I do not become hungry because I *say* that I am hungry. Rather, it is the other way around; I say that I am hungry because I am experiencing a sense of hunger (unless of course, I am lying to myself and those around me). Thinking in movement and thinking in words are two different modes of thinking. A logos that is different in kind, not degree. Sheets-Johnstone (2011) demonstrate that in both an evolutionary and ontogenetical perspective, movement is not pre-linguistic, and that "if anything, *language is post-kinetic*" (p. 438 [emphasis in original]). This should not be confused with the notion that humans do not think using language at all; humans are indeed linguistic beings and linguistic concepts have a very real influence on experience. Neither should it be confused with the notion that movements cannot, conceptually, be broken into behaviours and acts, because this is exactly what happens with language. However, it is not the mission of language to capture "the qualitative dynamic metaphysics of aliveness" (Sheets-Johnstone 2011, p. 436). While movement is indeed discernible into meaningful patterns that can linguistically be broken into behaviours and acts, movement is *primarily* a continual dynamic flow. Wilful activities find a place in parallel with automatic processes. They all glide into each other and overlap.

The human capacity to *think in movement* does not diminish as we age, it is gradually submerged and becomes hidden by our capacity and practice of thinking in words. When human infants learn language, it starts not firstly with concepts, but with discovering oneself as a sound maker, and this initial discovery leads to a further "discovery of themselves as articulators" (Sheets-Johnstone 2011, p. 329). As we mature, linguistic capabilities strengthen. In parallel with solidifying a sense of basic corporeal powers, we learn their vernacular labels. Slowly as navigating the environment turns habitual, our attention shifts from discovering and feeling corporeal powers, toward expressing our thinking, not as movement but as bounded states and actions.

6.3 A PRESENCE-BASED SOCIALITY

In the traditional, cognitivist perspective on thinking, the human way of everyday environmental sense-making takes place through a process of cognising bodily sensory impressions inside the brain, where knowledge about the world is structured by mental models. In this chapter, we have seen that another explanation is offered by Sheets-Johnstone's perspective on thinking as a corporeal phenomenon, a perspective telling us that our concepts about the world are corporeally, not mentally, forged. "Our capacity for self-movement and our experience of self-movement are indeed cornerstones of our sense-makings" (Sheets-Johnstone 2011, p. 506). On the basis of their animation, living beings have a natural disposition toward meaning, which Sheets-Johnstone (2011) points to as a semantic-kinetic relationship. Indeed, exchanging and finding meaning is by no means limited to the human animate form, but exists as a *readiness toward meaning* found throughout all of biology. All animate forms of life are capable of making sense of their environment. If they weren't, they would soon cease to be animate. Hand in hand with our animation, Sheets-Johnstone explains, there is a capacity to *respond*, a capacity that "emerges with life itself" (2011, p. 343). In the same manner animate forms come into the world moving, they come into the world with a disposition toward meaning, a pan-animate *being-toward-meaning* (Sheets-Johnstone 2011, p. 344) making it possible to respond and adapt to the ever changing dynamic environment.

> Living creatures, ourselves included, are thus responsive in a sense beyond the sense in which biology texts speaks of responsivity. We are all of us *semantically* responsive, just as we are *semantically* receptive. We are all of us inherently meaning-seekers and meaning-finders. Meaning-seeking readily explains why it is the receiver, the displayed-to-animal, that solidifies meaning and indeed, why it assents to meaning in the first place. It straight away recognises the world and other creatures in its world as having semantic value. For all of us, to be intentionally active is to move spontaneously toward meaning and in virtue of meaning (Sheets-Johnstone 2011, p. 344).

That animate forms phenomenologically *turn-toward* in attention, Sheets-Johnstone (2011, p. 343; p. 503) explains, means there is a semantic congruency between movement and meaning: "How we move — as well as how we perceive the movement of others — is concordant with intentionality in a phenomenological sense, that is, in the sense of meaning" (Sheets-Johnstone 2011, pp. 504). It follows, Sheets-Johnstone explains, that there is a "*built-in semiotic specificity in the movement of living bodies*" (p. 302 [emphasis in original]) by which we intuitively make sense of the movement of others. We take "what is living to be that which moves itself and to apprehend *what is not moving and has never moved* to be precisely inanimate" (Sheets-Johnstone 2011, p.

116 [emphasis added]). This fits with the suggestion that perceptual animacy is rooted in a kind of perceptual grammar (Scholl and Tremoulet 2000), though it would probably be described as a *kinetic grammar* by Sheets-Johnstone. Should such a kinetic grammar exist, we might imagine it to contain the semantics of liveliness. Animate forms understand the world in terms of that which is familiar, and with movement, we share a most intimate familiarity. Indeed, animate forms relate to each other "in and through movement, through a *kinesthetically* and *kinetically inflicted* intercorporeality" (p. 515 [emphasis in original]). In other words, our concept of *aliveness* is grounded in movement, and thus, when we experience the presence of other animate forms, we experience them as *living beings that move*. In and through our corporeal consciousness—which gives us our sense of ourselves as agents—we also recognise that the movements of other animate forms as equally agentive. Sheets-Johnstone describes a sociality that does not necessitate a purposeful and engaged social interaction with the other agents. A sociality simply rooted "in a dynamic intercorporeality" (Sheets-Johnstone 2011, p. 515). This forms the basis for what I take to be a *kinesthetic intercorporeality*, a sociality that is at its most foundational level presence-based.

Brincker (2016) address the issue of *perceptual agency* as one of "how one, on the basis of some objective features of body and behaviour, can judge others as having subjectivity, as being someone rather than merely something" (p. 443). For traditional cognitive sciences, consciousness and the mind are viewed as purely internal structures. In this view, the mind, which is at the seat of agency, must necessarily be hidden. Any "perceived" sense of agency must be the result of an inference on the side of the perceiver, not of any intrinsic qualities of the perceived entity. In light of Sheets-Johnstone's evolutionary rooted description of animate forms, it is difficult to deny a linkage between animacy and agency, as her phenomenological explications demonstrate time and again that these are inseparable. If it is so that animate movement is at the source of a creature's corporeal consciousness, and the root of our own felt sense of agency, then agency must be an inseparable part of the corporeal power and (self-)movement. It is in this sense that the capacity for movement—animation—is equal to and inseparable from agency (Sheets-Johnstone 2011, p. xxii). It holds then that, if the animate movement of others witnesses their aliveness, it necessarily also witnesses their agency. Indeed, if we take what Sheets-Johnstone suggest about thinking in movement seriously, it becomes obvious that "the mind" is indeed not located on any such "inside" of agents, hidden for the world to see. Therefore, if one, as Brinker (2016) suggests, "does not predefine minds as necessarily hidden and diagonally opposed to behaviour, then minds, subjectivity, agency and even the ability to perceive are at least possibly perceptible" (p. 444). In the cognitivist perspective, the phenomenon of perceptual animacy must naturally be the result of mental acts—as sensory input cognised and giving meaning by one's existing mental models. Sheets-Johnstone shows us, however, that there is nothing mental about them. Mental models only occlude what is perfectly clear about the movement of other agents. We may say that—if not the *thoughts*, per se—the consciousness of other living beings will not be hidden, but apparent in their every movement.

If agents are "those having the power to act" (Sheets-Johnstone 2011, p. 51), then acts, here understood as the self-movement rooted in corporeal powers, witness the presence of agency in a kinesthetic intercorporeality. The kinetic grammar that reveals animacy must necessarily also reveal agency. Instead of understanding agency as being a hidden quality of entities, the agentive movements we see in perceptual animacy are not of *mental activity*, but of a corporeal consciousness. In addition to recognising the liveliness of others, we share in a kinesthetic intercorporeality many corporeal concepts, depending of course on shared characteristics of our forms. "We instinctively know what it is like to stare and to be stared at [...] we know what it is because we know what it is to be an animate form" (Sheets-Johnstone 2011, p. 304). In the natural attitude, we take for granted this basic and foundational sociality shared by all animate beings. "We take it for granted, as it were, that, whatever the situation, we will find other creatures meaningful and the world in general meaningful; and indeed, both are consistently full of meaning for us. They are consistently meaningful for nonhuman creatures as for human ones" (Sheets-Johnstone 2011, p. 344). What is described here is co-existence as a most fundamental form of sociality rooted in kinaesthetically understanding of being in the presence of another animate and agentive being. Thus, it may be assumed that is not single movements by themselves that "reveal" the agency of the moving entity, but a quality of the dynamic expression, a richness so to speak, that also reveals the capability to turn-toward in meaning. In turn, agency and intentionality are inseparable.

If we buy into the idea that animate sociality is founded in a kinesthetic intercorporeality, we must also acknowledge that the liveliness and presence of others is of our immediate interest, and that when perceiving movement, we naturally turn-toward it in interest. "What moves straightaway capture our attention; it is consistently at the focal point over what is not moving" (Sheets-Johnstone 2011, p. 116). What happens when catching a glimpse of something falling in the corner of your eye is an immediate recognition of something that is intercorporeally meaningful, even if turns out that it was not caused by an animate other. Catching a glimpse of something falling in the corner of your eye can give a short-lived and momentary experience of the presence of others. This misconception ceases as soon as the motion falls to rest and the source of the movement is upon closer examination identified as some inanimate object. While Sheets-Johnstone proclaims that animate forms "intuitively equate aliveness with movement" (2011, p. 117), Gelman et al. (1995) argue that the animate—inanimate categorisation is not made on the basis of visual dynamic expression (observable movement) alone, but that we also pay attention to material qualities and physical form of the entity. This is one of the ways we attune our skilful intentionality. On the basis of this knowledge, entities will be categorised as animate (and agentive) even when they are resting or feigning stillness. Similarly, dead leaves will be identified as inanimate even if they are forced to motion and swept away by the wind.

As Sheets-Johnstone's phenomenology of movement demonstrates, *movement* is more than just a shift of location in space. Animate movement, *animation*, is a fundamental property of all forms of life. So is intentionality, an innate disposition of turning-toward in meaning: "Flowers turn toward the sun; pill bugs curl into spheres; lambs rise on untried legs, finding their way into patterned coordination. The phenomenon of movement testifies to animation as the foundational dimension of the living" (Sheets-Johnstone 2011, p. 453). The paradox, so to speak, is that not all cultures recognise the fundamental animation and turning-toward in meaning of, for instance, floral animate forms (Santazilia 2020). Yet we find the movements of shapes as simple as triangles lively enough to perceive them as purposeful—as coming from a place of intentionality. What is easily overlooked about the movements of the shapes in the study by Heider and Simmel (1944) however, is that they do not move in this way by their own devices. They have precisely been *animated by* another being. Moreover, these shapes have been purposefully animated in a way to express a narrative. That the movements of these shapes are different from the falling of things, or leaves being propelled into motion by the wind, is not trivial. That they seem to move purposefully, as if capable of turning-toward the other shapes in meaning, when we know they cannot, may well be because this apparent intentionality is not really bestowed these shapes by their onlookers. It essentially emanates from the turning-toward of an animate form, in this case, a human animator. When observing the moving shapes, one sees an intentionality that was already there, albeit not originating from the shapes themselves. Similarly, I contend, the intentionality found in robotic artefacts may well be sourced to the turning-toward in meaning belonging to its human creators.

7 Discussion

The ships hung in the sky in much the same way that bricks don't.
— *Douglas Adams, The Hitchhiker's Guide to the Galaxy*

7.1 INTRODUCTION

In this thesis, I have held a phenomenological gaze on human–robot interactions. Through three different, but adjacent theories, I have maintained a focus on meaning as subjective and the first-person perspective as central to understanding experience. The primary focus has been on the manner in which a robotic artefact can appear to have better interactive capabilities than they do.

In Chapter 4, I suggested to use *artificial Umwelt* to illustrate how the environment might appear for a robotic lawnmower. The example used in this chapter is taken from the analysis made by Soma and Herstad (2018). The analysis provides a new perspective on how robotic artefacts are susceptible to failure, and subsequently be dependent on human assistance. It does so by offering insight into how robotic artefacts perceive their environments. In turn, it will be possible to identify what aspects of the environment it will be able to discern and react to and vice versa. In Chapter 5, I explicate some of the differences between Gibsonian affordances and Norman's (1988) interpretation of them. As affordances are the meaningful relationship between a person and their environment, it cannot be fully predicted exactly what behaviours will be offered to a person through the affordances they perceive. For design to support mutually intelligible interactions between humans and robotic artefacts, attention should be directed toward the way people skilfully turn toward the opportunities for behaviour they evaluate to be meaningful and appropriate in a given situation. In Chapter 6, I explore the Maxine Sheets-Johnstone's (2011) phenomenology of movement. Her phenomenology explains how movement is the primary mode of thinking for all animate forms, humans included, and that all forms of life come into the world with a readiness to turn-toward in meaning. I explain how one in Sheets-Johnstone's phrase *kinesthetic intercorporeality* can see the contours of what may be characterised as presence-based sociality. This sociality springs from our own knowledge of ourselves as animate forms, as living beings with the corporeal power to act. Through knowledge of our own bodily-felt agency, we recognise the agency of others when we observe their animate movements. Thus, through this presence-based sociality, I explore the intertwined nature of animacy, agency, and intentionality.

In this chapter, I will discuss the perspectives presented in Chapters 4, 5, and 6 in combination with the central terminological contributions of the papers included in this thesis. In the first part of the discussion, I use the notion of *datanomy* (Soma et al. 2022) to characterise the robotic *artificial Umwelt*—discussed both in Soma and Herstad (2018) and Chapter 4—and of the ways this matters for human–robot interaction. In the second part, I relate *movement acts* (Schulz, Soma, and Holthaus 2021) to the presence-based sociality explored in Chapter 6. Finally, I discuss how what has been presented in this thesis and in the attached papers can be used to outline *kinetic affordances* for human–robot interaction.

7.2 DATANOMOUS TECHNOLOGIES AND THEIR ARTIFICIAL UMWELTEN

In Soma et al. (2022), we explicate some of the differences between human and machine autonomy, as well as discuss the ways in which such technologies can support or limit human autonomy. The perspective we take on human autonomy focuses exclusively on its nuances in a socio-material and socio-technical environment. It highlights that, in addition to being physically situated in an environment, all humans are socially and culturally situated. In turn, all human action is situated. Obviously, we are hardly the first to make this point. Our perspective on the situated human is not surprisingly inspired by the insight offered by Suchman (1987) on this subject. However, we also address some slightly different nuances.

These nuances are discernible through the notion of *space for action*, a notion referring to the "full range of possibilities available to a human in a situation" (p. 12). This is not to be confused with the number of choices available. A person's space for action says something about their independence and freedom in acting. More importantly, it emphasises their ability to do so in terms of their competence. In other words, it refers to both their possibility *and* ability to redefine the situation they are finding themselves in. Machine autonomy is implemented to give the technology operational independence. That is independence from human intervention in the form of control or command over the responses or functioning of the system. Our discussion intends to show that autonomy is not a *feature*, but a state of being where human subjects have the space for action to redefine their own situation. In contrast, the operational independence of so-called autonomous technologies is restrained by a general inability to make contextually appropriate adaptations. The profoundly limited situational awareness of these technologies implies they also lack the space for action to redefine the situation they are in. In our paper, we introduce the term *datanomy* to emphasise exactly this contrast between what it means to be self-driven (autonomous human) and data-driven (datanomous technology).

The theories presented in Chapters 4, 5, and 6 all acknowledge the situatedness of the subject in their own ways. That is, they hold in common the realisation that *all* kinds of living beings—animate forms—are highly adaptable to the dynamic environment because they are inseparable from it. The animate capacity to *think in movement*, which can also be understood as an innate capability for kinetic spontaneity, enables them to always act appropriately in the dynamic environment. For this reason, Sheets-Johnstone (2011) holds, responses to the dynamic environment cannot be pre-programmed.

> A moment's serious reflection on the matter discloses a major reason why this sensitivity to movement is both basic and paramount: no matter what the particular world (Umwelt) in which animal lives, it is not an unchanging world. Hence, whatever the animal, its movement cannot be absolutely programmed such that, for example, at all times its particular speed and direction of movement, its every impulse and stirring, its every pause and stillness, run automatically on something akin to a lifetime tape (Sheets-Johnstone 2011, p. 55).

The term datanomy is meant to highlight that *data-driven* technology interacts with the world exclusively through data. To characterise data, I wish to point to two dimensions of ecological information it does not possess. First, data lack the qualitative richness, the "endless resolution" of ecological information. Second, it is not abundantly available. In Soma et al. (2022) we describe data as an "indicator or representation of a phenomenon, like fever indicates illness" (p. 20). By this, we mean to highlight that data, as items of information, come in the form of sets of values or variables. Another essential aspect of data is that is it collected by people. It is "purposefully curated and intended to assist in identifying the meaning represented by the values [...] data is designed" (p. 20). Data is limited, curated, and pre-defined; data is always historical. This makes it difficult to adapt and continue independent operation should the situation change in a way that has not been anticipated during either programming or training (depending on the technology a specific system is built on). Because data lack contextual awareness, so does the data-driven piece of technology. Accordingly, the information available in data pales in comparison to the ecological and sociomaterial information that specify human opportunities for action.

The *artificial Umwelt* can be used as an illustration of the data-driven perspective of the robotic artefact. What artificial Umwelt shows us is that *datanomy* is the foundation of a robotic artefact's capacity to act and respond to the environment. When we realise that the artificial Umwelt of the robotic artefact is *datanomous*, it becomes clear that the sensory range or sensitivity alone is not what limits the robotic capability for contextual adaptability. A robotic artefact can be equipped with a camera that captures light as good as, or with higher resolution than, the human eye is capable of. It can also be equipped with sensors able to take in a wider range of stimulus energies than those available to the human perception, like for instance Wi-Fi or Bluetooth signals. However, what the sensors capture must be converted into digital values, what we usually refer to as data, for them to be useful to the system. This process necessitates selecting which data that should represent the environmental information.

This is not to imply that robotic technology is not situated, just that its version of being placed in a situation is fundamentally dissimilar from animate *situatedness*. Building on the argument made by Suchman (1987), Robertson and Loke (2009) explain how interactive technologies “act in accordance with the constraints and opportunities” (p. 3) of their situations, just like people do with theirs. However, the digitally represented version of a given situation will be profoundly different from how the situation is understood by the human interacting with it, exactly because the resources available to people are different from those available to the system. The data values that constitute a system’s foundation for situational awareness have all been pre-defined by human programmers and developers. Dourish addresses a similar issue when he discusses the role of context in systems design, arguing that context cannot be encoded and represented. Context cannot be modelled because it is not a preexisting fact. It is “an outcome, rather than a premise” (Dourish 2004, p. 22). The issue of making machines and interactive systems aware of context is a tricky one, and the notion of artificial Umwelt provides further insight into why this is. As we explain in Soma et al. (2022), a “system does not know anything about the phenomena it measures or the operations it performs. Just like a robot vacuum cleaner does not know anything about dust, the insulin pump does not know anything about blood or insulin” (Soma et al. 2022, p. 21). As described by Soma and Herstad (2018), the robot lawnmower does not know anything about gardens or grass. In fact, it knows very little about anything concerning its environment. What the robot senses, or can *make sense of*, is limited by the data format prescribed by a combination of training and programming. When apples fall to the ground the robot lawnmower will not register their presence. Not only is there no such thing as apples in the artificial Umwelt of this particular lawnmower robot. What is made obvious by the illustration in Figure 4.1 is that there are in fact no such things as lawns. This is similar to the functionality of the smart insulin pump described in Soma et al. (2022). This device is only independent (and smart) in the closed-off context of a patient and their blood circulatory system. However, the pump cannot make any decisions based on other contextual information relating to the life of the patient.

Both examples illustrate that datanomous systems can come across as independent—operating without human intervention. This independence does, however, presuppose that the data values needed are perfectly fitted to the task. A fallen apple on a lawn will seem trivial to the human Umwelt but will be decisive to the robotic lawnmower’s independent functioning. Even minuscule changes to the environment will become momentous hindrances because, in its empty artificial Umwelt, the apple does not exist. Although new sensors, new actuators, and new data formats can be added, this adaption of the artificial Umwelt is not made by means of the technology itself. Precisely this is the most salient distinction between the self-driven and the data-driven: The robotic artefact will never be the *creator* of its own artificial Umwelt, like the animate form is of its own. What is meant by the organism being the constructor of its own Umwelt is not that it is free to add new perceptual systems to itself. It means that the organism skilfully evaluates which affordances deserve attention and action, and which can safely be ignored. The individual constantly evaluates what is relevant in a constantly changing environment. For robotic artefacts, programming naturally contributes to determining which sensory data should be prioritised, in which order, and how it should react to it when active. But this prioritising has precisely been programmed and prescribed. It is never situationally decided and determined. When apples become stuck in the rotating blades of a robotic lawnmower, the artefact does not suddenly understand what an apple is. Consecutively, it is not its armless morphology that first and foremost renders it unable to remove the apple and resuming operation. The robotic artefact can neither define nor redefine what is meaningful to it and for *what*. It is in this sense the data-driven artificial Umwelt lacks its own space for action. The decisions of the robotic artefact essentially originate from a place previous to or outside of the situation. The world is constantly changing, but the artificial Umwelt is *static*. This makes it challenging to endow robotic artefacts with an appropriate contextual awareness; any such awareness will soon be obsolete.

7.2.1 Analysing the artificial Umwelt

That a robotic system is making *datanomous* rather than *autonomous* decisions tells us something important about the structure of these decisions. A datanomous technology only acts independently inside the frames of its given context, and it can only interact with the environment through its specific data forms. In this respect, the concept of artificial Umwelt can be used with a similar purpose to that of Suchman's (1987) analytical framework on human–computer interaction.

THE USER		THE MACHINE	
I	II	III	IV
Actions not available to the machine	Actions available to the machine	Effects available to the user	Rationale

Figure 7.1: Suchman's framework presented in *Plans and Situated Actions* (Suchman 1987, p. 76).

What Suchman's framework showed, was that while there is a shared understanding between the user and the system (columns II and III) through the user interface, both parties have a situational understanding that will be unintelligible to the other (columns I and IV). For exactly this reason Robertson and Loke (2009) found the framework useful as a design tool. The perspective offered by this framework made it easier to determine the contextual awareness of the interactive system they were prototyping. In regard to the robotic artefact, column IV can be paralleled with what I have presented as a robotic artefact's datanomy, and its rationale for reacting to interactional exchanges. Column I pertain to everything about the user or an environment that will be unavailable to the artificial Umwelt. In Chapter 4.2.1, I demonstrated the practice of partaking in participant observations of a robotic artefact's functional cycles. This practice aims at gaining insight into what it is about an interactional situation the robotic artefact will be aware of. One can, by arriving at what constitutes a meaningful object for a robotic artefact, determine which actions will be available to the artificial Umwelt.

Through participant observation of the artificial Umwelt, one can analyse the *actual* interactional capabilities of the robotic artefact. When the presence of a person and their actions are viewed as meaningful objects within an artificial Umwelt, these can be understood in terms of which artificial functional cycles they initiate. As is made clear by the analysis of Roberto's artificial Umwelt in Chapter 4, it is not the length of the grass, or even the existence of grass, that enables the robotic lawnmower to cut it. What Figure 4.1 reveals is that the garden is almost entirely empty. Indeed, only very few things about the human Umwelt is intelligible to the robotic artefact. Lawns and grass do not appear as meaningful objects to the robotic artefact—and neither do apples. The things it can react to are almost exclusively the electric fence that the human facilitator sets up before use. Otherwise, it will only be able to register the presence of objects large enough to trigger its crash-sensor. This means that there is, practically speaking, only one functional cycle during runtime interaction (i.e., not preprogrammed, such as its schedule), which can be described as sensing a crash, stopping, turning, and going in another direction.

The knowledge gained from this way of analysing will be best used to determine how well a robotic artefact currently fits into its intended environment because it makes obvious both what it will and will not be able to respond to within it. By analysing the robotic artefact's functional cycles, the structure of its datanomous decisions can be made intelligible to humans. Similar to how the biologist partakes in participant observation when imagining the Umwelt of other organisms, the designer partaking in participant observation of an artificial Umwelt should ask themselves some questions. The biologist should, according to Uexküll (cited by Brentari 2015, p. 80) ask themselves questions such as: "which parts of the world are accessible to animals?" and, "which qualities of the objects surrounding us have an influence on the meaning organs of single animal species?"

The questions appropriate for robot development will naturally be different and focus not on understanding a way of life, but on achieving mutual intelligibility between interactants. Keeping the meaningful objects and their functional cycles central, the procedure can elucidate what and how in the environment the robotic artefact will be able to discern, and perhaps more importantly, what it will *not* be able to discern. Relevant questions can be akin to these suggestions: "What part of the users' actions or behaviours are accessible to the robotic artefacts? What can the user do that will be meaningful, and what is an appropriate functional cycle in response? What can be misinterpreted, and what is an appropriate or inappropriate functional cycle?" These questions focus on whether the functional cycles currently existing in the interactional procedure of the robotic artefact resonate with its intended purpose of use.

7.3 MOVEMENT ACTS IN A PRESENCE-BASED SOCIALITY

That the robotic artefact collects, interprets, and acts on sensory data—and does so without human intervention—makes it appear functionally independent as long as its environment is constrained sufficiently in regard to the contextual changes it is able to handle. It is exactly this functional independence, and the subsequent expression of the robotic artefact, that originally sparked the wonder that has directed the phenomenological gaze maintained in this thesis. In Soma (2020), I attempt to *language the experience* (Sheets-Johnstone 2011, p. 466) of encountering a robot lawnmower as it is carrying out its designated task. In the paper, I elucidate the conflicting feeling this encounter elicited in me:

> When I look at the robot lawnmower, I see the wheels spinning, moving the robot across the lawn. They are not being spun by any one or any other external force, it is unattached to wires that might drag it, untouched by a person who could push it. The robot spins the wheels. It 'turns' around, 'finding' a new direction and 'follows' that path. I become aware that whenever I describe what is physically taking place as I observe robot the lawnmower's mechanical movements, I cannot escape using verbs, semantically implying someone doing something, such as the robot 'coming' toward me, or 'going' away from me (Soma 2020, p. 106).

What I describe here is that I *see* an artefact—a machine, a non-living and inanimate object—that looks like *it knows* what *it's doing*. Yet, I am convinced that this *sight* does not resonate with my knowledge of what a robotic artefact *is*. Herein lies the specific contradiction subject to my phenomenological gaze. The robot is exactly an object in motion, not an animate being. Its movements are simply dislocation in space. *It* does not have a phenomenological kinesthetic experience of the movements it is executing. Nonetheless, my description utilises an intentionalist vocabulary to describe actions or behaviours, rather than attempt to describe the robot's movement in terms of pure physics or vectorial changes. I do so on purpose, as I hold it is not trivial to my experience of perceiving these movements that they belong to an independently moving entity.

I understand what I describe here to be a mode of encounter. To illuminate, I contrast it to Heidegger's (1962) *readiness-to-hand*, a mode of encounter where objects withdraw from the world and merge with the subject when it is skilfully manipulated as equipment (Wheeler 2014). Readiness-to-hand describes how a subject's experience merges with the object. In this mode of encounter, the object is essentially an extension of their intentionality—enabling both acting and perceiving *through* it as if it were an inseparable part of themselves. The subject and the object become *one*. As explained at length in Chapter 6, Sheets-Johnstone explicates animate beings as attuned to animation, intuitively finding animate movement meaningful. I further explore her notion of kinesthetic intercorporeality and characterise it as a presence-based sociality. The peculiar thing is, of course, that the robotic artefact is no other, but an object. This means that I cannot share a kinesthetic intercorporeality with it *as such*. However, by viewing it as a mode of encounter, I call attention to the realisation that *my* only way of relating to these movements nonetheless comes from my corporeal knowledge. The meaningfulness of the perceived movement of others comes from a place of kinesthetic intercorporeality. Rather than engaging in a skilful merge with the robotic artefact as I would were it to be a piece of *equipment*, I become engaged in a skilful kinesthetic navigation with this object as another. The functional independence of robotic artefacts establishes the entity precisely as an *other* in relation to myself. In this mode of encounter, the subject and the object exactly become *two.*

A parallel perspective to a presence-based sociality has been addressed in the context of human–robot interactions by Seibt et al. (2020). They present Type(s) of Experienced Sociality (TES), a framework that encompasses a "classification of complex phenomenological contents which in first approximation can be characterized as feelings of co-presence or 'being-with'." (2020, p. 59). What they describe here are feelings of co-presence that *approximate* "what it feels like to be with agents with different responsive capacities" (2020, p. 59) such as humans or other species of animals. The basic level of this framework acknowledges that a sense of *being-with* can arise even when the felt sociality is highly asymmetrical. Seen in relation to kinesthetically rooted corporeal knowledge, I hold that the conflicting feeling described in Soma (2020) is rooted in the two different kinds of thinking—thinking in movements and thinking in words—as described by Sheets-Johnstone (2011, p. 436).

To specify further, I mean to say that what I think about the robot through my linguistic terms differs from what I think about it through my corporeal concepts. In a presence-based sociality, *presence* is in itself meaningful. Exactly this sentiment is reflected by the first axiom of communication by Watzlawick et al. (1967), who state: "First of all, there is a property of behaviour that could hardly be more basic and is, therefore, often overlooked: behaviour has no opposite. In other words, there is no such thing as non-behaviour, to put it even more simply: one cannot *not* behave" (p. 29). This axiom, together with well as Searle's *speech acts* (1969), inspired what we specify as *movement acts* in Schulz et al. (2021). Movement acts acknowledge, also in the context of human–robot interactions, there are multiple layers of meaning available in every single act or doing (e.g., Rommetveit 1980). This is echoed by a parallel to ethology highlighted by Korcsok and Korondi (2023). Within ethology, *inaction* is considered to be a kind of behaviour, rather than a non-behaviour. As such, a robotic artefact in "stand-by" mode is currently unmoving, but not without behaviour.

To arrive at this term, we analysed qualitative comments made during a series of experiments with a Fetch robot[1]. The original intention of the experiment (described in Schulz et al. 2019) was to look at one animation principle, slow in and slow out, and see how it affected people's perception of the robot. In the experiment, the Fetch robot moved between three locations. The robot's actions were partially conducted using Wizard of Oz; Fetch was in charge of calculating its own path to its next location but did not make the decision about *when* to move or where its next location was. The interactions between the Fetch and the participants were simple. Fetch moved toward the participant, who was sitting designated spot. The first time the Fetch approached them, they would place a cup inside a paper bag attached to it. The second time, they would remove a questionnaire, fill it out, and place it into the paper bag again. Sporadically throughout the study, Fetch would be unable to calculate its path to the next location. When this happened, it found itself to be stuck—even if there was never anything blocking its path. To make itself unstuck, it went into recovery mode. This specific procedure involved the robot turning once, slowly, around its own axis. It would resume normal function after this. It was primarily this phenomenon we wanted to study.

We went into the analysis primarily interested in what, if anything, the participants had expressed about Fetch's recovery procedure. However, during analysis, we found that not only the recovery procedure, but that also Fetch's shorter and longer pauses would elicit commentary on what they surmised Fetch was, for instance, *doing, thinking,* or *feeling*. Some people would drastically change the way they described Fetch after having witnessed the recovery procedure, while others remained indifferent to the incident. Some would focus on Fetch's short or longer pauses and explain, for instance, how they felt it made the robot appear *confused*. In light of the first axiom of communication (Watzlawick et al. 1967), we arrived at the conclusion that *all* of Fetch's movements were instances of *movement acts*. Consequently, the term—intended to cover all kinds of robotic movement—does not discriminate between movements implemented exclusively for interactional purposes (movements intended to facilitate and engage users in interaction) and those meant only to have functional purposes (movements necessary for the robotic artefact to carry out the task to which it is put).

In Schulz et al. (2021), we primarily discuss movement acts in terms of what the Fetch robot seems to *communicate* to the participants in the study. Specifically, movement acts can be used to identify what we distinguish as intended and unintended communication between Fetch and the participants[2]. While I do not think that it is incorrect, per se, I find the term communication too narrow for the extensive range of possible meanings robotic movement acts may encompass. What a movement act will mean during interaction cannot be fully predicted because, as discussed earlier, meaning cannot be predetermined. This particular notion is central to speech act theory, which dictates that the words in themselves are not enough to understand the speech *act* in itself. Movement acts highlight that presence is itself meaningful, also in human–robot interactions. Thus, a robot's movement act will be expressive beyond its specific displacement in space. Indeed, movement acts underscore the recent empirical realisation about movement highlighted by Brincker (2016) that "our movements contain much more information than what has often been assumed in the philosophy of mind" (p. 450). Nonetheless, in light of Sheets-Johnstone's phenomenology of movement, also this description becomes an enormous understatement of the meaningfulness inherent in movement. By paying close attention to the unique perspective on the primordial significance of movement presented by Sheets-Johnstone, we may appreciate that movement is more than merely displacement in space.

Similar to how the interpretations of speech acts are situationally contingent on a socio-cultural context, so are the expressions of movement acts. In this sense, expressivity is not a particular feature or property of the robotic artefact, but a quality arising through the relationship between those involved in an interactional situation. To find a more accurate way of describing the relationship between displacement and the unfolding nature of movement, I now take the first step in outlining *kinetic affordances*.

7.4 OUTLINING KINETIC AFFORDANCES

One of Sheets-Johnstone's (2011) essential insights is that animate forms experience in the form of self-movement. Thus, self-movement is neither a means of perception nor in service of it. While Sheets-Johnstone commends Gibson and his ecological approach (1966, 1979) for recognising the former, she criticises him for overlooking the latter. According to her, Gibson is unable to escape the natural attitude view on movement. She further explains how movements are, in Gibson's perceptual systems, made known to the active observer only through affordances. In a quite offhand remark, she suggests that Gibson might as well have called movement *kinetic affordance*: "Though reduced to locomotion in service of perception, movement is what Gibson might well have termed 'kinetic affordance'" (Sheets-Johnstone, 2011, p. 203). I read into her comment that Sheet-Johnstone intends this "kinetic affordance" to be ecological information about the self and one's own locomotion and position in the environment. Self-movement, as it is conceived by Gibson, is only disclosed to experience in terms of the subject's relation to its environment. This perspective overlooks what Sheets-Johnstone recognise about the primacy of movement, namely, that it is experience itself. Animate forms are not familiarised with their own movements primarily through their perceptual systems but through a felt sense of kinesthesia. We, animate beings, precisely feel self-movement as *our own*. Gibson's perspective on movement—and indeed everyone else instrumentalising movement—has basically an *outside* perspective of it. Rather than a kinesthetic experience, it exists merely as a kinetic expression. It is in exactly this manner I believe *kinetic affordances* will be useful to human–robot interactions.

In the following quote—a paragraph recited in its entirety due to its descriptive precision—Sheets-Johnstone gives us a clue as to what kind of information could be structured in kinetic affordances.

> Let us imagine ourselves walking with resolute step. We find in this way of walking a tensional quality that is taut and hard. We have a sense of our bodies and our moving gait as firm and strong. We find projectional quality that we might describe in terms of a sharp and even striding, or a flat and heavy clumping; in either case, our projection of force is measured, unhesitant, deliberate. We find linear qualities describable in terms of strait-line bodily contours and straight-line paths of movement, undeviating direct in each instance. We find amplitudinal qualities describable in terms of a controlled but unconstrained bodily spatiality, that is, a controlled but unimpeded range of movement as we carve an unobstructed space. All of these qualities coalesce in the global phenomenon we imagine: "walking with resolute step." Together they articulate an overall spatio-temporal dynamic, a dynamic that coincides with the intended image: "walking with resolute step." Accordingly, the dynamic is there in the imagined movement. Similarly, when we actually walk with resolute step, the dynamic is there in the actual movement. An examination of our own experience thus demonstrates to us that no configuration of qualities exist apart from its creation: there is no firm and strong tensional quality, no sharp and even striding, no straight-line designs and patterns, no controlled but unimpeded amplitudes short of their imaginary or perceptual instantiation in movement. In actually walking with resolute step, we can sense ourselves creating this spatio-temporal dynamic and attend specifically to any of its qualities; any time we care to turn our attention to them, they are there. We find, then, that in moving, we bring a certain play of forces to life and spatialise and temporealise them in the process. An overall dynamic with distinctive qualities are created by our movement and experienced in our kinesthetic consciousness of movement. (Sheets-Johnstone 2011, pp. 126-7)

The exclusive focus of Sheets-Johnstone on self-movement experience itself is clearly incompatible with robotic movement, as robotic artefacts lack the capacity for bodily felt experiences, and in turn for kinesthetic qualities. However, in this attempt to outline kinetic affordances, our interests are found in another perspective. Sheets-Johnstone's phenomenology of movement demonstrates that the kinesthetic qualities of movement, experienced from a first-person perspective, constitute the foundation for animate beings' understanding of themselves. Indeed, it is possible to imagine what it means to walk with resolute steps because we, humans, corporeally know what it means to do so. Even if imagining itself does not give rise to the spatio-temporal dynamics, such as actual movement, it points to an important aspect of these dynamics. Namely, various ways of moving have their own distinct dynamic quality, a *kinetic expression*. In this manner, kinesthetic qualities also constitute the foundation for our understanding of ourselves in the world and in relation to others. This is particular to kinesthetic intercorporeality. The experiential dimension of self-movement cannot be ignored when attempting to understand how we relate ourselves to the movement of others. Corporeal knowledge of kinesthetic qualities is essential in the mode of encounter pertaining to human–robot interaction, as the robotic artefact establishes itself as another *to me* by virtue of its kinetic expression. I am proposing that kinetic affordances are meaningful in light of our presence-based sociality, as they are specified by the kinesthetic qualities we are experientially familiar with.

7.4.1 Cardinal structures of movement

Sheets-Johnstone specifically calls attention to how movement, as experience itself, creates an unfolding spatio-temporal dynamic. In the natural attitude view on movement, we are used to thinking of movement as taking place in space. Sheets-Johnstone, through her emphasis on movement as primarily unfolding, enable us to recognise that movements are not merely a displacement in space. While movements indeed need a space to unfold in, their primary characteristic is their unfolding nature. To characterise the unfolding spatio-temporal dynamics of movement, Sheets-Johnstone discloses four primary qualitative structures of movement, which she refers to as the *cardinal structures of kinesthetic experience*.

> Any movement has a certain *felt tensional quality, linear quality, amplitudinal quality,* and *projectional quality*. In a very general sense, the tensional quality has to do with our sense of effort; the linear quality with the felt linear contour of our moving body and the linear paths we sense ourselves describing in the process of moving; the amplitudinal quality with both felt expansiveness or contractiveness of our moving body and the spatial extensiveness or constrictedness of our movement; the felt projectional quality with the way in which we release force or energy. (Sheets-Johnstone 2011, p. 123 [emphasis added])

What she describes here has to do with force or effort, with space, and with time. About the way these structures intertwine, Sheets-Johnstone explains that spatial aspects are described by linear and amplitudinal qualities, while temporal aspects are described by the combination of tensional and projectional qualities. Especially the temporal dynamics are a result of how tensional and projectional qualities combine. The kinetic expression is determined by the intensity (or intensities) of these temporal qualities. In this way, they pertain to the complex spatio-temporal dynamic of movement (Sheets-Johnstone 2011, p. 123-4). Hence, these four cardinal structures are the way we, animate forms, are intimately familiarised with our movements.

Below, I attempt to sort these four cardinal structures of movement (Table 7.1) and explain each in isolation. To go with these explanations, I suggest a "guiding question" intended to capture the essence of each cardinal structure.

Spatial dynamics	*Linear qualities* — Directional progression of movement *Amplitudinal qualities* — Range of movement
Temporal dynamics	*Tensional qualities* — The displayed force of movement *Projectional qualities* — Purpose in or of movement

Table 7.1: Sorting Sheets-Johnstone's (2011) cardinal structures of movement. It should be noted that these four cardinal structures of movement qualities are highly intertwined and that in the actual spatio-temporal dynamics of experience, they will arise collectively.

However, I want to be explicit that these four structures will not appear separated in this sense in our experience. "These qualitative aspects of movement are of course separable only reflectively, that is, analytically, after the fact; experientially, they are all a piece in the global qualitatively felt dynamic phenomenon of self-movement" (Sheets-Johnstone 2011, p. 123). Further, I do not mean to say that the experiential dimension of these cardinal qualitative structures can be programmed to elicit a phenomenological experience in the robotic artefact. Self-movement *creates* these qualities in animate beings, and the movements of robotic artefacts do not. However, if it is indeed so that kinesthetic qualities are discernible in kinetic expressions, we intuitively navigate using these cardinal structures of movement in the mode of encounter pertaining to kinesthetic intercorporeality. In a presence-based sociality, the kinetic expression of another moving entity, also a robotic artefact, will be relevant to us as a *relation*.

Linearity

In the natural attitude view, this constitutes the most obvious quality of movement as it speaks directly of spatial displacement. In light of *movement acts*, no linear progression will also be relevant. In this manner, this cardinal structure is important in highlighting that *presence* is meaningful in itself. Relationally, a movement's directionality will be important in determining whether the entity is moving toward or away from the observer. Further, a transition from a forward (or backward) linear progression to no linear progression, into a full stop or pause, and perhaps transitioning again into a change of direction will be relationally significant. It shows that also linear progression must be tightly interwoven with temporality. Direction pertains to spatial dynamics, but the progression itself is necessarily temporally specified.

Guiding question: *"Where is it going?"*

Amplitude

The range of a movement is nearly indistinguishable from its linear qualities. However, as the entity directionally moves towards or away from the observer, the extent of its reach holds significance in terms of the entity's directionality *in relation* to them. This particular relation pertains to whether the observer will be within the moving entity's reach.

Guiding question: *"What can the entity reach through its movement?"*

Tension

The tensional quality speaks about the apparent forces involved in the movements, for instance as speed or acceleration. While tension primarily pertains to the expression of the forces involved, it matters for interaction in combination with projection. Through this combination, the observer will be able to discern how the forces matter for the outcome of the movements. Hence, it may also speak about the precision of the current movement. High effort often leaves less room for precision, and conversely, low intensity allows for greater precision.

Guiding question: *"How forceful does the movement appear?"*

Projection

As a qualitative dimension of temporal dynamics, it is in combination with tension that projection gains significance, as it relates to the projection of the forces involved. One may also presume that, in interactional situations, the perceived projectionality of the movement is highly contingent on context as we use sociomaterial environmental information to discern its purpose. For animate beings, projectionality obviously pertains to how its intentionality comes to show, as the forces involved are turned-toward various aspects of the environment.

Guiding question: *"What is the movement toward (in the sense of purpose)?"*

7.4.2 Applying kinetic affordances to Fetch

These are the cardinal qualities that structure *kinetic affordances*. While I have here attempted to explain them in isolation, it is only together they form an affordance. Moreover, by proposing that the meaningfulness in the movement of others is perceivable through kinetic affordances, I contend they must be structured by the cardinal qualities of movements. I will try to exemplify how these qualities are discernible through four of Fetch's movement acts, all of which took place during the series of experiments described in Schulz et al. (2021): pausing, spinning, moving straight with a linear velocity profile, and moving straight with a *slow in, slow out* velocity profile.

When Fetch pauses, the tensional quality might be characterised as low. Simultaneously, its linear progression is none, and thus this particular low tension may not come across as projected toward anything specific. In turn, the amplitudinal quality makes it clear that nothing beyond the robot's current location will be within its reach. Naturally, one cannot predict exactly which behaviours the kinetic affordance of this movement will offer, as it is subjectively and situationally contingent. However, when Fetch pauses, I may be compelled to go check on it, or I may be compelled to wait. Perhaps I will simply ignore it.

As Fetch spins around its own axis, the linear progression is still none. And while the tension may still not seem to be directed at anything specific, the projectionality of this non-progressional yet tensional movement will be unclear in a different manner. Rather than not being directed towards anything in particular, the spinning may appear to have a specific projection, but exactly what it is remains unclear to me as an onlooker. In an interactional situation, I may be curious as to why it spins, and go check on it. Perhaps I am frightened and decide it will be best to get away from it.

In either of these instances, whether Fetch is pausing in its trajectory to the next location, or spinning around in one spot, the influence of the implemented velocity profiles on the linear quality is limited. However, when Fetch is moving trouble-free across the room, both velocity profiles will give the tensional quality a different expression. However, it does so only when considered in combination with the projectional quality I become aware of as the overall movement unfolds. The importance of the amplitudinal quality becomes clear when both velocity profiles are compared.

When Fetch moves using a linear velocity profile—in other words, with constant speed—the projectionality of the tension is less clear, as I am unable to discern what it is moving toward. Thus, the apparent amplitude makes itself highly relevant to the overall relation between Fetch and me. When both linear progression and tension is unchanging, while projectionality is unclear, it becomes of greater interest to me whether I am within the movement's current reach. In other words, if it looks like Fetch is on a collision course toward me, I am compelled to move myself away. In contrast, when Fetch is moving toward me using a *slow in, slow out* velocity profile, the movements' projectional quality changes because the slowing out—speeding down—will make it clear to me that Fetch will not be on a collision course with me. By gradually gaining speed, and especially by gradually slowing down toward the end of its trajectory, I will not feel compelled to move. Thus, in an interactional setting where the robot is heading straight toward me, the overall kinetic affordance will be slightly different in the two velocity profiles.

All four cardinal structures can be identified in all of Fetch's movements; however, we find that their interplay differs in the various ways of moving. Thus, it becomes obvious that it is not possible to describe the kinetic affordance without involving all four cardinal structures of movement. Indeed, the two different velocity profiles implemented in Fetch during the experiments will give the robot two different tensional expressions throughout its moving from start to stop.

These structures are highly intertwined in experience and will thus be highly intertwined in kinetic affordances. However, it also becomes clear that their interplay will give rise to different kinetic affordances. By virtue of being *relations* between the subject and the object they are encountering as another, kinetic affordances tell the perceiver something about the movement and the entity the movements belong to *in relation* to themselves. In perceiving kinetic affordances, the subject enters a mode of encountering the moving entity as another, thereby relating its movement to itself.

7.5 SEEING HUMAN–ROBOT INTERACTIONS THROUGH A PHENOMENOLOGICAL GAZE

In this thesis, I have demonstrated what it means to hold a *phenomenological gaze* on human–robot interactions. The study maintains a phenomenological gaze on a specific experience—the experience of *seeing* a robot lawnmower as animate, while simultaneously *knowing* it to be inanimate. By approaching the phenomenon of interest with curiosity, the phenomenological gaze widened the explorative space of this study. Rather than maintaining a specific aim of explaining *why* I experience a robotic lawnmower as animate, I used a phenomenological perspective to make the familiar *unfamiliar*. This particular endeavour allowed me to change focus toward *how* and *what.* By centring my attention on *how* the experience takes place in my conscious experience (Soma 2020), I uncovered a difficulty with describing exactly *what* robot movement is in my experience of it—that I am compelled to use verbs rather than vectors, describing robot movement as *behaviours* rather than mere relocation in space. That robot movement is described as behaviour even if robotic artefacts are, strictly speaking, not able to *behave* bypasses the inherent social asymmetry of human–robot relations problematised by Seibt as "the description problem of human-robot interactions" (2016, p. 107).

I approached the description problem faced by HRI (Seibt 2016) with a specific experience as my point of departure. I explicated three distinct, but thematically related theories. I utilised Uexküll's Umwelt theory to illuminate the question of a robotic artefact's actual interactional capabilities. I explored Gibson's theory of affordances to better understand the relationship between meaning and behaviour, as a step toward grasping how mutual intelligibility may be achieved between humans and robotic artefacts during interaction. Lastly, I elucidated the aspects of Sheets-Johnstone's phenomenology and reached an understanding of human–robot interactions as a specific *mode of encounter*. Thus, I in this thesis introduce the relevance of a kinesthetic intercorporeality, what I characterise a presence-based sociality, to the field of HRI.

Finding ways of accurately describing human–robot interactions is important for ethical and moral issues within the HRI field (Seibt 2016). I would argue, that so is non-technical and available descriptions of the functional mechanics of robot technology. Thus, the theoretical contributions can be summarised in two interrelated, but slightly nuanced ways. In any interaction between humans and robotic artefacts, the *meaning* of robot movement resides in the relationship between the interactants. This relationship is highly contextual, and thus unpredictable outside the situation itself. Phenomenological concepts demonstrate that humans understand robotic technology through human corporeal concepts. This is the fundament for mutual intelligibility.

Because robots are not animate forms, they cannot experience, and thus we cannot ask them about their version of the environment. Nonetheless, as a metaphorical interactional partner, this perspective matters greatly for the mutual intelligibility of a human–robot interaction. Thus, the *artificial Umwelt* can be seen as an effort toward understanding robot technology with concepts easily relatable to human corporeal concepts. Building on this, the concept of *datanomy*, as a characteristic of the artificial Umwelt, facilitates a description of, say, why robots do now know anything about grass that can be understood even without the technical knowledge of robot technology. In this thesis discussion, I explain how the *artificial Umwelt* of the robotic artefact is *datanomous* and static in nature, making it clear that the robotic artefact is not always successful in its independent navigation in the environment. This perspective is important for mutual intelligibility when interacting with robots because it touches upon aspects of the balance between the apparent capabilities of a robotic artefact, and its actual abilities. When the robotic artefact's artificial Umwelt is unable to handle the unexpected changes in a dynamic environment, the appearance of its functional independence is revealed as precisely functionality, rather than capabilities. Specifically, we are made aware that the robotic artefact's context differs greatly from our own. Further, datanomy more accurately specifies the robot's foundation for understanding and navigating the environment. Thus, these concepts can aid designers in keeping both human-specific and robotic-specific perspectives more clearly in mind.

Further, the concept of *movement acts* highlights the perspective that the meaning of movement resides in the relation between the interactants. Through my discussion of *movement acts* and *kinesthetic intercorporeality*, I identify this presence-based sociality as a mode of encounter between humans and moving entities. In this manner is it a key aspect of human–robot interactions. By understanding human–robot interactions as a mode of encounter founded in a kinesthetic intercorporeality, it is surmised that we recognise and understand the robotic movement first and foremost through our intimate familiarity with self-movement. As all movements—and our kinesthetic experience of them—are structured by four cardinal qualities *amplitude*, *linearity*, *tension*, and *projection*, I contend that these four cardinal structures further specify what I describe as *kinetic affordances*. The nature of affordances is relational—kinetic affordances are the meaningful relation between an active observer and a moving entity. To summarise, I arrive at the concept of *kinetic affordances* by combining concepts developed and explored in the papers included in this thesis with concepts described and explored in this thesis.

Kinetic affordance is a way to describe robot movement in terms of relation to the interactants. It is based on (universal) human experience of movement based on the structures familiar to us. Rather than specifying exactly what the robot *does*, kinetic affordances can be used to describe its interactional expression in relation to us. This is descriptive without attributing animacy because the *meaning* of the movement resides in the relationship between humans and robotic artefacts.

Driven by a sense of genuine wonder, I have throughout this thesis been concerned with the mutual intelligibility of human–robot interactions, primarily focusing on how the interactional exchanges of robots may be intelligible for humans. In light of our mode of encountering moving entities coming from a place of kinesthetic intercorporeality, we naturally find robotic movement to possess animate qualities. The concepts *artificial Umwelt*, *datanomy*, *movement acts*, and *kinetic affordances* are all, in their own way, pertinent to elucidate the interplay between a robotic artefact's apparent capabilities, and its actual capabilities.

Notes

Chapter 1

1 Multimodal Elderly Care Systems (MECS) project, is funded by Research Council of Norway under grant agreement 247697

2 Other results from this study can be found in (Verne 2020). The results described by Verne (2020) were not published at the time of writing Turning Away from an Anthropocentric View on Robotics (Soma and Herstad, 2018).

Chapter 2

1 "Hanson Robotics' most advanced human-like robot, Sophia" https://www.hansonrobotics.com/sophia/

2 Jeppe Hein: 360° presence. ARoS October 2009 https://www.youtube.com/watch?v=95RZUuA-tJI (cited in Levillain and Zibetti 2017)

Chapter 3

1 While the subject will be considered outside the scope of this thesis, there is an entire direction of phenomenology springing out from the work of Don Ihde, post-phenomenology. Post-phenomenology focuses primarily on the mediating role of technology.

Chapter 4

1 Although I myself prefer "surrounding-world." This phrase preserves the semantic richness of the German and the prefix um, as it connotes apprehension from a centre.

2 An excellent introduction to the particularities of these concepts can be found in O'Neil's *Translaters Introduction* in the newest translation of *Streifzüge durch die Umwelten von Tieren und Menschen* (1934), *A Foray into the Worlds of Animals and Humans* (2010).

3 While Soma and Herstad (2018) was written using the republishing of the 1957 original translation, *A Stroll through the worlds of animals and men: A picture book of invisible worlds* (von Uexküll 1992), I will in this chapter use the terms from the newest translation, *A Foray into the Worlds of Animals and Humans* (von Uexküll 2010).

4 However, to my knowledge Umwelt theory is, little known in the HCI community and as far as I know, not been used to analyse design (as an analytical

tool meant to help identify how a robotic artefact perceives or senses the world) by anyone else.

5 I am inclined to say that while the artificial Umwelt can be imagined, but I believe the artificial Innenwelt cannot.

Chapter 6

1 While Sheets-Johnstone has published numerous volumes of collected works—books consisting of republished selected philosophical essays that all focus in some way on the specified topic—the focus on this section will be on the collection of works in the second edition of *The Primacy of Movement* (2011) in an attempt to limit the focus of this chapter.

2 The use of the pronouns "we" and "us" as a collective for animate forms will in this chapter be used interchangeably like it is by Sheets-Johnstone (2011).

3 Uexküll believed that one of the fundamental differences between living organisms and mechanisms was that there was no growth in mechanisms (Ziemke and Sharkey 2001, p. 708).

4 While modern dictionaries seem to prefer the spelling "kinaesthesia" I will consequently stick to "kinesthesia" as this is the spelling used by Sheets-Johnstone

Chapter 7

1 The analysis and analytical procedure are explained in detail in Schulz et al. (2021).

2 This specific dimension has been incorporated into a typology of robot signals in Holthaus et al. 2023).

List of references

Bartneck, Christoph. 2023. 'Godspeed Questionnaire Series: Translations and Usage'. Pp. 1–35 in.

Bartneck, Christoph, Dana Kulić, Elizabeth Croft, and Susana Zoghbi. 2009. 'Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots'. *International Journal of Social Robotics* 1(1):71–81. doi: 10.1007/s12369-008-0001-3.

Bianchini, Samuel, Florent Levillain, Armando Menicacci, Emanuele Quinz, and Elisabetta Zibetti. 2016. 'Towards Behavioral Objects: A Twofold Approach for a System of Notation to Design and Implement Behaviors in Non-Anthropomorphic Robotic Artifacts'. Pp. 1–24 in *Dance Notations and Robot Motion*, *Springer Tracts in Advanced Robotics*, edited by J.-P. Laumond and N. Abe. Cham: Springer International Publishing.

Bødker, Susanne. 2015. 'Third-Wave HCI, 10 Years Later---Participation and Sharing'. *Interactions* 22(5):24–31. doi: 10.1145/2804405.

Bradshaw, Jeffrey M., Robert R. Hoffman, David D. Woods, and Matthew Johnson. 2013. 'The Seven Deadly Myths of "Autonomous Systems"'. *IEEE Intelligent Systems* 28(3):54–61. doi: 10.1109/MIS.2013.70.

Bratteteig, Tone. 2021. *Design for, med og av brukere*. Universitetsforlaget.

Breazeal, Cynthia. 2003. 'Toward Sociable Robots'. *Robotics and Autonomous Systems* 42(3):167–75. doi: 10.1016/S0921-8890(02)00373-1.

Breazeal, Cynthia, Kerstin Dautenhahn, and Takayuki Kanda. 2016. 'Social Robotics'. Pp. 1935–72 in *Springer Handbook of Robotics*, *Springer Handbooks*, edited by B. Siciliano and O. Khatib. Cham: Springer International Publishing.

Brentari, Carlo. 2015. *Jakob von Uexküll: The Discovery of the Umwelt between Biosemiotics and Theoretical Biology*. Vol. 9. Dordrecht: Springer Netherlands.

Brincker, Maria. 2016. 'Dynamics of Perceptible Agency: The Case of Social Robots'. *Minds and Machines* 26(4):441–66. doi: 10.1007/s11023-016-9405-2.

Brooks, Rodney A. 1991. 'Intelligence without Representation'. *Artificial Intelligence* 47(1–3):139–59.

Bruineberg, Jelle, Anthony Chemero, and Erik Rietveld. 2019. 'General Ecological Information Supports Engagement with Affordances for "Higher" Cognition'. *Synthese* 196(12):5231–51. doi: 10.1007/s11229-018-1716-9.

Burneleit, Eva, Fabian Hemmert, and Reto Wettach. 2009. 'Living Interfaces: The Impatient Toaster'. Pp. 21–22 in *Proceedings of the 3rd International Conference on Tangible and Embedded Interaction*. Cambridge United Kingdom: ACM.

Čapek, Karel. 1920. *R.U.R.*

Carpinella, Colleen M., Alisa B. Wyman, Michael A. Perez, and Steven J. Stroessner. 2017. 'The Robotic Social Attributes Scale (RoSAS): Development and Validation'. Pp. 254–62 in *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction - HRI '17*. Vienna, Austria: ACM Press.

Coeckelbergh, Mark. 2011. 'Humans, Animals, and Robots: A Phenomenological Approach to Human-Robot Relations'. *International Journal of Social Robotics* 3(2):197–204. doi: 10.1007/s12369-010-0075-6.

Costall, Alan. 1995. 'Socializing Affordances'. *Theory & Psychology* 5(4):467–81. doi: 10.1177/0959354395054001.

Damiano, Luisa, and Paul Gerard Dumouchel. 2020. 'Emotions in Relation. Epistemological and Ethical Scaffolding for Mixed Human-Robot Social Ecologies'. *HUMANA.MENTE Journal of Philosophical Studies* 13(37):181–206.

Danto, Arthur C. 1999. *The Body/Body Problem: Selected Essays*. Berkeley: University of California Press.

Dautenhahn, Kerstin. 1997. 'I COULD BE YOU: THE PHENOMENOLOGICAL DIMENSION OF SOCIAL UNDERSTANDING'. *Cybernetics and Systems* 28(5):417–53. doi: 10.1080/019697297126074.

Dautenhahn, Kerstin. 2014. 'Human-Robot Interaction'. Retrieved 6 May 2024 (https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/human-robot-interaction).

Devin, Sandra, Camille Vrignaud, Kathleen Belhassein, Aurélie Clodic, Ophélie Carreras, and Rachid Alami. 2018. 'Evaluating the Pertinence of Robot Decisions in a Human-Robot Joint Action Context: The PeRDITA Questionnaire'. Pp. 144–51 in *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*.

Dewey, John. 1980. *Art As An Experience*. Perigee.

Dourish, Paul. 2001. *Where the Action Is: The Foundations of Embodied Interaction*. MIT Press.

Dourish, Paul. 2004. 'What We Talk about When We Talk about Context'. *Personal and Ubiquitous Computing* 8(1):19–30. doi: 10.1007/s00779-003-0253-8.

Dreyfus, Hubert L. 1965. 'Alchemy and AI'. *RAND Corporation*.

Dreyfus, Hubert L. 1986. *What Computers Can't Do: The Limits of Artificial Intelligence*. Rev. ed. 7th print. New York: Harper & Row.

Duarte, Emanuel Felipe, and M. Cecilia Baranauskas. 2016. 'Revisiting the Three HCI Waves: A Preliminary Discussion on Philosophy of Science and Research Paradigms'. Pp. 1–4 in.

Ehn, Per. 1988. 'Work-Oriented Design of Computer Artifacts'.

Emmeche, Claus. 2001. 'Does a Robot Have an Umwelt? Reflections on the Qualitative Biosemiotics of Jakob von Uexküll'. 2001(134):653–93. doi: 10.1515/semi.2001.048.

Fong, Terrence, Illah Nourbakhsh, and Kerstin Dautenhahn. 2003. 'A Survey of Socially Interactive Robots'. *Robotics and Autonomous Systems* 42(3):143–66. doi: 10.1016/S0921-8890(02)00372-X.

Fong, Terrence, Charles Thorpe, and Charles Baur. 2001. 'Collaboration, Dialogue, Human-Robot Interaction.' Pp. 255–66 in.

Forlizzi, Jodi. 2007. 'How Robotic Products Become Social Products: An Ethnographic Study of Cleaning in the Home'. Pp. 129–36 in *Proceedings of the ACM/IEEE International Conference on Human-robot Interaction*, *HRI '07*. New York, NY, USA: ACM.

Formosa, Paul. 2021. 'Robot Autonomy vs. Human Autonomy: Social Robots, Artificial Intelligence (AI), and the Nature of Autonomy'. *Minds and Machines*. doi: 10.1007/s11023-021-09579-2.

Frauenberger, Christopher. 2019. 'Entanglement HCI The Next Wave?' *ACM Transactions on Computer-Human Interaction* 27(1):2:1-2:27. doi: 10.1145/3364998.

Gasser, Georg. 2021. 'The Dawn of Social Robots: Anthropological and Ethical Issues'. *Minds and Machines* 31(3):329–36. doi: 10.1007/s11023-021-09572-9.

Gaver, William W. 1991. 'Technology Affordances'. Pp. 79–84 in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, *CHI '91*. New York, NY, USA: Association for Computing Machinery.

Gelman, Rochel. 1990. 'First Principles Organize Attention to and Learning about Relevant Data: Number and the Animate-Inanimate Distinction as Examples'. *Cognitive Science* 14(1):79–106. doi: 10.1016/0364-0213(90)90027-T.

Gelman, Rochel, Frank Durgin, and Lisa Kaufman. 1996. 'Distinguishing between Animates and Inanimates: Not by Motion Alone'. Pp. 150–84 in *Causal Cognition*, edited by D. Sperber, D. Premack, and A. J. Premack. Oxford University Press.

Gibson, James J. 1966. *The Senses Considered as Perceptual Systems*. Boston: Houghton Mifflin Company.

Gibson, James J. [1979] 2015. *The Ecological Approach to Visual Perception*. Classic edition. New York; Hove, England: Psychology Press.

Goodman, Russell. 2017. 'William James'. in *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta. Metaphysics Research Lab, Stanford University.

Goodrich, Michael A., and Alan C. Schultz. 2007. 'Human-Robot Interaction: A Survey'. *Foundations and Trends® in Human-Computer Interaction* 1(3):203–75. doi: 10.1561/1100000005.

Gunn, Janet Varner. 1998. 'Review of The Roots of Thinking; The Roots of Power: Animate Form and Gendered Bodies, Maxine Sheets-Johnstone'. *Hypatia* 13(3):177–81.

Harrison, Steve, Phoebe Sengers, and Deborah Tatar. 2011. 'Making Epistemological Trouble: Third-Paradigm HCI as Successor Science'. *Interacting with Computers* 23(5):385–92. doi: 10.1016/j.intcom.2011.03.005.

Heidegger, Martin. 1962. *Being and Time*. New York, NY: Harper & Row.

Heider, Fritz, and Marianne Simmel. 1944. 'An Experimental Study of Apparent Behavior'. *The American Journal of Psychology* 57(2):243. doi: 10.2307/1416950.

Hewett, Thomas, Ronald Baecker, Stuart Card, Tom Carey, Jean Gasen, Marilyn Tremaine, Gary Perlman, Gary Strong, and William Verplank. 1992. 'ACM SIGCHI Curricula for Human-Computer Interaction'. doi: 10.1145/2594128.

Ho, Chin-Chang, and Karl F. MacDorman. 2010. 'Revisiting the Uncanny Valley Theory: Developing and Validating an Alternative to the Godspeed Indices'. *Computers in Human Behavior* 26(6):1508–18. doi: 10.1016/j.chb.2010.05.015.

Holthaus, Patrick, Trenton Schulz, Gabriella Lakatos, and Rebekka Soma. 2023. 'Communicative Robot Signals: Presenting a New Typology for Human-Robot Interaction'. Pp. 132–41 in *Proceedings of the 2023 ACM/IEEE International Conference on Human-Robot Interaction*. Stockholm Sweden: ACM.

Hornecker, Eva, and Jacob Buur. 2006. 'Getting a Grip on Tangible Interaction: A Framework on Physical Space and Social Interaction'. Pp. 437–46 in *Proceedings of the SIGCHI conference on Human Factors in computing systems*. ACM.

Ihde, D. 1990. *Technology and the Lifeworld, from Garden to Earth*. Bloomington: Indiana University Press.

James, William. 1890. *The Principles of Psychology*. Henry Holt and Company.

Kahn, Peter H., Aimee L. Reichert, Heather E. Gary, Takayuki Kanda, Hiroshi Ishiguro, Solace Shen, Jolina H. Ruckert, and Brian Gill. 2011. 'The New Ontological Category Hypothesis in Human-Robot Interaction'. Pp. 159–60 in *Proceedings of the 6th international conference on Human-robot interaction*. ACM.

Käufer, Stephan, and Anthony Chemero. 2015. *Phenomenology: An Introduction*. Cambridge ; New York: Polity.

Korcsok, Beáta, and Péter Korondi. 2023. 'How Do You Do the Things That You Do? Ethological Approach to the Description of Robot Behaviour'. *Biologia Futura* 74(3):253–79. doi: 10.1007/s42977-023-00178-z.

Levillain, Florent, and Elisabetta Zibetti. 2017. 'Behavioral Objects: The Rise of the Evocative Machines'. *Journal of Human-Robot Interaction* 6(1):4. doi: 10.5898/JHRI.6.1.Levillain.

Loke, Lian, and Toni Robertson. 2013. ‘Moving and Making Strange: An Embodied Approach to Movement-Based Interaction Design’. *ACM Transactions on Computer-Human Interaction* 20(1):1–25. doi: 10.1145/2442106.2442113.

Luo, Yuyan, Lisa Kaufman, and Renée Baillargeon. 2009. ‘Young Infants’ Reasoning about Physical Events Involving Inert and Self-Propelled Objects’. *Cognitive Psychology* 58(4):441–86. doi: 10.1016/j.cogpsych.2008.11.001.

Malle, Bertram F. 2011. ‘Attribution Theories: How People Make Sense of Behavior’. Pp. 72–95 in *Theories in social psychology*. Hoboken, NJ, US: Wiley Blackwell.

van Manen, Max. 2014. *Phenomenology of Practice: Meaning-Giving Methods in Phenomenological Research and Writing*. Walnut Creek, California: Left Coast Press.

Maturana, Humberto R., and Francisco J. Varela. 1980. *Autopoiesis and Cognition: The Realization of the Living*. Dordrecht, Holland ; Boston: D. Reidel Pub. Co.

McGrenere, Joanna, and Wayne Ho. 2000. ‘Affordances: Clarifying and Evolving a Concept’. Pp. 179–86 in *Proceedings of the Graphics Interface 2000 Conference*.

Michaels, Claire F. 2003. ‘Affordances: Four Points of Debate’. *Ecological Psychology* 15(2):135–48. doi: 10.1207/S15326969ECO1502_3.

Michotte, A. 1963. *The Perception of Causality*. Basic Books.

Mutlu, Bilge, Nicholas Roy, and Selma Šabanović. 2016. ‘Cognitive Human–Robot Interaction’. Pp. 1907–34 in *Springer Handbook of Robotics*, *Springer Handbooks*, edited by B. Siciliano and O. Khatib. Cham: Springer International Publishing.

Myers, Michael. 1997. ‘Qualitative Research in Information Systems’. Retrieved 13 May 2015 (http://www.qual.auckland.ac.nz/).

Norman, Donald A. 1988. *The Psychology of Everyday Things*. New York: Basic Books.

Robertson, Toni. 1997. ‘Cooperative Work and Lived Cognition: A Taxonomy of Embodied Actions’. P. 205 in *Proceedings of the fifth conference on European Conference on Computer-Supported Cooperative Work*.

Robertson, Toni, and Lian Loke. 2009. ‘Designing Situations’. Pp. 1–8 in *Proceedings of the 21st Annual Conference of the Australian Computer-Human Interaction Special Interest Group: Design: Open 24/7*, *OZCHI ’09*. New York, NY, USA: Association for Computing Machinery.

Rommetveit, Ragnar. 1980. ‘On “Meanings” of Acts and What Is Meant and Made Known by What Is Said in a Pluralistic Social World’. Pp. 108–49 in *The Structure of action*, edited by M. Brenner. Oxford: Blackwell.

Sagan, Dorion. 2010. ‘Introduction: Umwelt after Uexküll’. Pp. 1–34 in *A foray into the worlds of animals and humans: with A theory of meaning*, *Posthumanities*. Minneapolis: University of Minnesota Press.

Sanders, John T. 1997. ‘An Ontology of Affordances’. *Ecological Psychology* 9(1):97–112. doi: 10.1207/s15326969eco0901_4.

Santazilia, Ekaitz. 2020. ‘The Two Faces of Animacy’. *Studies in Language* 44(4):812–30. doi: 10.1075/sl.19089.san.

Schaefer, Kristin E. 2016. ‘Measuring Trust in Human Robot Interactions: Development of the “Trust Perception Scale-HRI”’. Pp. 191–218 in *Robust Intelligence and Trust in Autonomous Systems*, edited by R. Mittu, D. Sofge, A. Wagner, and W. F. Lawless. Boston, MA: Springer US.

Scholl, Brian J., and Patrice D. Tremoulet. 2000. ‘Perceptual Causality and Animacy’. *Trends in Cognitive Sciences* 4(8):299–309. doi: 10.1016/S1364-6613(00)01506-0.

Scholtz, J. 2003. ‘Theory and Evaluation of Human Robot Interactions’. P. 10 pp.- in *36th Annual Hawaii International Conference on System Sciences, 2003. Proceedings of the*.

Schulz, Trenton, Patrick Holthaus, Farshid Amirabdollahian, Kheng Lee Koay, Jim Torresen, and Jo Herstad. 2019. ‘Differences of Human Perceptions of a Robot Moving Using Linear or Slow in, Slow out Velocity Profiles When Performing a Cleaning Task’. Pp. 1–8 in *2019 28th IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. New Delhi, India: IEEE.

Schulz, Trenton, Rebekka Soma, and Patrick Holthaus. 2021. ‘Movement Acts in Breakdown Situations: How a Robot’s Recovery Procedure Affects Participants’ Opinions’. *Paladyn, Journal of Behavioral Robotics* 12(1):336–55. doi: 10.1515/pjbr-2021-0027.

Schulz, Trenton, Jim Torresen, and Jo Herstad. 2018. ‘Animation Techniques in Human-Robot Interaction User Studies: A Systematic Literature Review’. *arXiv:1812.06784 [Cs]*.

Searle, John R. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Seibt, Johanna. 2016. ‘" Integrative Social Robotics " -A New Method Paradigm to Solve the Description Problem And the Regulation Problem?’ Pp. 104–15 in.

Seibt, Johanna. 2018. ‘Classifying Forms and Modes of Co- Working in the Ontology of Asymmetric Social Interactions (OASIS)’ edited by M. Coeckelbergh, J. Loh, M. Funk, J. Seibt, and M. Nørskov. *Envisioning Robots in Society - Power, Politics, and Public Space - Proceedings of Robophilosophy 2018 / TRANSOR 2018* 133–46. doi: 10.3233/978-1-61499-931-7-133.

Seibt, Johanna, Christina Vestergaard, and Malene F. Damholdt. 2020. ‘Sociomorphing, Not Anthropomorphizing: Towards a Typology of Experienced Sociality’. in *Frontiers in Artificial Intelligence and Applications*, edited by M. Nørskov, J. Seibt, and O. S. Quick. IOS Press.

Sheets-Johnstone, Maxine. 2011. *The Primacy of Movement*. Expanded 2nd ed. Amsterdam ; Philadelphia: John Benjamins Pub. Co.

Sheridan, Thomas, W. Verplank, and T. Brooks. 1978. 'Human and Computer Control of Undersea Teleoperators'.

Sirkin, David, Brian Mok, Stephen Yang, and Wendy Ju. 2015. 'Mechanical Ottoman: How Robotic Furniture Offers and Withdraws Support'. Pp. 11–18 in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction - HRI '15*. Portland, Oregon, USA: ACM Press.

Smith, David Woodruff. 2018. 'Phenomenology'. in *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta. Metaphysics Research Lab, Stanford University.

Soltanzadeh, Sadjad. 2019. 'A Practically Useful Metaphysics of Technology'. *Techné: Research in Philosophy and Technology* 23(2):232–50. doi: 10.5840/techne2019924103.

Soma, Rebekka. 2020. 'Engaging in Deep Wonder at the Experience of Encountering a Lawnmower Robot'. in *Frontiers in Artificial Intelligence and Applications*, edited by M. Nørskov, J. Seibt, and O. S. Quick. IOS Press.

Soma, Rebekka, Tone Bratteteig, Diana Saplacan, Robyn Schimmer, Erik Campano, and Guri Verne. 2022. 'Strengthening Human Autonomy. In the Era of Autonomous Technology'. *Scandinavian Journal of Information Systems* 34(2).

Soma, Rebekka, and Jo Herstad. 2018. 'Turning Away from an Anthropocentric View on Robotics'. *Frontiers in Artificial Intelligence and Applications* 53–62. doi: 10.3233/978-1-61499-931-7-53.

Søraa, Roger, and Dr Fostervold. 2021. 'Social Domestication of Service Robots: The Secret Lives of Automated Guided Vehicles (AGVs) at a Norwegian Hospital'. *International Journal of Human-Computer Studies* 152:102627. doi: 10.1016/j.ijhcs.2021.102627.

Spatola, Nicolas, Barbara Kühnlenz, and Gordon Cheng. 2021. 'Perception and Evaluation in Human–Robot Interaction: The Human–Robot Interaction Evaluation Scale (HRIES)—A Multicomponent Approach of Anthropomorphism'. *International Journal of Social Robotics* 13(7):1517–39. doi: 10.1007/s12369-020-00667-4.

Stake, Robert E. 2005. 'Qualitative Case Studies'. Pp. 443–66 in *The Sage handbook of qualitative research, 3rd ed*, edited by N. K. Denzin and Y. S. Lincoln. Thousand Oaks, CA: Sage Publications Ltd.

Steinfeld, A. 2004. 'Interface Lessons for Fully and Semi-Autonomous Mobile Robots'. Pp. 2752-2757 Vol.3 in *IEEE International Conference on Robotics and Automation, 2004. Proceedings. ICRA '04. 2004*. New Orleans, LA, USA: IEEE.

Steinfeld, Aaron, Terrence Fong, David Kaber, Michael Lewis, Jean Scholtz, Alan Schultz, and Michael Goodrich. 2006. 'Common Metrics for Human-Robot

Interaction'. Pp. 33–40 in *Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction*, *HRI '06*. New York, NY, USA: Association for Computing Machinery.

Stoffregen, Thomas A. 2000. 'Affordances and Events'. *Ecological Psychology* 12(1):1–28. doi: 10.1207/S15326969ECO1201_1.

Stoffregen, Thomas A. 2003. 'Affordances as Properties of the Animal-Environment System'. *Ecological Psychology* 15(2):115–34. doi: 10.1207/S15326969ECO1502_2.

Suchman, Lucille Alice. 1987. *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge [Cambridgeshire] ; New York: Cambridge University Press.

Svanæs, Dag. 2000. 'Understanding Interactivity: Steps to a Phenomenology of Human-Computer Interaction'.

Svanæs, Dag. 2013. 'Interaction Design for and with *the Lived Body*: Some Implications of Merleau-Ponty's Phenomenology'. *ACM Transactions on Computer-Human Interaction* 20(1):1–30. doi: 10.1145/2442106.2442114.

de Swart, Peter, and Helen de Hoop. 2018. 'Shifting Animacy'. *Theoretical Linguistics* 44(1–2):1–23. doi: 10.1515/tl-2018-0001.

Thinès, Georges, Alan Costall, and George Butterworth, eds. 1991. *Michotte's Experimental Phenomenology of Perception*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Thrun, Sebastian. 2004. 'Toward a Framework for Human-Robot Interaction'. *Human-Computer Interaction* 19:9–24. doi: 10.1207/s15327051hci1901&2_2.

Turkle, Sherry, Will Taggart, Cory D. Kidd, and Olivia Dasté. 2006. 'Relational Artifacts with Children and Elders: The Complexities of Cybercompanionship'. *Connection Science* 18(4):347–61. doi: 10.1080/09540090600868912.

von Uexküll, Jakob. 1921. *Umwelt und Innenwelt der Tiere*. Berlin, Heidelberg: Springer.

von Uexküll, Jakob. 1992. 'A Stroll through the Worlds of Animals and Men: A Picture Book of Invisible Worlds'. *Semiotica* 89(4). doi: 10.1515/semi.1992.89.4.319.

von Uexküll, Jakob. 2001. 'An Introduction to Umwelt'. 2001(134):107–10. doi: 10.1515/semi.2001.017.

von Uexküll, Jakob. 2010. 'A Foray into the Worlds of Animals and Human'. Pp. 41–138 in *A foray into the worlds of animals and humans: with A theory of meaning*, *Posthumanities*. Minneapolis: University of Minnesota Press.

Uexküll, Thure von. 1982. 'Introduction: Meaning and Science in Jakob von Uexküll's Concept of Biology'. 42(1):1–24. doi: 10.1515/semi.1982.42.1.1.

Van Manen, Max. 2023. *Phenomenology of Practice: Meaning-Giving Methods in Phenomenological Research and Writing*. Second edition. New York London: Routledge.

Veling, Louise, and Conor McGinn. 2021. 'Qualitative Research in HRI: A Review and Taxonomy'. *International Journal of Social Robotics* 13(7):1689–1709. doi: 10.1007/s12369-020-00723-z.

Wagemans, Johan, Rob van Lier, and Brian J. Scholl. 2006. 'Introduction to Michotte's Heritage in Perception and Cognition Research'. *Acta Psychologica* 123(1):1–19. doi: 10.1016/j.actpsy.2006.06.003.

Walsham, Geoff. 2006. 'Doing Interpretive Research'. *European Journal of Information Systems* 15(3):320–30. doi: 10.1057/palgrave.ejis.3000589.

Watzlawick, Paul, Janet Beavin Bavelas, and Don D. Jackson. 1967. *Pragmatics of Human Communication: A Study of Interactional Patterns, Pathologies and Paradoxes*. W. W. Norton & Company.

Wheeler, Michael. 2014. 'Martin Heidegger'. in *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta.

Winograd, Terry, and Fernando Flores. 1986. *Understanding Computers and Cognition: A New Foundation for Design*. Intellect Books.

Winthrop-Young, Geoffrey. 2010. 'Afterword: Bubbles and Webs: A Backdoor Stroll Through the Readings of Uexküll'. Pp. 209–43 in *A foray into the worlds of animals and humans: with A theory of meaning*, *Posthumanities*. Minneapolis: University of Minnesota Press.

Young, James E., JaYoung Sung, Amy Voida, Ehud Sharlin, Takeo Igarashi, Henrik I. Christensen, and Rebecca E. Grinter. 2011. 'Evaluating Human-Robot Interaction'. *International Journal of Social Robotics* 3(1):53–67. doi: 10.1007/s12369-010-0081-8.

Ziemke, Tom, and Noel E. Sharkey. 2001. 'A Stroll through the Worlds of Robots and Animals: Applying Jakob von Uexkülls Theory of Meaning to Adaptive Robots and Artificial Life'. *Semiotica* 2001(134). doi: 10.1515/semi.2001.050.

Papers

I

Paper 1

Turning Away from an Anthropocentric View on Robotics

Soma, R., & Herstad, J. (2018). Turning Away from an Anthropocentric View on Robotics. Frontiers in Artificial Intelligence and Applications, 53–62.
https://doi.org/10.3233/978-1-61499-931-7-53

Paper 2

II

Engaging in Deep Wonder at the Experience of Encountering a Lawnmower Robot

Soma, R. (2020). Engaging in Deep Wonder at the Experience of Encountering a Lawnmower Robot. In M. Nørskov, J. Seibt, & O. S. Quick (Eds.), Frontiers in Artificial Intelligence and Applications. IOS Press.
https://doi.org/10.3233/FAIA200905

II

Paper 3

Movement Acts in Breakdown Situations: How a Robot's Recovery Procedure Affects Participants' Opinions

Schulz, T., Soma, R., & Holthaus, P. (2021). Movement Acts in Breakdown Situations: How a Robot's Recovery Procedure Affects Participants' Opinions. Paladyn, Journal of Behavioral Robotics, 12(1), 336–355. https://doi.org/10.1515/pjbr-2021-0027

III

III

Research Article **Open Access**

Trenton Schulz*, Rebekka Soma, and Patrick Holthaus

Movement Acts in Breakdown Situations

How a Robot's Recovery Procedure Affects Participants' Opinions

Abstract: Recovery procedures are targeted at correcting issues encountered by robots. What are people's opinions of a robot during these recovery procedures? During an experiment that examined how a mobile robot moved, the robot would unexpectedly pause or rotate itself to recover from a navigation problem. The serendipity of the recovery procedure and people's understanding of it became a case study to examine how future study designs could consider breakdowns better and look at suggestions for better robot behaviors in such situations. We present the original experiment with the recovery procedure. We then examine the qualitative responses from the participants in this experiment to see how they interpreted the breakdown situation when it occurred. Responses could be grouped into themes of sentience, competence, and the robot's forms. The themes indicate that the robot's movement communicated different information to different participants. This leads us to introduce the concept of movement acts to help examine the explicit and implicit parts of communication in movement. Given that we developed the concept looking at an unexpected breakdown, we suggest that researchers should plan for the possibility of breakdowns in experiments and examine and report people's experience around a robot breakdown to further explore unintended robot communication.

Keywords: Non-verbal cues and expressiveness, Movement, Recovery, Human-Robot Interaction, Study Design, Adoption of technology, Trust

***Corresponding Author: Trenton Schulz:** Norwegian Computing Center, Oslo, Norway, E-mail: trenton@nr.no
Rebekka Soma: University of Oslo, Oslo, Norway, E-mail: rebsaurus@uio.no
Patrick Holthaus: University of Hertfordshire, Hatfield, England, E-mail: p.holthaus@herts.ac.uk

1 Introduction

Robots are developed to do specific tasks, and people interacting with them expect them to perform these tasks correctly and efficiently. In a dynamic and unpredictable environment, however, robots are vulnerable to unforeseen issues. If, for instance, a robot suddenly becomes unaware of where it is, it will have to reorient itself. Even in controlled environments, robots can still function incorrectly, and people seeing the robot will inevitably interpret its malfunction.

In an earlier experiment we ran, participants collaborated with a mobile robot to tidy up in a home environment [1]. The goal of the experiment was to see if the way robot sped up and slowed down changed people's opinion about the robot. During the experiment, an *unplanned* event sometimes occurred where the robot would become "stuck" in the navigation stack. This made the robot pause or go into a recovery procedure to free itself. The experiment did not lead to an interesting quantitative result, but participants remarked about the recovery procedure when answering questions during the experiment. So, we used the serendipity of the situation to examine if statements from the participants could provide insights into future study design or help to develop new recovery procedures.

In this article, we present a case study to systematically evaluate unanticipated breakdown situations that occurred in the original experiment. We analyze the participants' qualitative responses on how well the robot handled the task. We identify three themes in the responses after the robot paused or ran its recovery procedure. The themes show that the robot's movement communicated different things to the participants. We introduce the concept of *movement acts* to examine different aspects of a movement's implicit and explicit communication to better communicate with human participants. The participants' statements show that examining unplanned breakdown situations can yield interesting data that might otherwise be ignored.

In particular, insights from our analysis help to better understand the nature of a robot's social signals and they are thus valuable for application in real-world scenarios. People need to trust robots to work with them or accept their services, and a mismatch between the expectation and reality can lead to a loss of trust [2]. Further, even single violations can lead to a significant reduction of trust in the technology [3]. It is therefore important to design robots to compensate for possible negative feelings or concerns. By examining people's opinions in an HRI scenario where the robot does not work as expected, we may get a better understanding of people's feelings towards robots in other breakdown situations as well. Thus, the study might help to identify factors that could affect trust in encounters where the robot faces an issue but also in those with a flawless robot performance. Finally, there is a benefit from examining breakdowns in an experiment. The examination may produce interesting quantitative results to inform future study design and supplement already suggested best practices [4]

We begin by presenting how a robot's movement can carry meaning explicitly and implicitly (Section 2). We then review other studies that have examined breakdown situations in HRI (Section 3). Then, the original experiment design is presented (Section 4), which is the setting for the case study. Next, the case study is presented with an elaboration on the unplanned recovery procedure, a description of our analytical procedure, and the presentation of results that include common themes we identify from participants' opinions (Section 5). We discuss the communicative nature of each themes, introducing and discussing the term *movement acts* (Section 6). We provide suggestions for incorporating unexpected movement acts in study designs and limitations of our analysis before concluding (Section 7).

2 Social signals and cues

A central challenge in social robotics is to understand how humans interpret the meaning of a robot's actions and behaviors [5]. Since information between humans and robots is typically exchanged through seeing, hearing, and touch [6], a person can interpret a robot's capabilities and intentions through explicit non-verbal communication such as gestures, facial expressions, or movement in space. These have been called *communication modalities* [7]. Each modality can be thought of as having an *explicit* and an *implicit* dimension [8].

Before presenting the case study, let us establish some background on how unintended movement can also communicate social cues. We start first with examining how meaning can be found in movement. Then, we will review how robot behavior (e.g., movement) can be interpreted socially. This will help to explain why it is interesting to consider a robot's movement in a breakdown situation.

2.1 Meaning in movement

The *speech act theory* tells us that humans are attuned to a speaker's intended meaning (i.e., the content of the words and sentences themselves) *and* to the speaker's utterances (i.e., the acts of speaking or not speaking). The utterance itself can contain "requests, warnings, invitations, promises, apologies, predictions, and the like" [9, p. 1]. The theory draws parallels to Watzlawick et al.'s *first axiom of communication* that states "… no matter how one might try, one cannot *not* communicate. Activity or inactivity, words or silence all have message value" [10, p. 30]. That is, it is impossible to not communicate and there is no such thing as a *non-behavior*. Expanding this to include movement, humans, as social beings, are sensitive to both the implicit and explicit dimensions of movement as well. They actively look for and interpret signals of social behavior.

While explicit communication signals convey information purposely from the sender to the receiver with a defined and intended meaning, implicit communication signals convey information that lacks this intention and purpose [11]. Instead, information is inadvertently conveyed that may or may not be incidental. This information could include the sender's emotional state, inner motivation, or intention behind an utterance or an action [7], and can be interpreted by the receiver consciously and unconsciously. That is, information can be sent and received without an intended message, and the unintended message can lead to misunderstandings. For example, some movements are intended to explicitly signal a message (e.g., waving to a friend). Upon receiving such a signal, the receiver might interpret the intended message, but at the same time be sensitive to what is implicit in the act of waving [12]. Often, however, movements and behaviors are just incidental. For example, a friend's apparent wave turns out to be only stretching with no intention to signal "hello".

The transmission of information is not limited to human movement. It is now generally recognized that most people will assume intentions of objects and figures that move in a certain way, even though they are aware that the objects and figures are not actually alive. The phenomenon, usually referred to as *anthropomorphizing*, was demonstrated in a study where humans observed the movements of geometrical shapes and the observers assigned the shapes agency, motive, and personality [13]. Recently the phenomenon was categorized as a *type of experienced sociality* [14]. A related, but slightly different kind of experienced sociality is *sociomorphing*. It occurs when a person interacts with a non-human agent and attributes it social capabilities although it might not necessarily have human-like properties [14].

How do we examine this phenomenon? One solution is to use *semiotics*, the study of signs and their usage. The most common understanding of *signs* is a dyadic relationship between the signifier and the signified: A sign represents its object in some respect. Semiotics is often associated with text and media analysis, but signs do not necessarily need to be linguistic symbols. Further, the study of signs is not exclusively looking for symbolism and hidden meaning in the different forms of storytelling in text and media. In Pierce's pragmatic tradition of semiotics [15], a sign is *not* dyadic relationship; instead a sign is a *triadic* relationship between the signifier, the object signified, and an interpreter (or "translator") of what is represented. The study of signs in the pragmatic tradition of semiotics is thus concerned with the study of how meaning is *generated* in this triadic relation.

So, by using the pragmatic tradition of semiotics, communication can be unconscious and pre-reflexive, and forms of unconscious communication and sign processing exist beyond human language [16]. Thus, anthropomorphizing and sociomorphing are two examples of pre-linguistic meaning-making phenomenon occurring in everyday experiences.

Because of the asymmetrical social capabilities of humans and robots [17], the first axiom of communication might not translate perfectly to HRI; robots move and behave in human social spaces, but cannot truly be considered to have feelings, moods, purpose, etc. The semiotic perspective, however, is sensitive to any layer of meaning implicit within any and every movement, and enable us to analyze robot movements and non-movements, for example in a breakdown situation, as meaningful, even if no message was intended to be communicated to a user.

2.2 Interpreting robot behaviors socially

A robot's core functionality is often enhanced using social features to make the interaction more robust [18]. That is, by using shapes that can be socially interpreted or by actively communicating the robot's current state, it is easier for people to interpret the robot's function and behavior [19]. Knepper et al. [20] argued that actions performed in collaboration between humans and robots will be interpreted as functional *and* communicative. Humans interpret a robot's signals and cues even when these signals and cues might not have an explicitly defined or well-designed social meaning. That is, the robot's blinking lights, noises from motors, or body movements sometimes have an unintended effect on a robot's social perception [8]. For example, even though robots may deliberately make sounds intended to communicate with people (*intentional sounds*), the noise produced by actuation servos for robot functionality (*consequential sounds*) also shaped people's interaction with the robot [21]. Because consequential movements and noises are inevitable to get the robot to move, designers and developers were encouraged to consider what might be implicitly communicated to the user through these modalities, especially considering that robots do not need to be anthropomorphic to be sociable [22]. For example, the Fetch robot (Fig. 1) uses its pan-tilt camera in its head to support its navigation algorithm. In our experiment (Section 4), the robot's movements of this part could be misinterpreted as head movements bearing social gaze.

Modeling and exhibiting social signals appropriately can aid the robot in communicating its current state [19] and guide users through an interaction situation [23]. Several studies have found that using embodied cues, such as verbal, vocal, gaze, gestures, and proximity, can influence people's opinion of the robot [24]. Some examples of embodied cues influencing people's opinion include using motion that communicate the robot's collision avoidance strategy instead of its destination [25] or expressing the internal state of the robot by timing the robot's movements [26]. Techniques from animation, such as the twelve animation principles [27], have also been used to communicate a robot's intent to people watching or working with a robot [28].

Purely movement-based interactions can also be successfully implemented. The creators of a mechanical ottoman made it move in such a way as to ask if the person in the room was willing or available to interact with it [22]. The study illustrated that the designers were aware that there is an explicit and an implicit dimension to the ottoman's movement. Another study had a robot move its arm using what the researchers characterized as legible motion. The human collaborator could better infer the robot's goal and resulted in better collaboration on a shared task [29]. Cooperation between humans and robots also improves when developers carefully consider how to use a robot's movement for explicitly expressing its purpose, intent, state, mood, personality, attention, etc. [30].

There are also examples of what can happen when robot motion does not take into account how a robot may appear socially, even when it is not regarded as a social robot. A mismatch between the expectation and reality may, for example, lead to a loss of trust [2]. In one instance, a military robot was deactivated after it made unanticipated movements, and people distrusted it [31]. Another example is in a study where people showed tendencies towards anxiety and discomfort when they were uncertain how a robot arm would move as they worked together in proximity on a task [32]. In one study, people viewed a robot in virtual reality and on video sorting balls according to color. Participants watching the video trusted the robot when it moved fluidly, but less when it trembled doing its task. This finding was not confirmed when the robot and person cooperated on the same task [33]. This suggest a robot's motion may be more noticeable when the person is only watching the robot instead of working directly with the robot.

In summary, robots' actions communicate information even when their actions are not intended to communicate anything. Thus, it might be interesting to examine an unexpected breakdown situation in an experiment and investigate how people interpret and understand the situation and the robot's explicit and implicit communication. In our study, we apply a semiotic perspective to the perception and interpretation of robot motion as we have a special interest in the sociability of robots. We are interested in examining how a person might see the movement of a robot as a sign of "something".

3 Studies examining breakdown situations with robots

There are several related studies that address breakdown situations in HRI. In contrast to this article, all these studies examined breakdowns that happened as a part of the study design. Accordingly, their participants may not have known about the breakdown beforehand, but the people running the study did.

In this section, we first detail how these studies identify negative effects of breakdown situations on people's opinion of a robot, such as a loss of trust, in a controlled way. This provides a starting point to analyze the observations from our original experiment and see if it can confirm previous studies' findings or introduce new lines of thought. This section also presents studies that develop mitigation strategies to repair negative effects of breakdowns and salvage the interaction as a basis for our later discussion and identification of themes. Finally, we look at a study where the systematical documentation of accidental breakdowns in pre-tests can help to prevent them later to get an inspiration how other researchers have learned from them.

3.1 Effects on user perception

A number of studies investigate how a planned breakdown alters people's perception of a robot and hence how such situations influence the robot's acceptability and usefulness. In general, it appears that different contexts lead to different implications for a robot's breakdown or errant behavior. For example, in one study where children were to engage with a robot, a robot that displayed unexpected behavior elicited more engagement from the children than one that behaved as expected [34]. In a different study, a human and robot worked together on memory and sequence completion tasks. When the robot made mistakes, it triggered a positive attitude for the human, but lowered human performance [35]. Yet another study found that participants preferred a robot that made mistakes in social norms and made small technical errors in an interview and instruction-giving process than one that performed flawlessly, but the study found no differences in the robot's perceived intelligence or anthropomorphism [36]. In contrast, we suspect that participants in our original experiment might have been frustrated or irritated by the breakdown instead.

Some studies have examined specifically how breakdowns affect human trust in robots. A meta-analysis of factors influencing trust in HRI found that the robot's task performance had a large impact on people's trust [37]. In another study, researchers looked at how willing people were to follow odd commands, such as watering a plant with orange juice, from a robot that was acting faulty [38]. Although the robot's behavior affected participants' opinion of the robot's trustworthiness and the participants had different opinions about the odd requests, many of the participants honored the requests. The researchers speculated that this could be due to some participants feeling they were in an experiment and actions therefore had low stakes. Similarly, we are interested in examining how erroneous behavior that cannot easily be interpreted might have affected the users and their perception of the robot.

Another study provided different ways that a robot could handle a breakdown while playing a cooperative game with someone and looked at people's trust in the robot afterwards [39]. For some participants, the robot would freeze while speaking in mid-sentence during the game. It would then either start from the beginning or pick up from where it left off. It could then provide a justification for why it froze or offer no explanation. The robot's freeze had a negative effect on the participants' perceived trust of the robot, but restarting the interaction had a more negative impact on the perceived trust than if the robot continued. Robots that continued and provided a justification for freezing further reduced the negative perceived impact of trust. Similarly to our study, we are interested in examining if a robot's freezing and recovery behaviors might have caused negative effects on the users' perception.

3.2 Repair and mitigation strategies

Some studies have evaluated different mitigation techniques to salvage an interaction despite the occurrence of breakdowns. One experiment investigated whether some robot action can repair the situation after a planned breakdown [40]. The experiment let people observe a scenario with a robot (with either a humanlike or non-humanlike form) and person. The observers then rated the robot and the service it provided. The robot's breakdown had a negative influence on how observers rated their satisfaction with a robot and the service, but different mitigation techniques could change the observer's opinion. If the robot apologized for the breakdown, the observer tended to rate the service provided by the robot higher. Similarly, if the robot offered compensation for the error, the observer felt the current interaction went well. There was also a correlation between an observer's orientation to service (more relational versus more utilitarian) and how well the mitigation performed. Finally, if some mitigation was provided, observers rated the robots as more human-like regardless of the robot's form.

A different approach is the strategy of *calibrated trust* where the person's expectations are tuned to the robot's shortcomings or potential malfunctions [2]. For example, participants in the study above by Lee et al. rated the task more difficult for the robot if the robot warned early that it might not complete the task correctly [40]. This article examines both aspects, i.e., it provides help with planning and adjusting the robot's behaviors to the participants' expectations, and it provides tools to design fallback strategies for unexpected cases.

3.3 Unexpected breakdowns

Most breakdowns do not happen according to plan and study data from these breakdowns is often discarded to make data analysis easier. Few have argued that there is value hidden in data discarded due to robot breakdowns and other error situations [41]. For example, Barakova et al. have documented a robot's unexpected behavior in pilot studies with children with autism and how the unexpected behavior affected the children [42]. The unexpected errors were the result of a mistake by the leader of the session, software problems, or issues with the robot. The experiences lead the researchers to document their redesign of the study and changes to the robot's software to eliminate the issues for the final study [42]. The documented changes are useful for other researchers designing similar studies.

Our purpose here was to look at people's opinion of the robot's breakdown situation in an experiment that was *not* designed for a breakdown and was *not* present in the pilot study. We wanted to examine the participants comments and see what lessons we could learn for future experiments and robot design. We performed this examination through the lens of explicit and implicit communication.

4 Case setting: earlier experiment

The case study focuses on people's opinion of a robot during a temporary breakdown and self-recovery of its navigation system. The data is collected from an earlier experiment that examined how people reacted to a robot that moved using two different velocity curves [1]. To introduce the case, we document the robot and setting that was used in the earlier experiment, its procedure, and the data that was collected. Although the experiment's method was documented previously [1], we present an expanded description of the experiment here to highlight some constraints and challenges in the design.

The original intention of the experiment was to look at one animation principle, *slow in and slow out*, and see how it affected people's perception of the robot. The slow in and slow out animation principle states that the speed an object moves at changes through its journey: motion is slower at the beginning and at the end [27]. Using the slow in and slow out principle should lead to a motion that appears more "natural" and less "robot-like".

Given the constraints of designing and running the experiment, we went for a within-subjects design for the experiment. This decision likely had an effect on the quantitative results (for example there could be a learning effect between studies [4]), it is less important for the purposes of a case study, especially given that the breakdown was unplanned and occurred throughout the whole experiment.

As we designed the experiment, we were concerned that if we simply presented the robot moving using a velocity profile using the slow in and slow out animation principle or the standard linear velocity profile and asked people their opinion, they would manufacture a response to satisfy our question, and we would not get their actual perception. We decided that participants would take part in a task that was dependent on them watching the robot's movement and seeing the movement from different angles, but participants were not explicitly asked about the robot's movement. The participants' answers would focus on the way the robot performed the task and not on how the robot moved. The experiment would see if the way that the robot moved affected the participants' opinion of the robot.

4.1 Experiment setting, questionnaire, robot, and navigation system

The procedure was approved by the University of Hertfordshire Health, Science, Engineering and Technology Ethics Committee (Protocol Number COM/SF/UH/03491) and took place at the University of Hertfordshire's Robot House. Robot House is a place that people can visit and experience robots and sensors in a home environment instead of a typical lab environment. Since the overarching goal of the research is to have a robot be a part of a home and that the robot's movement should appear more friendly and ultimately lead to better trust in the robot, it seemed appropriate to run the experiment in a physical area that resembled a home environment rather than a lab.

The questionnaire for the original study included the Godspeed series [43]. We also included an additional Likert item about how well the person could predict where the robot would go, and an open question: "What do you think about how the robot handled this task?" We included the prediction item as we wondered if the different velocity profiles would affect how easy the person could predict the robot's movement. The results from the Godspeed series were reported previously [1]. The open question gathered qualitative information and is the basis of our analysis below (Section 5).

Fig. 1. The Fetch robot at Robot House, its arm configuration, and the basket used for the experiment.

The robot we used was a Fetch Mobile Manipulator from Fetch Robotics [44] hereafter referred to as Fetch (Fig. 1). We selected Fetch since it can move at a rate of 1 meter per second (m/s). This speed is slower than an average person's walking speed of 1.4 m/s [45], but accelerating up to this speed takes enough time that it is possible to create different velocity profiles.

The linear and slow in and slow out velocity profiles were based on the algorithm described in Schulz et al. [46] and adapted to a plugin for the local navigation planner in Fetch's navigation stack, which is the navigation stack from the Robot Operating System (ROS) [47]. Fetch's local planner uses the trajectory roll out scheme [48]. This method of integration is similar to a set up suggested by Gielniak et al. [49] for integrating stylized motion into a velocity profile for a task. The plugin included dynamic parameters for setting the velocity profile (linear or slow in and slow out). This allowed us to change the velocity profile without restarting the robot's navigation system. The changes only affected Fetch's linear velocity (i.e., moving forward); the angular velocity (i.e., turning in place) was unchanged from the original plugin and thus always used a linear velocity profile.

We considered ignoring the environment and simply issuing pre-recorded velocity commands to Fetch. This technique would have resulted in smoother velocity curves, but we were concerned that small inaccuracies would occur while turning, starting, and stopping would lead to large inaccuracies as Fetch moved through the house. Fetch's navigation stack had already been extensively tested for moving the robot around and avoiding obstacles. After investigating both, we found that Fetch's navigation stack with our developed plugin worked better than any solution we could develop from scratch in the time given for the experiment.

Fetch moved between several preassigned destinations in the house: (Positions 1, 2, and 3 respectively on Fig. 2). Each spot had two poses, one for facing the person and one for facing away from the person towards the next location. The two poses per location was done to keep the performance of Fetch's navigation similar across conditions. Position 1 had a slightly different locations for its poses to make it easier to remove and add items to the basket without the participant noticing.

4.2 Experimental Procedure and Data Collection

Participants that had consented to being part of the experiment entered Robot House and filled out demographic information of age, gender, and if the participant had any experience with robots.

After the participant filled out this questionnaire, we went through safety information with the participant for interacting with the robot. We explained that they would be interacting with a Fetch robot during the experiment and that two of us would be constantly monitoring the robot. Fetch was brought over and controlled with the remote control during this explanation. We told the participants that we did not expect any safety issues, but advised them not to approach the robot while it was moving and that if the emergency stop was engaged, that the robot would keep its momentum and move unexpectedly. Participants were told they could end the experiment at any time if they felt unsafe (none of the participants ended their participation). As the safety information was being explained, one of the facilitators remotely controlled the robot and moved it towards the participant so the participant could see how the robot moved and see its size. Participants could ask additional questions regarding safety at this time.

Then, the scenario was explained. The participant was visiting a friend's house to help in cleaning up the home (The facilitator that had controlled the robot was introduced as the friend). Cups had been placed on the dining table and the coffee table near some couches. These cups needed to be returned to the kitchen. The robot would aid in the cleanup by collecting cups from the participant and taking them to the kitchen. Since we did not want to draw attention to the robot's motion, we explained we were interested in how the robot handles the hand over of objects from the participant. We instructed the participant where to stand, what to do, and what the robot would be doing (Fig. 2).

The facilitators and the participant would then take their positions. Video recording of the procedure from one camera was started for participants that consented. One facilitator

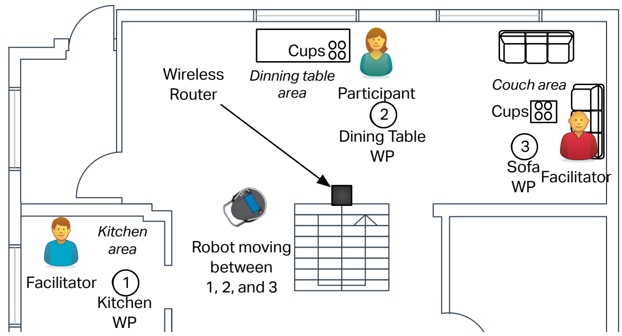

Fig. 2. Floor plan and position of people for the experiment. The robot would move between the numbered positions, starting at Position 1, using either a linear or slow in and slow out velocity profile.

would stand in the kitchen (near Position 1 in Fig. 2); the participant and the facilitator helping in cleaning the house would stand near the dining table and sit on the couch respectively (near Positions 2 and 3 respectively in Fig. 2). Fetch would be sent to Position 1 in Fig. 2. The remote control was placed on the table near the sofa to indicate the robot was not being teleoperated, but the remote control was in easy reach of the facilitator if something were to go wrong.

Starting at Position 1, the procedure was the following: (*1*) The robot moved from Position 1 to Position 2. (*2*) The participant took one of the cups from the dining table and put it in the robot's basket (Fig. 1). (*3*) The robot moved from Position 2 to Position 3. (*4*) The facilitator on the couch took a cup from the coffee table and put it in the robot's basket. (*5*) The robot moved to Position 1. (*6*) The facilitator in the kitchen removed the cups and put a copy of the questionnaire in the basket. (*7*) The robot moved to Position 2. (*8*) the participant took the questionnaire from the robot and filled it out. (*9*) Once the questionnaire was complete, the participant put the questionnaire back in the robot's basket. (*10*) The robot moved to Position 1. (*11*) Finally, the facilitator in the kitchen removed the questionnaire and prepared the robot for the next iteration.

This procedure was performed for four iterations: two times the movement was with a linear velocity profile, and two times the movement was with a slow in and slow out velocity profile. The profiles were counterbalanced to avoid ordering effects. The counterbalancing was achieved by taking the six possible combinations of two linear and two slow in and slow out velocity profiles, and randomly selecting an ordering for each participant.

After the final iteration, any video recording was stopped and participants went through an ending procedure. They filled out a questionnaire with open-ended questions concerning the overall interactions: "Do you have any general impressions about the robot during your interaction with it?" and "Do you have any questions you would like to ask us?"

We also informed participants that we were actually interested in the robot's movement and not the handover. We used this opportunity to answer their questions and go into more technical details about how the robot sensed its environment and moved around the house. Participants were encouraged to ask any additional questions about the set up, the robot, and the experiment. Finally, we thanked participants for their time and, as noted in the informed consent form, gave them a £10 gift card for Amazon as compensation for time and traveling to Robot House.

Since we were concerned how the robot's movement affected people's perception of the robot, we removed some confounding factors to improve the internal validity of the experiment. For example, we chose to use a basket for collecting cups and the questionnaire since Fetch's arm movement is not deterministic and would confuse participants. In addition, Fetch only moved and did not use speech recognition or sound. A pilot study revealed that it was confusing for the person to know when it was OK to put a cup in the basket. To signal to the person that Fetch was ready to receive a cup or take and return a questionnaire, it would raise its torso 10 cm. when it had arrived at the pose facing that person.

Participants were asked to stand if able while giving the cup to the robot and receiving the questionnaire (all participants were able to stand). They could sit while filling out the questionnaire. The primary reason was to allow a better view of Fetch and keep the base for participants' perceptions similar since a standing participant is taller than the robot, which might not be true with a sitting participant. A lesser, secondary reason was to make participants feel safer as the robot approached since we reasoned that participants may feel easier to move away from a robot when they are already standing versus having to get up from a chair. For an additional level of safety, having a person in the kitchen and the couch also allowed two people to watch the robot and activate an emergency stop if Fetch was going to run into something.

Fetch was *partially* controlled via Wizard of Oz. In line with recommendations from Riek [50], we include the additional information about our use of Wizard of Oz. The wizard, the facilitator in the kitchen and this paper's first author, acted as the robot's eyes and as a conductor for the robot. The wizard was in charge of noticing when the participant or the facilitator had put the cup into the basket. Then, the wizard would signal for the robot to go to the next pose. The robot would then navigate to the next position using its navigation stack. We chose to use a Wizard of Oz component to reduce variability of time for the experiment with the robot detecting the cup or questionnaire was added to the basket. Given that Fetch traveled a fixed route and the participant's role was rigidly defined, the Wizard of Oz component could have been eliminated given enough time. The wizard also noted down observations for each iteration.

For additional data, we collected the robot's odometry information and the time from when a request to move was made to move to the next location until the time that the robot arrived at the location and raised its torso.

5 The case study

In this section we document the unexpected robot's recovery procedure, and describe how we analyzed the data collected from the previous experiment for the case study in this paper.

5.1 Robot's recovery procedure

From the Godspeed questionnaire, the participants' responses were not different enough between the linear or slow in and slow out velocity curve [1]. When we were examining the qualitative, free-text comments, we noticed how some comments expressed feelings of discomfort, curiosity, or confusion from interacting with Fetch, especially during its recovery procedure. The unplanned phenomenon became part of the interaction in a substantial amount of the trials, happening 68 times in total or around 9% of the time when the robot moved. 31 of the 38 participants experienced at least one of these issues.

The recovery procedure occurred when Fetch encountered problems in calculating its path for moving across the room. The recovery procedure caused the robot to move differently than its intended behavior and was not planned for by the facilitators. As part of the navigation stack, the recovery procedure was likely implemented to get Fetch to move again without considering what an observer would see. During the pilot and testing, the breakdown situation did *not* occur. Once discovered, the functionality could have been disabled, but it would have increased the chance the robot did nothing, which would have caused an even larger interruption during the experiment.

Our interest is in the case of the unexpected breakdown. Further, the irregularity of its occurrence makes it unfit for quantitative analysis. Instead, we analyze the comments qualitatively and cross-reference them with our notations of when Fetch had issues, and how that issue manifested in the robot's movement and behavior.

Normally, the robot navigated competently between the positions in Fig. 2. When Fetch received an instruction to proceed to the next navigation point, its head would look up and down as it calculated the path and speed to travel. The head movement would pan-tilt the depth camera inside Fetch's head

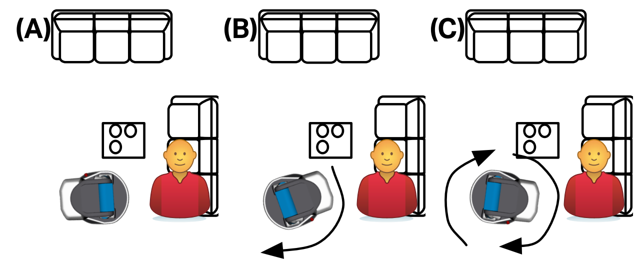

Fig. 3. Example of the recovery procedure when Fetch traveled from Position 3 to Position 1. In A, Fetch is at Position 3 and has received the cup. Then, it rotates to go to Position 1 (B). Once the rotation finishes, it tries to compute the route to Position 1. If this takes longer than 25 seconds, Fetch is considered "stuck" and rotates 360 degrees to get unstuck (C) before continuing on its path.

and support its navigation algorithm. When a path had been calculated, it would straighten its head and proceed on the path.

This process would normally be completed within a second or two, and Fetch would begin to move. Sometimes, however, it encountered problems in calculating its navigation path. In these situations, it would continue trying to calculate a path and the head would continue to move up and down until one of the following things happened: (*1*) it succeeded in calculating the path, and started on the path after the delay or (*2*) if the navigation software had not calculated the path after 25 seconds, it decided that that Fetch was "stuck".

If Fetch was stuck, the navigation software would rotate the robot 360 degrees to make the robot "unstuck" (Fig. 3). After completing the rotation, Fetch would quickly find its path and proceed. The turn took around 8 seconds to complete. Adding the 25 seconds from attempting to calculate the path and around 4 seconds for Fetch to turn after receiving the cup results in the participant experiencing an approximate 40 seconds wait during Fetch's recovery procedure.

5.2 Analytical procedure

To begin our analysis, we looked at the logs of the robot's performance and noted when it paused or when it became stuck and initiated its recovery procedure. We then arranged the participants' responses to the open questions on each iteration and their overall opinions into tables, arranged first by trial and later by participant. This resulted in tables that charted the journey of the robot, and we could easily follow the robot, the problems it had, and the responses from the participant. The answers from all the participants were manually annotated through emergent coding, first deductively and then inductively [51]. The themes presented in the results emerged from the coding during the inductive approach, which was performed by two researchers who then met to harmonize on the themes. Exactly how many iterations and re-reading of the comments were not counted: Each researcher read through as many times as they needed to make sense of the data.

Table 1. Count of occurrences the robot had problems moving from position to position split by velocity curve for linear and slow in and slow out respectively ($n = 760$); positions are as documented in Fig. 2 and the columns follow the procedure documented in Section 4.2.

	1→2	2→3	3→1	1→2	2→1	Total
Linear						
Stuck	1	0	5	0	2	8
Delay	2	1	5	0	1	9
Other	0	1	0	0	1	2
No problem	73	74	66	76	72	361
Slow in & slow out						
Stuck	0	3	17	0	5	25
Delay	1	3	10	1	2	17
Other	0	0	0	0	1	1
No problem	75	70	49	75	68	337

5.2.1 Fetch's performance

The study had 38 participants; 19 identified as female and 19 identified as male. The participants' ages were from 18 to 80 years old (Average age: 37.39 years, median age: 34.5 years, SD: 15.74 years). 22 of the 38 (around 58%) participants had previous experience with robots. Each participant had four iterations of the cup cleaning task (two times with slow in and slow out and two times with linear) for a total of 152 encounters (76 for slow in and slow out and 76 for linear). Fetch's journey for each iteration can be divided into separate stages or legs (e.g., in one leg, Fetch traveled from Position 3 to Position 1). Each iteration had five legs. The total number of legs over all iterations is 760.

With our focus on the breakdown, we examined the videos of participants and noted Fetch's behavior. Fetch's behavior was divided into four classifications: (*1*) *no problem*: the robot worked as intended, (*2*) *delay*: Fetch made a longer calculation than normal, (*3*) *stuck*: Fetch was stuck and went into recovery mode, and (*4*) *other*: an event that could not be placed in the other behaviors (those three events are described below). These classifications were checked against the observation notes from the facilitator and could also be confirmed using Fetch's odometry logs. This also enabled us to classify for participants that did not wish to be recorded on camera (one participant chose not to be recorded). The coding for the videos was obvious and the classification was in agreement.

Table 1 shows counts for events as the robot moved from position to position split by velocity curve. The three events marked as *other* were: (*1*) the software crashed after the final questionnaire was filled out, (*2*) a near collision with the table at the sofa when the robot traveled from Position 2 to Position 3, and (*3*) the robot shook as it returned from Position 2 to Position 1.

The area that had the most problems was when the robot moved from the sofa area (Position 3) back to the kitchen (Position 1). The robot was stuck 17 times during the slow in and slow out curve and 5 times for the linear velocity curve. Overall, Fetch was stuck 25 times when it used the slow in and slow out curve versus the 8 times when it used the linear curve. If we look at these numbers in terms of percentages, approximately 92% of the legs had no problems. Splitting it by the linear and slow in and slow out the percentages of legs with no problems were 95% and approximately 89% respectively.

We were unsure about why there is a difference between the linear and slow in and slow out curves. Due to implementation reasons, the navigation stack ran on a separate computer and not directly on Fetch. Both curves, however, use the same code path and the only difference was the maximum speed the plugin allowed at the start and stop (in general, the slow in and slow out curve has slower maximum speeds at the start and stop). This should not have caused a problem in picking reasonable trajectories.

Table 2. Themes and underlying codes and number of comments regarding Fetch's movement; **Bold** indicates the theme and the total number of all its codes.

Theme	Number of comments
Sentience	**51**
Mood/emotion	6
Hesitate/wait	13
Checking/recognizing	15
Deciding/confused	9
Helping out	2
Intelligent/smart	5
Asking for cup	1
Form	**23**
Eyes/Look	9
Head/Nod	11
Arm/Extend	3
Competence	**53**
Handle/Do well	32
Slow	21

5.2.2 Analysis of comments

Using the tables that charted the journey of the robot with the responses belonging to each iteration in each trial, the we qualitatively analyzed each response. During the deductive stage, the focus was on whether the responses descriptively commented on what happened, or if metaphors were used to rationalize Fetch's behavior.

The inductive stage focused on how Fetch's sociability presented itself to the participant. Through the process of manually coding in iterations, the themes emerged. Coding the data using a semiotic lens on meaning-making made it easier to stay focused on what a participant's response could tell us about how they made sense of the movements of the robot (both during breakdown, but also when it worked according to plan).

Based on this work, we identified the categories forming the themes of our annotation scheme. The themes were: (*a*) *sentience*, participants associating abilities to Fetch and guessing its intention (51 comments); (*b*) *form*, participants commenting about Fetch's form or body parts (23 comments); and (*c*) *competence*, how Fetch performed its tasks and participants' confusion and uncertainty with the recovery procedure (53 comments), while a few (four participants) only reported on the movement with no underlying associations that we could identify. The themes are not mutually exclusive and some comments were coded into multiple themes. Table 2 breaks down the semiotic themes and their corresponding codes. Overall, the comments and themes showed up evenly distributed among all the iterations (Table 3).

Table 3. Breakdown of comment theme versus which iteration the comment was written, or if it was written at the end of all iterations; some comments are counted in multiple themes

Theme	It. 1	It. 2	It. 3	It. 4	End	Total
Sentience	11	9	12	8	11	51
Form	3	6	2	5	7	23
Competence	12	11	11	13	6	53

5.3 Results: Themes based on comments

The comments from participants when Fetch performed correctly were generally positive about how Fetch handled the task. While the responses made after an iteration where Fetch did not have any navigational issues are not excluded from our analysis. Going through the tables, however, it was clear that the more interesting comments were made when Fetch did not perform as expected.

We report participants' comments from the themes of form, competence, and sentience around the delays and recovery procedures, and another unrelated event that came up often enough that we include it as well. Since there were more problems with the slow in and slow out profile iterations, there are more comments from those iterations. We did *not* find a difference in the nature of the comments and therefore do not differentiate between the velocity profiles below. To avoid repetition, we do not report on *all* comments from the legs that had problems, but instead report a representative amount.

5.3.1 Sentience

Comments categorized into this theme included words commonly used to describe actions of living, sentient beings. They are examples of participants anthropomorphizing or sociomorphing Fetch, trying to explain or rationalize what they expected the robot would do or what they thought the robot intended to do. A few participants speculated about its *mood*. For example, one participant felt that "it looks oddly happy doing what its doing." (*Participant 35*). Another participant commented that "he [*sic*] looked sad on the last go" (*Participant 8*). We have not, however, attempt to speculate on whether the participants were purposefully attempting to anthropomorphize or sociomorph Fetch using these words.

When Fetch performed its task without delay or getting stuck, the participants tended to comment that the robot "handled the task well." But participants came to different conclusions about what was happening when the robot would pause. Some participants thought that the robot paused because it "... checked surroundings very well before moving back" (*Participant 2*), or that it "... felt like it was taking a bit more time to make [a] decision" (*Participant 10*). A third participant (*Participant 27*) felt that Fetch was quick and safe, and he could predict Fetch's movements after the second iteration. When it paused on the third iteration, however, he commented that Fetch "seemed to scan its surroundings more before moving." This was less predictable, but he felt that Fetch "was taking more precaution so completing [the] task safer."

Other participants interpreted Fetch's delay as confusion, "[The delay] evoked an impression of slight confusion" (*Participant 33*). When one participant witnessed the recovery procedure, he noted that the robot had "more confusion than last time" (*Participant 17*). When there were no issues on the next iteration, he declared that the robot had become "more confident." One participant that experienced Fetch's delay going from Position 2 to Position 3 and rotating before going to the kitchen felt that something may have been wrong with its sensors, "[Fetch was] more unpredictable, as if it couldn't sense as well the environment," (*Participant 21*). Another participant liked the recovery procedure. She "liked when it did a little twirl, but I thought that made it seemed confused..." (*Participant 28*). Finally, one participant had concern for Fetch when it executed the recovery procedure, "[the behavior] made me want to come over and check on him [*sic*]" (*Participant 36*).

Though not related to the recovery procedure, there was some confusion on when participants should hand over the cup in the first iteration. Several participants needed a hint on the first iteration, "... an indicator it would stop then raise up would have been helpful," (*Participant 35*). One participant (*Participant 19*) comment that she "wasn't sure" when to hand over the cup when Fetch first stopped. In the fourth iteration, there was a delay in Fetch raising its body, and she "wasn't sure when exactly he [*sic*] would be finished and I could hand over the cup." In other iterations, she felt that Fetch performed the task well and "wanted to say, 'Thank You.'" Fetch would reach a position and then rotate to face the person. Sometimes it would overshoot its stop position and need to rotate back. This also lead to some confusion: "It wasn't really clear when I was supposed to give the cup. Then it's moving around gave me the impression it was waiting" (*Participant 6*).

Fetch raised itself 10 cm for each participant. There would sometimes be a delay between arriving at a position and raising. During one delay, one participant (*Participant 18*) thought Fetch's delay in raising its body was due to calculating the person's hand height, "Maybe it took a little time for it to adjust to where my hands were."

Participants also had ideas about Fetch *should* move in some situations. One participant (*Participant 1*) noted, "The movement could be slower near obstacles. The trajectory would be more reassuring of a minimal accident possibility." Another participant was curious about " ... how it would react to a change in conditions (fallen cup, user movement, etc.)" (*Participant 27*). Another participant shared this curiosity, and she noted that Fetch "Moves quite smoothly. Avoids obstacles (perhaps there are insufficiently many obstacles to show this)" (*Participant 18*). A different participant (*Participant 22*) felt that Fetch could have moved faster and gotten closer, "[Fetch was] too slow for me; could have come a bit nearer to me to collect the cup and the questionnaire." Finally, one participant (*Participant 32*) commented that Fetch's approach could have been better since "... sometimes the movement adds a fear to the user (whether it will stop or not)."

5.3.2 Form

Comments categorized into this theme specifically described characteristics associated with distinct "body parts" and actions supported by them.

Fetch's head and arm gave participants certain expectations about how they should interact with it. Several participants expected Fetch would use its arm when getting the cup or at least extending the bag (*Participant 6*, *Participant 7*, and *Participant 35*). As mentioned in Section 4.2, the arm was disabled for safety and consistency.

Fetch's head and its rising and lowering left some participants thinking that Fetch was doing more. "[Fetch] keeps lowering its eyes towards my groin. Is this normal?" asked one participant (*Participant 1*). A different participant commented that Fetch didn't make eye contact (*Participant 29*). Another participant found it strange, "It felt odd that the laser scanner (or whatever it is) never tilted upwards to 'look at' me" (*Participant 33*). Yet another participant (*Participant 22*) commented, "I am not sure if the upping and downing of the head piece was assessing me or even waiting for me to react." She also complained, "The sound when Fetch was going up and down was a bit annoying." On the other hand, one participant "liked the way the robot bobs its head; it is quite humanlike," (*Participant 28*).

One participant gave Fetch more abilities than it had. Upon first interacting with Fetch, one participant commented that Fetch was "smart to sense objects around it." (*Participant 24*). In the second iteration, Fetch ran the recovery procedure twice. The participant maintained that Fetch "... handled [the] task, but [was] slow in process, although it's smart." Later, she commented that Fetch could "sense the obstacles in between or around it and make its way back" and felt that Fetch was "quick to respond" in the final encounter. Generally, she felt that Fetch was "friendly and smart".

Finally, aside from moving after receiving the cup, Fetch did not react to participants' actions. This made some participants (*Participant 7* and *Participant 33*) question whether the robot actually looked at them, even though they used both "eyes" and "look" to describe their thoughts.

5.3.3 Competence

Comments categorized into this theme encompassed a level of confusion or uncertainty with the participants around the recovery procedure. Although some participants observed the delays and recovery procedures and questioned the robot's intentions, other participants did not like when this happened.

One participant (*Participant 3*) experienced several emotions of the course of his iterations with the robot. After the first iteration, he commented that "It was a bit unpredictable at first, but I got used to its actions." On the second iterations, after Fetch miscalculated the path from Position 2 to Position 3 and almost hit the table by the couch, and executed the recovery procedure, he expressed confusion: "I struggled to understand what the robot was doing." The third iteration went better, but he still expressed worry: "I could predict what the robot was doing, but it felt like it was going too fast. It felt rushed when putting my cup in the basket." The fourth encounter had a recovery procedure from Position 3 to Position 1 that did not make him feel comfortable: "I felt relaxed until I saw the robot pause after picking up the second cup. I wasn't sure why it kept looking up and down, this made me a bit uncomfortable". He finally summed up all iterations optimistically, but not fully convinced: "I felt overall comfortable with how the robot was helping me. The robot pausing for a long time made me feel uneasy at times."

These concerns are also echoed by another participant's experiences (*Participant 15*). By the second iteration, she commented that the "robot handled [the] task well; [I] felt more comfortable with the robot so made the experience more comfortable." This continued with the third iteration: "Robot handled task smoother I feel than previous two tasks. Robot speed also feels like it increase, but it felt smoother." This comfort disappeared after there were multiple recovery procedures in the fourth iteration: "Robot paused and was stationary for a while. The robot then turned around in a circle unexpectedly when picking up the second cup. This made it seem as if the robot lost control". Overall, she commented that spending time with Fetch helped with the interaction: "After 1st interaction. Robot feels more natural and more easy going."

Another participant expressed annoyance when Fetch was delayed or executed the recovery procedure. "Weird actual interaction triggers, slow turn on a spot" commented one participant (*Participant 16*). This annoyance continued in the second encounter "More annoyed at the slow turn in front of me [and] with being slightly stuck in the corners." This lead to different feelings on the third encounter "[it] lingered after being handed the cup. Made me feel weird/uneasy." The delay at the start of the fourth iteration was also classified as strange: "weird long linger before the hand off made me nearly give him [*sic*] the cup too early."

Concerns of unpredictability and confusion were raised by another participant (*Participant 21*). Initially, there was a delay after giving Fetch the cup: "[the] reaction after handing [over] the cup was too slow, [I] didn't know if it recognized it." Additional delays in the second iteration did not help: "More unpredictable, as if it couldn't sense as well the environment." The delays in raising and lowering Fetch's body also caused issues "The pauses before/after asking for the cup make it seem more unnatural and unpredictable." He summed up all the iterations as needing improvement: "the movement made it seem very artificial and unpredictable"

6 Discussion

In this section, we discuss the three themes and what might be learned. Next, we introduce the concept of *movement acts* for examining a robot's motion, followed by an application of this concept to the Fetch robot in our experiment and provide other examples where it can be applied. We end with a challenge for researchers to make lemonade out of the lemon in a breakdown situation.

6.1 Examining the themes

The codes that emerged during the analysis of the responses were categorized into three themes: sentience, form, and competence. These themes emerged from the coding process and were not mutually exclusive.

Sentience had the largest variety of codes. Yet, what the codes all have in common is the clear use of either directly or metaphorically describe what the robot did or did not do. Following this line of thinking, we also noticed that it was possible to further distinguish the comments (especially) belonging in the *sentience* category between those that directly anthropomorphize (or sociomorph) Fetch, and those where such anthropomorphism is implicit in the language used to describe it. For instance, in many of the comments coded within the sentience theme, participants explain how it "felt like," or "was as if," or "seemed to me like" the robot did something that gave it a life-like character. Even though Fetch was not intended to be sociable in the experiment and thus limited in its social function, this was a recurring pattern. Perhaps it was due to the context of cooperating being of a social nature. It is also possible that the Fetch's movements in the room gave the participants an experienced sociality (such as sociomorphing or anthropomorphizing) in their interactions with Fetch.

Form had the fewest occurrences of codes related to it, and the codes that did emerge concerned only three "body parts." Still, there were some interesting trends that appeared. During coding, we noticed a certain overlap between the comments that described Fetch's form using anthropomorphic or zoomorphic terminology and comments that sociomorphed Fetch. For instance, small movements of the camera were perceived as social gazes or nods of a head. On the other hand, the comments about the arm were all about the lack of its function. Fetch not extending its arm when it stopped to receive the cup from the participant might be one reason for the lack of experienced sociability toward this particular part of the interaction. Perhaps putting Fetch's arm in a sling would have helped indicate that its arm could not be extended in the interaction. On the other hand, this change could have other unintended effects on the robot's sociability and might look inappropriate to the participants. Heider and Simmel's study [13] shows that the exact visual appearance of a moving object is only marginally perceived by a human observer. Instead, the quality of the movement is the predominant characteristic. The experiment does, however, show that this is an area that could be examined further.

Competence had the smallest variety in codes, of which there were only two. Yet, these two codes occurred more often than the others. A large number of responses explaining how the participants felt that the robot was slow. We understood these comments as concerning the nature of the interaction in a collaborative action, which often gave rise to frustration at the robot being slow at completing the task, even if it felt safe to collaborate with it. Another reason for this high occurrence of codes regarding Fetch's competence could be that the participants were answering an open question specifically asking how well they thought the robot handled the task.

While this is not included in our coding one interesting observation made going through the data was an apparent distinction in the responses of participants who appeared to use words with a *sentience* connotation without any apparent inner strife, and those who appeared less inclined to do this, but seemed to either not find, or not to bother finding other words to describe what is happening in the interaction. Fussell et al. [52] argued that it is easy for people to anthropomorphize robots in casual descriptions of robots because they use "ordinary" words. This can be related to the work of Seibt [53] who describes that varieties of "as if" (either explicit or implicit) in descriptions of human interactions with robots masks the social asymmetry of the interaction.

In general, it is difficult to discern whether the participants perceived Fetch as actually having the social and sentient abilities their words in their answers described, or if they were applied for a lack of a better way of expressing the experience. Many comments regarding Fetch's competence were direct answer to how well the robot handled the task, and might not be the result of any sociomorphing or anthropomorphism. On the other hand, if a participant described that they felt Fetch was "checking the room" it would imply a kind of perceived competence in the robot, as checking could be characterized as knowing what to look for and getting an overview of the situation. Further, a checking function like "looking" requires intent and purpose, which are relying on sentience. Further still, it would also be a kind of comment on how the participant perceived Fetch's form, because "checking" then also requires having visual perception. Further, a comment regarding Fetch "checking" could thus belong to all three categories, "looking" in two. Several of the comments were annotated with codes belonging to two or even all three themes. One method that could be used to examine this more thoroughly is the Linguistic Category Model [54], as is done by Fussell et al. [52] to examine linguistic anthropomorphism at different abstraction levels.

6.2 Movement acts

Knepper et al.'s [20] classification of intentional and consequential sounds can also apply to robot movement. A similar categorization for movements can enable us to better understand the implicit and explicit communication in Fetch's movements. During normal operation Fetch's movements were primarily functional, even though Fetch did not have any movements that were designed purposefully for social interaction and giving social cues. Because Fetch did not communicate explicitly with language or sound in the experiment, the communication was purely expressed through Fetch's movement across the room and what was explained via the facilitators. This was due to the original experiment examining different velocity profiles. The purpose was to see if the difference in the profiles communicated different information to the participants.

For example, Fetch's journey in each iteration was functional and intentional to collect cups and return it to the kitchen, but Fetch's rotations were functional and consequential as the movement "calculated a path" and "performed a recovery procedure" respectively without communicating any intended message. Still, it does not cover how meaning arises in a semiotic, triadic relationship between signifier, signified, and interpreter. That consequential movement or non-movement is present in the world for all present to observe, which can result in unintended interpretations of what that movement or non-movement *meant* [12].

Before the experiment, the participants were explicitly told that Fetch would be collecting cups. They were therefore aware that Fetch moved to collect cups and knew that this would be the purpose of the robot's approaching and stopping (having the implicit meaning of "now's the time to give the cup"). This means that even if the participants knew what the purpose was, *when* and *how* they should hand over the cup became unclear to many participants because they were expecting a social cue and hence still waited for the robot.

That the intention behind the implementation of the rotation has no explicit communication purpose, however, does not invalidate the experiences of people who interpret robot movement with a different meaning than intended—even if they are not quite sure what to make of it. Our case study has further confirmed that functional or consequential movements still communicate "something." But as the "message" being interpreted was not intentionally sent, what this "something" ends up meaning to an observer can be difficult to predict. As this case study has demonstrated, a robot's movement in the breakdown situation leads to different, possibly incompatible, interpretations by the participants.

Another of Fetch's consequential movements, or rather *non-movement* was its occasional delays, where it paused longer than usual before leaving a station. These pauses also brought forth puzzled comments from the participants about what the purpose of the delay was. During our analysis, we clearly saw that these pauses, even when the participants were not quite sure what to make of it, did not go by unnoticed.

Currently, there is no framework or concept that covers the triadic relationship of different meanings that might arise during interaction with robots and that acknowledges both movement and non-movement as social signs. Therefore, we draw inspiration from the concept of speech act [55] and introduce the concept of *movement act*. A movement act entails the understanding that both intentional and expressive movements *and* intentional and consequential movements might be interpreted by an observer as communication of inner state and intention. Just as *not speaking* is itself an act open to interpretation by the surroundings and its inhabitants, so will also *not moving* be an act (as a conscious or less conscious choice) open to interpretation. For example, a pause may only be a pause, but it may also imply a sense of insecurity or confusion.

That the situation is interpreted differently based on the robot's movements is in line with what others have already suggested: A robot, despite its limited social capabilities, is capable of communicating implicitly and explicitly using movements only. Using the concept of movement acts, we can isolate, identify, and characterize this phenomenon. We can then take each movement act and individually examine its implicit and explicit dimension. Movement acts can make sense of what a robot's movement communicates explicitly (or lacks to communicate). Being aware of the implicit dimension allows one to systematically look for interpretations that might happen during an interaction. The notion of movement acts facilitates a behavior design process that aims for an effective and clear communication between a robot and the people who interact with it.

6.3 Applying movement acts to robots

If we apply the movement acts concept to the original experiment, it can help explain some issues or provide suggestions for a better movement design.

First, although a robot's movement can communicate information, the original experiment did not find any significant difference in the perception of the slow in and slow out and regular velocity curves. So, was the *slow in and slow out* motion worth the effort? The previous article [1] outlined multiple reasons why that might have been the case. Yet given the participant's comments in the case study, it would appear that the robot's motion in a breakdown situation drew attention away from any other type of motion. That is, the movement acts in the rotation and delay captured more attention than the movement act in the velocity profile. Although the slow in and slow out movement act was meant to be implicit in its communication, it could have been too subtle. Perhaps a slow in and slow out velocity profile cannot be used alone, and may need to be used in concert with one or more animation principles—for example, *exaggeration* or *anticipation*—to capture sufficient attention.

The movement act of Fetch tilting its head up and down as it calculated its path gave depth information to the navigation stack and provided some context to participants watching that something was happening, but the participants' comments indicated that this movement act was ambiguous and communicated different information. The act must communicate more explicitly that Fetch needed more time. One way to do this could be additional movements such as slowing its head movement or performing a quick "double take" when the calculation started to take more time. Another possibility could be to combine the movement with other cues such as sound and light.

Likewise, Fetch's rotation movement act focused on the functional purpose for the navigation stack (re-calibrating its obstacles and position). On the one hand, we could have put more effort to avoid the situation entirely in the original experiment. One the other hand, this movement act could be modified to communicate its purpose to observers as an explicit, communicative motion. For example, perhaps Fetch might quickly raise and lower its torso before rotating, or it could just lower its head completely in a sign of defeat before rotating. As it's unlikely to avoid all breakdown situations, we recommend paying attention to the implicit dimension of all movement acts, including functional ones, to may make it easier to communicate a robot's current state.

This is where knowledge from other studies may be helpful. A model for mitigating breakdowns in HRI has been proposed based on a literature review [56]. The model suggested using visual indicators (LEDs, icons, emojis), secondary screens, and audio [56], but motion is not mentioned. The responses from the participants in our case study showed that motion communicates information as well. So, incorporating motion with these other modalities could strengthen communication for mitigating a breakdown. But, as Aéraïz-Bekkis et al. already reported, the discomfort that some participants expressed can be related to uncertainty about the robot's movements and its intentions [32]. Hence, if a robot's unexpected movement behaviors are causing discomfort (or fear), trust in the robot might be eroded, as the robot's performance is a large factor affecting trust [37]. This is congruent with the observation by Ogreten et al. [31] that soldiers never used a specific kind of robot in the field due to this robot's unexpected movements. Isolating the movement into movement acts can help identify where and why the uncertainty is happening, and provide places where additional or different motion may communicate more explicitly and remove the uncertainty.

Returning to the humans' expectations of a robot based on the robot's appearance [57], designing a movement act to express the navigation issue may help calibrate people's expectation that the robot may not be an expert navigator yet. Similarly, using movement acts to isolate the motion in a breakdown situation could lead to more legible motion for people to understand what is happening in the situation [58]. Using movement acts may also show that there is a need to add additional functionality to the robot (e.g., adding sound, lights, or extra moving parts) to aid in legibility or provide multiple modalities for communication.

There are many areas designers can turn to for inspiration to explicitly or implicitly communicate information through motion. Some sources of inspiration can be from animals or art. For example, Koay et al. [59] looked at how hearing dogs use movement to communicate with their deaf owners and transferred it to a humanoid robot. Participants were able to understand the robot's movement as communication and act upon them to solve a problem even when they had not been told the nature of the study. The original experiment drew inspiration from animation [28], but other areas such as puppetry [60] or dance [61] also offer inspiration. All these fields have dealt with issues of designing motion that can be understood by others, provide some expression, and set expectations by the people viewing the motion.

6.4 Making unexpected breakdowns expected

There are multiple ways to reflect on the case study. The case study might be seen as a cautionary tale. Researchers can try to control as much of the variables in an experiment, but issues still can show up. In this case, the robot may have built-in behavior that will take over if things don't work. It is good that a built-in behavior can resolve a problem, but one should consider how the people interacting with a robot will interpret the behavior. One might conclude that researchers should prioritize making robot robust, making the experiment meticulously planned, or controlling the entire experience by filming it and having participants watch it.

We would instead present this as a call to embrace the unexpected and design the breakdown situation into a study. Using the metaphor from Hoffman and Ju's designing with movement in mind [30], we would encourage researchers to design their experiments with the possibility of "robot breakdowns in mind". This does not absolve researchers and engineers from designing robust robots and well-designed experiments, but to accept that a breakdown may occur and have a plan to get data out of those situations. Moreover, we want to encourage authors to extensively report unexpected breakdowns to gain a deeper understanding of HRI.

Since these breakdowns may not happen for every encounter in a study, researchers will likely need to employ qualitative methods to explore the breakdown. One way of doing this could be to have a qualitative, semi-structured interview with the participants if a breakdown situation and see how they interpreted the breakdown or even if they noticed any sort of breakdown. This may mean that even if the participants' quantitative data may not be useful due to a breakdown, they can still provide qualitative information about their experiences and interpretations of the breakdown situation.

If experimenters desire more control and consistent experience, they could intentionally insert or trigger a breakdown situation during an experiment, even if the experiment doesn't primarily look at breakdowns. Since the breakdown is known in these cases, experimenters could design better ways of gathering data from the participants about the breakdown and how the participants interpret it. An inspiration for this approach comes from a long-term case study where participants developed their mental models of a robot shoe rack over several encounters with the robot changing behaviors every two weeks (with some unintentional errors from the Wizard) [62].

For example, if we had designed our experiment from Section 4 with breakdowns in mind, we could have used the opportunity to go deeper on things participants wrote and explored their opinions. It might have been possible to examine what participants meant when they said the robot was "waiting" or was "confused"? What actions from the robot made them think this? What made them feel uncomfortable and why? Alternatively, if the person felt that everything worked fine, why do they think that? Yet another approach could have been explored in the built-in navigation recovery. We could have found a reliable way to trigger the error to make the breakdown part of the experiment.

Answers to the qualitative questions may not be directly connected to the quantitative question being investigated in a study (our case study was not linked to the earlier experiment). The data collected from the interview questions, however, can provide a better understanding in future robot design and interaction. This could lead to insight into how to make breakdown situations easier to understand, or make people feel safer and more comfortable when such a situation occurs.

Of course, quantitative scales may also be useful for getting data about breakdowns. In the original experiment, there may have been an issue that the Godspeed Series might not have been sensitive enough to capture the change in perception during the breakdown situations. A different scale, such as the robot social attributes scale (RoSAS) [63], might have picked up participants' different perceptions of the robot that occurred during the breakdown situation.

Designing with breakdowns in mind could be formalized so that it can be part of any HRI experiment. During an experiment's design phase, experimenters could dedicate time to imagining possible breakdowns or other things that could go wrong. Additional sources for inspiration could include breakdowns that occurred in other pilot studies or experiments. From this work, the list of breakdowns would provide a starting point to determine what breakdowns could be prevented. For the breakdowns that are not prevented—either because they are unpreventable or they can be triggered in a controlled way—the experimenters could then plan qualitative or quantitative measures to record participants' reactions. This results in a set of breakdowns that the experimenters can prevent; a set of breakdowns that are not preventable, but expected; and a set of breakdowns the experimenters could choose to trigger. These sets will never be complete, but the steps in creating them provide preparation to handle the unexpected breakdowns not in the sets as well. Additionally, following these steps should add more realism in experiments, whether they are in the field or in a lab.

Breakdowns happen in many situations, inside and outside of HRI. It benefits all researchers to gather data from breakdowns to help improve future experiments and to understand how HRI can help improve a breakdown situation. Being willing to collect data from random, but expected, breakdowns in an experiment also is compatible with calls for bold HRI research [64] and to try research that goes beyond experimental psychology [65]. We can expect that as robots spend more time in less well-controlled environments, it will be necessary to also understand the extreme cases when interactions do not go as planned and researchers armed with methods to examine this area will find rich data that will improve future breakdown situations and HRI.

6.5 Limitations

We mentioned in Section 4 that we chose a within-subjects design. This decision could have affected the quantitative results, for example there could be a learning effect throughout iterations. Counter-balancing can help mitigate this effect, but it is difficult to say if it had an effect here. We also mentioned that a choice of within-subjects is less important for the case study as we are interested in participants' opinions during the breakdown situation and *not* the answers to the Godspeed questionnaire.

Regardless, even though all participants had a different experience with the breakdown situation, there still may be some learning effect for some participants who witnessed the breakdown situation more than once. This is where using the qualitative data in the case study is useful as the goal here is not to generalize, but to examine a phenomenon and learn from it to create better future interactions.

Having the kind of data set we had, our analysis could have benefited from using a framework such as the Linguistic Category Model [54], and would have strengthened this study. It would allowed us to conduct an analysis in which the descriptive action verbs used by participants to describe the robot were examined.

One could also argue that the participants from the studies about breakdown situations presented earlier also were not aware of the planned malfunctions. Those studies *are* looking at results that they can generalize. Our goal here was to show that even when things in an experiment are unexpected for the researchers and the participants, there are still possibilities to get data out the situation that may be useful. Here we were limited to one qualitative question that did not specifically consider the breakdown situation, but the themes from the comments lead us to developing a concept for better analyzing motion and understanding how the motion communicates information to humans.

Breakdowns are often an opportunity to return to the study design. If breakdowns have a fair enough chance of happening (e.g., one could argue 9% is fairly often in our case), it might be a good idea to spend time incorporating the breakdown into the study.

7 Conclusion

In an experiment that was designed to look at how a robot moves, we ended up with an unplanned phenomenon of a robot's recovery procedure although we had not designed the experiment to investigate this phenomenon. Given the serendipity of the situation, we used comments from participants to examine the phenomenon and found themes related to the implicit interaction of the robot's delays and the recovery procedure.

The themes reiterate that the robot's movement or lack thereof can be seen as a communicative act (i.e., a movement act). A movement act will be interpreted differently by people interacting with the robot. The robot's movement act, if it creates confusion or uncertainty, can possibly lead to humans losing trust in the robot (e.g., [31, 37]). But splitting a robot's movement into movement acts lets designers focus on the implicit and explicit communication and the act and create better communication.

Robot designers should consider that a robot's movement in a breakdown situation may cause an observer to be confused and try to interpret what it is doing. The movement act concept allows us to isolate the motion and examine the implicit and explicit information that is communicated by the motion. By focusing on what the motion communicates, it is possible to make the message clearer to participants and observers. Providing expressive signals, perhaps by using techniques from animation, may make the robot's movement easier to understand and thereby raise the human users' trust in the robot.

In addition, this study shows that there is additional information that can be extracted from experiments that may not have been originally under investigation. It is still important to strive for error-free operation, but there are things that may be examined even when breakdowns happen with a robot's performance. Breakdowns may also have consequences on how well a robot is able to learn or cooperate with a participant [66]. This points to additional considerations when designing a study to better capture unplanned situations that occur and still find interesting data from a study instance that might have otherwise been ignored in the search of answering different research question.

Breakdown situations have the potential to overshadow other effects that might have come up during the experiments otherwise. One way to eliminate these situations is more extensive pre-testing. But even when they occur, better post-experiment analysis and reporting of such occurrences can lead to better HRI research. We certainly plan on using the movement act concept and to gather and report data from breakdown situations in our future experiments involving robot movement.

Acknowledgments

This research was partly funded by the Research Council of Norway as part of the Multimodal Elderly Care Systems (MECS) project, under grant agreement 247697. We thank Robot House at the University of Hertfordshire for allowing us to use their facilities to design and run the experiment described in this paper. Thanks also to the interesting discussions and feedback from the SCRITA workshop at RO-MAN 2019 and the reviewers and colleagues who have read revisions of the article and provided insightful feedback. Finally, thanks to Wolfgang Leister who lent us his proofreading skills on later revisions.

References

[1] T. Schulz, P. Holthaus, F. Amirabdollahian, K. L. Koay, J. Torresen, J. Herstad. Differences of Human Perceptions of a Robot Moving using Linear or Slow in, Slow out Velocity Profiles When Performing a Cleaning Task. In: 2019 28th International Conference on Robot & Human Interactive Communication (RO-MAN) (IEEE, New Dehli, India, 2019)

[2] J. D. Lee, K. A. See. Trust in Automation: Designing for Appropriate Reliance. Hum Factors, 2004. 46(1), 50–80

[3] S. K. Devitt. Trustworthiness of Autonomous Systems. In: H. A. Abbass, J. Scholz, D. J. Reid, editors, Foundations of Trusted Autonomy, Studies in Systems, Decision and Control, 161–184 (Springer International Publishing, Cham, 2018)

[4] C. L. Bethel, Z. Henkel, K. Baugus. Conducting Studies in Human-Robot Interaction. In: C. Jost, B. Le Pévédic, T. Belpaeme, C. Bethel, D. Chrysostomou, N. Crook, M. Grandgeorge, N. Mirnig, editors, Human-Robot Interaction: Evaluation Methods and Their Standardization, Springer Series on Bio- and Neurosystems, 91–124 (Springer International Publishing, Cham, 2020)

[5] T. B. Sheridan. Eight ultimate challenges of human-robot communication. In: International Workshop on Robot and Human Communication (IEEE, 1997) 9–14

[6] M. A. Goodrich, A. C. Schultz. Human–Robot Interaction: A Survey. HCI, 2008. 1(3), 203–275

[7] S. Lackey, D. Barber, L. Reinerman, N. I. Badler, I. Hudson. Defining Next-Generation Multi-Modal Communication in Human Robot Interaction. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2011. 55(1), 461–464

[8] F. Hegel, S. Gieselmann, A. Peters, P. Holthaus, B. Wrede. Towards a Typology of Meaningful Signals and Cues in Social Robotics. In: 2011 20th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN) (2011) 72–78

[9] M. Green. Speech acts. In: E. N. Zalta, editor, The Stanford Encyclopedia of Philosophy (Metaphysics Research Lab, Stanford University, 2017), winter 2017 edition

[10] P. Watzlawick, J. B. Bavelas, D. D. Jackson. Pragmatics of Human Communication: A Study of Interactional Patterns, Pathologies and Paradoxes (W. W. Norton & Company, 1967)

[11] J. Maynard Smith, D. Harper. Animal Signals (Oxford University Press, 2003)

[12] R. Rommetveit. On 'Meanings' of Acts and What is Meant and Made Known by What is Said in a Pluralistic Social World. In: M. Brenner, editor, The Structure of Action, 108–149 (Blackwell, Oxford, 1980)

[13] F. Heider, M. Simmel. An Experimental Study of Apparent Behavior. The American Journal of Psychology, 1944. 57(2), 243–259

[14] J. Seibt, C. Vestergaard, M. F. Damholdt. Sociomorphing, Not Anthropomorphizing: Towards a Typology of Experienced Sociality. In: M. Nørskov, J. Seibt, O. Quick, editors, Frontiers in Artificial Intelligence and Applications (IOS Press, Vienna, Austria, 2020)

[15] A. Atkin. Peirce's Theory of Signs. In: E. N. Zalta, editor, The Stanford Encyclopedia of Philosophy (Metaphysics Research Lab, Stanford University, 2013), summer 2013 edition
[16] M. Q. Patton (SAGE Publications, Inc, 2015), 4th edition
[17] J. Seibt. Classifying Forms and Modes of Co-Working in the Ontology of Asymmetric Social Interactions (OASIS). In: M. Coeckelbergh, J. Loh, M. Funk, J. Seibt, M. Nørskov, editors, Frontiers in Artificial Intelligence and Applications (IOS Press, Vienna, Austria, 2018)
[18] C. Breazeal, C. D. Kidd, A. L. Thomaz, G. Hoffman, M. Berlin. Effects of Nonverbal Communication on Efficiency and Robustness in Human-Robot Teamwork. In: International Conference on Intelligent Robots and Systems (IEEE, 2005) 708–713
[19] A. Esposito, L. C. Jain. Modeling Social Signals and Contexts in Robotic Socially Believable Behaving Systems. In: A. Esposito, L. C. Jain, editors, Toward Robotic Socially Believable Behaving Systems - Volume II, volume 106, 5–11 (Springer International Publishing, Cham, 2016)
[20] R. A. Knepper, C. I. Mavrogiannis, J. Proft, C. Liang. Implicit Communication in a Joint Action. In: Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction - HRI '17 (ACM Press, Vienna, Austria, 2017) 283–292
[21] D. Moore, W. Ju. Sound as Implicit Influence on Human-Robot Interactions. In: Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction - HRI '18 (ACM Press, Chicago, IL, USA, 2018) 311–312
[22] D. Sirkin, B. Mok, S. Yang, W. Ju. Mechanical Ottoman: How Robotic Furniture Offers and Withdraws Support. In: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction - HRI '15 (ACM Press, Portland, Oregon, USA, 2015) 11–18
[23] P. Holthaus, S. Wachsmuth. It was a Pleasure Meeting You - Towards a Holistic Model of Human-Robot Encounters. International Journal of Social Robotics, 2021
[24] B. Mutlu. Designing Embodied Cues for Dialog with Robots. AI Magazine, 2011. 32(4), 17–30
[25] C. I. Mavrogiannis, W. B. Thomason, R. A. Knepper. Social Momentum: A Framework for Legible Navigation in Dynamic Multi-Agent Environments. In: Proceedings of the 2018 ACM/IEEE International Conference on Human-Robot Interaction - HRI '18 (ACM Press, Chicago, IL, USA, 2018) 361–369
[26] A. Zhou, D. Hadfield-Menell, A. Nagabandi, A. D. Dragan. Expressive Robot Motion Timing. In: Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, HRI '17 (ACM, Vienna, Austria, 2017) 22–31
[27] F. Thomas, O. Johnston. The Illusion of Life: Disney Animation (Hyperion, New York, 1995), 1st hyperion ed. edition
[28] T. Schulz, J. Torresen, J. Herstad. Animation Techniques in Human-Robot Interaction User Studies: A Systematic Literature Review. ACM Transactions on Human-Robot Interaction, 2019. 8(2)
[29] A. D. Dragan, S. Bauman, J. Forlizzi, S. S. Srinivasa. Effects of Robot Motion on Human-Robot Collaboration. In: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI '15 (ACM, New York, NY, USA, 2015) 51–58
[30] G. Hoffman, W. Ju. Designing Robots With Movement in Mind. Journal of Human-Robot Interaction, 2014. 3(1), 89–122
[31] S. Ogreten, S. Lackey, D. Nicholson. Recommended roles for uninhabited team members within mixed-initiative combat teams. In: 2010 International Symposium on Collaborative Technologies and Systems (2010) 531–536
[32] D. Aéraïz-Bekkis, G. Ganesh, E. Yoshida, N. Yamanobe. Robot Movement Uncertainty Determines Human Discomfort in Co-worker Scenarios. In: 2020 6th International Conference on Control, Automation and Robotics (ICCAR) (2020) 59–66
[33] R. van den Brule, R. Dotsch, G. Bijlstra, D. Wigboldus, P. Haselager. Do Robot Performance and Behavioral Style affect Human Trust? International Journal of Social Robotics, 2014. 6(4), 519–531
[34] S. Lemaignan, J. Fink, F. Mondada, P. Dillenbourg. You're Doing It Wrong! Studying Unexpected Behaviors in Child-Robot Interaction. In: A. Tapus, E. André, J.-C. Martin, F. Ferland, M. Ammi, editors, Social Robotics, Lecture Notes in Computer Science (Springer International Publishing, Cham, 2015) 390–400
[35] M. Ragni, A. Rudenko, B. Kuhnert, K. O. Arras. Errare humanum est: Erroneous robots in human-robot interaction. In: 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN) (2016) 501–506
[36] N. Mirnig, G. Stollnberger, M. Miksch, S. Stadler, M. Giuliani, M. Tscheligi. To Err Is Robot: How Humans Assess and Act toward an Erroneous Social Robot. Front. Robot. AI, 2017. 4
[37] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. C. Chen, E. J. de Visser, R. Parasuraman. A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction:. Human Factors, 2011. 517–527
[38] M. Salem, G. Lakatos, F. Amirabdollahian, K. Dautenhahn. Would You Trust a (Faulty) Robot?: Effects of Error, Task Type and Personality on Human-Robot Cooperation and Trust. In: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction, HRI '15 (ACM, New York, NY, USA, 2015) 141–148
[39] F. Correia, C. Guerra, S. Mascarenhas, F. Melo, A. Paiva. Exploring the impact of fault justification in human-robot trust: Socially Interactive Agents Track. In: Proceedings of the International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS, volume 1 (2018) 507–513
[40] M. K. Lee, S. Kielser, J. Forlizzi, S. Srinivasa, P. Rybski. Gracefully Mitigating Breakdowns in Robotic Services. In: Proceedings of the 5th ACM/IEEE International Conference on Human-Robot Interaction, HRI '10 (IEEE Press, Piscataway, NJ, USA, 2010) 203–210
[41] N. Mirnig, M. Giuliani, G. Stollnberger, S. Stadler, R. Buchner, M. Tscheligi. Impact of Robot Actions on Social Signals and Reaction Times in HRI Error Situations. In: A. Tapus, E. André, J.-C. Martin, F. Ferland, M. Ammi, editors, Social Robotics, Lecture Notes in Computer Science (Springer International Publishing, Cham, 2015) 461–471
[42] E. I. Barakova, P. Bajracharya, M. Willemsen, T. Lourens, B. Huskens. Long-term LEGO therapy with humanoid robot for children with ASD. Expert Systems, 2015. 32(6), 698–

709
[43] C. Bartneck, D. Kulić, E. Croft, S. Zoghbi. Measurement Instruments for the Anthropomorphism, Animacy, Likeability, Perceived Intelligence, and Perceived Safety of Robots. Int J of Soc Robotics, 2009. 1(1), 71–81
[44] M. Wise, M. Ferguson, D. King, E. Diehr, D. Dymesich. Fetch & Freight: Standard Platforms for Service Robot Applications. In: IJCAI 2016 Workshop Autonomous Mobile Service Robots (2016)
[45] J. Rose, D. W. Morgan, J. G. Gamble. Energetics of Walking. In: J. Rose, J. G. Gamble, editors, Human Walking, 77–102 (Lippincott Williams & Wilkins, Philadelphia, PA, USA, 2006), 3rd edition
[46] T. Schulz, J. Herstad, J. Torresen. Classifying Human and Robot Movement at Home and Implementing Robot Movement Using the Slow In, Slow Out Animation Principle. International Journal on Advances in Intelligent Systems, 2018. 11(3 & 4), 234–244
[47] Open Source Robotics Foundation. About ROS. 2019
[48] B. P. Gerkey, K. Konolige. Planning and control in unstructured terrain. In: In Workshop on Path Planning on Costmaps, Proceedings of the IEEE International Conference on Robotics and Automation (ICRA) (2008)
[49] M. J. Gielniak, C. K. Liu, A. Thomaz. Stylized motion generalization through adaptation of velocity profiles. In: 2010 RO-MAN: The 19th IEEE International Symposium on Robot and Human Interactive Communication (IEEE, 2010) 304–309
[50] L. D. Riek. Wizard of Oz Studies in HRI: A Systematic Review and New Reporting Guidelines. Journal of Human-Robot Interaction, 2012. 1(1), 119–136
[51] M. B. Miles, A. M. Huberman, J. Saldaña. Qualitative Data Analysis: A Methods Sourcebook (SAGE, Los Angeles, 2020), fourth edition edition
[52] S. R. Fussell, S. Kiesler, L. D. Setlock, V. Yew. How people anthropomorphize robots. In: 2008 3rd ACM/IEEE International Conference on Human-Robot Interaction (HRI) (2008) 145–152
[53] J. Seibt. Towards an Ontology of Simulated Social Interaction: Varieties of the "As if" of Robots and Humans. In: R. Hakli, J. Seibt, editors, Sociality and Normativity for Robots (Springer International Publishing, 2017)
[54] G. Semin. The Linguistic Category Model. (2011)
[55] T. Winograd, F. Flores. Understanding Computers and Cognition: A New Foundation for Design (Intellect Books, 1986)
[56] S. Honig, T. Oron-Gilad. Understanding and Resolving Failures in Human-Robot Interaction: Literature Review and Model Development. Front. Psychol., 2018. 9
[57] M. Kwon, M. F. Jung, R. A. Knepper. Human expectations of social robots. In: 2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI) (2016) 463–464
[58] A. D. Dragan, K. C. Lee, S. S. Srinivasa. Legibility and Predictability of Robot Motion. In: Proceedings of the 8th ACM/IEEE International Conference on Human-Robot Interaction, HRI '13 (IEEE Press, Piscataway, NJ, USA, 2013) 301–308
[59] K. L. Koay, G. Lakatos, D. S. Syrdal, M. Gácsi, B. Bereczky, K. Dautenhahn, A. Miklósi, M. L. Walters. Hey! There is someone at your door. A hearing robot using visual communication signals of hearing dogs to communicate intent. In: 2013 IEEE Symposium on Artificial Life (ALife) (2013) 90–97
[60] M. Luria, G. Hoffman, B. Megidish, O. Zuckerman, S. Park. Designing Vyo, a robotic Smart Home assistant: Bridging the gap between device and social agent. In: 2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN) (2016) 1019–1025
[61] A. LaViers, L. Teague, M. Egerstedt. Style-Based Robotic Motion in Contemporary Dance Performance. In: Controls and Art, 205–229 (Springer, Cham, 2014)
[62] M. Rueben, J. Klow, M. Duer, E. Zimmerman, J. Piacentini, M. Browning, F. J. Bernieri, C. M. Grimm, W. D. Smart. Mental Models of a Mobile Shoe Rack: Exploratory Findings from a Long-term In-the-Wild Study. ACM Transactions on Human-Robot Interaction, 2021. 10(2), 16:1–16:36
[63] C. M. Carpinella, A. B. Wyman, M. A. Perez, S. J. Stroessner. The Robotic Social Attributes Scale (RoSAS): Development and Validation. In: Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction, HRI '17 (Association for Computing Machinery, Vienna, Austria, 2017) 254–262
[64] R. Brooks. A Brave, Creative, and Happy HRI. ACM Trans. Hum.-Robot Interact., 2018. 7(1), 1:1–1:3
[65] K. Dautenhahn. Some Brief Thoughts on the Past and Future of Human-Robot Interaction. ACM Trans. Hum.-Robot Interact., 2018. 7(1), 4:1–4:3
[66] S. K. Kim, E. A. Kirchner, L. Schloßmüller, F. Kirchner. Errors in Human-Robot Interactions and Their Effects on Robot Learning. Frontiers in Robotics and AI, 2020. 7

Paper 4

Strengthening Human Autonomy. In the Era of Autonomous Technology

Soma, Rebekka; Bratteteig, Tone; Saplacan, Diana; Schimmer, Robyn; Campano, Erik; and Verne, Guri B. (2022) "Strengthening Human Autonomy. In the Era of Autonomous Technology," Scandinavian Journal of Information Systems: Vol. 34: Iss. 2, Article 5.

IV