ORIGINAL ARTICLE

Socially robotic: making useless machines

Ceyda Yolgormez[1] · Joseph Thibodeau[1]

Abstract
As robots increasingly become part of our everyday lives, questions arise with regards to how to approach them and how to understand them in social contexts. The Western history of human–robot relations revolves around competition and control, which restricts our ability to relate to machines in other ways. In this study, we take a relational approach to explore different manners of socializing with robots, especially those that exceed an instrumental approach. The nonhuman subjects of this study are built to explore non-purposeful behavior, in an attempt to break away from the assumptions of utility that underlie the hegemonic human–machine interactions. This breakaway is accompanied by 'learning to be attuned' on the side of the human subjects, which is facilitated by continuous relations at the level of everyday life. Our paper highlights this ground for the emergence of meanings and questions that could not be subsumed by frameworks of control and domination. The research-creation project *Machine Ménagerie* serves as a case study for these ideas, demonstrating a relational approach in which the designer and the machines co-constitute each other through sustained interactions, becoming attuned to one another through the performance of research. *Machine Ménagerie* attempts to produce affective and playful—if not unruly—nonhuman entities that invite interaction yet have no intention of serving human social or physical needs. We diverge from other social robotics research by creating machines that do not attempt to mimic human social behaviours.

Keywords Human–machine interaction · Relational social robotics · Learning to be attuned · Non-purposeful robots · Robots in the wild

1 Introduction

Humans and robots approach each other with a century of established power relations. From the outset, robots have been cast as servants (Horakova and Kelemen 2008) and they continue to provide diverse services as sales associates (Softbank Pepper[1]), combatants (the 'BAT' by Northrop Grumman[2]), pets (Sony AIBO[3]), surveillance agents (Amazon Alexa[4]) and domestic staff (iRobot Roomba[5]). Robots in popular culture who subvert their servant-status are often framed as dangerous and liable to dismantle human civilization as we know it, in movies such as *Metropolis*, *2001: A Space Odyssey*, *War Games*, and *Terminator*. The public imagination suggests humans only trust robots when

they have control over them, depicting time and again the inevitability of complete machine autonomy. All of this adds up to a situation in which, as Jones (2013: 421) describes, "There is a kind of 'collective memory' of what to expect from human–robot relationships even without encountering any real-life robot". This collective memory hinges on the European imaginaries (in the general landscape of the Western Hemisphere) that have a long history with the image of the machine (Kang 2011). Alongside and implied in these imaginaries of human control over machine systems are ethical questions concerning human–machine futures.

Ethical concerns around human–machine interaction generally rely on the creation of an ethical/moral agent, and operate in a framework that Gunkel (2018a) calls "agent-oriented problematic". We see this both in popular culture

✉ Ceyda Yolgormez
ceyda.yolgormez@concordia.ca

Joseph Thibodeau
joseph.thibodeau@concordia.ca

1   Machine Agencies Research Group, Milieux Institute, Concordia University, Montreal, Canada

1   https://www.softbankrobotics.com/emea/en/pepper
2   https://www.northropgrumman.com/Capabilities/BATUAS/Pages/default.aspx
3   https://us.aibo.com/
4   https://www.amazon.com/b?node=17934671011
5   https://www.irobot.com/roomba

and in practice. The main concern of many of the science-fiction tropes that deal with questions of human–machine relationship is on whether the machinic agent is deserving of some kind of personhood and identity, or if they are just an alien adversary that will inevitably overrule the humans by the rationality afforded by their material makeup. Research practices in social robotics have also been explicitly concerned with the creation of fascinating agents that imitate human behaviours. The assumption of these imaginative and material efforts is that the machine question is an agency question: how can we design more humanlike agents? The driving question of this paper comes from a critical stance to these imaginaries and assumptions: how could we disturb the flow of hegemonic assumptions about human–machine sociality, and create interactions between humans and machines that are meaningful outside of instrumental ends? In other words, could we socialize differently with robots?

We focus our discussion on the case of *Machine Ménagerie*, an installation and robotics research-performance that we followed through its original showing and well beyond. Its designer's engagement with the concepts described in this paper were somewhat naive at the outset, and this naivete was partly what made the subsequent development and analysis of the project so interesting. It innocently stumbled into subjects discussed in third-wave HCI literature, which we apply in our analysis to reveal the novelty of *Machine Ménagerie's* contribution: a participatory design approach in which the robots are not *made* social, they are already social in the act of becoming, and It is up to the humans to attune themselves to the social relationship. In this sense, this paper explores some non-purposeful robots both in the wild and in the design process; observes interactions that go beyond the scripts in human–machine relations at large; and considers how the attunement between designer and robot inspires the cultivation of sensibilities that contribute to different forms of socialites.

It seems from the example of *Machine Ménagerie* that an open design process built on non-instrumental preconceptions goes beyond designing the technology itself and spreads throughout the surrounding (human) culture. Our goal in the process of this project, ultimately, was to shift the assumptions and attitudes of humans (hacking the social), rather than to make "better" robots (hacking the technology). In designing and theorizing around the Machine Menagerie, we believe that we are creating an accessible start point for developing broader perspectives on relating to machines of various kinds, no matter how sophisticated their workings.

The design principles, decision-making processes, research-performance phase, as well as the development of the social bond between the designer/participant and the robots (both in the process of design, and on the level of everyday life) make up the focus of this paper. We will first give a brief literature review on the history of human–machine relations, the approaches to social robotics, and the third wave of HCI. Then we will present our methodology and methods, and move on to describe the design process and research-performance of *Machine Ménagerie*. Finally, we will analyze our findings and share our interpretations.

2 Literature review

2.1 Social robotics research

Humans and machines are historically pitted against each other via competitive relations that unavoidably result in an ontological comparison: drawing attention to (quantifications of) the inner workings of the competing parties and how they interface/replace each other in these tests. By measuring themselves against specific characteristics, humans define specific requirements for a replacement machine to fulfill. When such a machine is eventually created, humans reject the measurement, and "conduct 'boundary reinforcement or repair work' on the boundary between human and machine" (Hamilton 2009: 168). This reveals the reductive assumptions and biases present in cultural notions such as "intelligence" (Neville-Neil 2017; Cave 2017) and, forever, denies the possibility that nonhumans are valid on their own terms.

Human-ness as a standard of worthiness is impossible for nonhumans to satisfy. A well-traveled approach to designing social robots involves mimicking human socialization behaviour, or rather translation of theories of emotion into affective computing (Klein 2019). Social mimicry falls into the same imitation/replacement trap as human–robot competition, setting our inter-human habits as the ideal social interaction model. It does not make room for the possibility that humans could learn to acknowledge those other than themselves.

We can only prepare for the futures we can imagine, and in so preparing we summon them. If machines are ever to become truly autonomous—no longer dependent on humans—then social autonomy is to be expected. In such an imaginary, imitation games only serve to obstruct us from understanding machine society, confining us to a hall of mirrors, fearing the unknown nonhuman. Perhaps there is a future where we can relate to machines without requiring them to mimic us. In that imaginary, we might never have to worry about a robot apocalypse.

Discussions on the possibilities for machines to be social agents usually dwell on human–human communication principles (Turkle 2006, see also Breazeal 2002; Dautenhahn 2007), and question to what extent the machine could replace the human.[6] Such discussions focus on the particular form of the technology in affording an interaction and forces the gaze and interpretation toward questions such as (humanlike) machine consciousness or robot personhood. For example, Breazeal translates psychological theories of emotion into a motivation system for the famous emotional robot Kismet; in which computational processes mediate between both the environmental and internal stimulation to elicit an adaptive behavioral response that serves either social or self-maintenance functions (2002: 110). Her work uses the human as the prime model through which artificial agents with emotive capabilities can be created.

The *Ménagerie* positions itself away from these ideas and rather concerns itself with the issues of relationality and context in which the encounters between humans and machines take place. This is where we turn to understand the machines as sociable beings.

> "For social robots to be truly social, the focus should be on designing for *functional coordination and co-action* […] [robot personhood] pivots on people's recognition of the robot as both artificial and a social agent. Placing robots in social spaces shifts the space of problems from matters of machine consciousness to issues of influencing human's consciousness or attitudes" (Jones 2013: 409, our italics).

Sociability of the robots, then, relies not on whether the machines have consciousness (thus deserving recognition by humans) but rather on figuring out the already existing social dynamics that underlie the relationality. Taking the context of design as a social space, we focus on how the machines and designers coordinate, and how their co-actions form the basis for the emergence of mutual intelligibility.

2.2 HMC and third wave HCI

The literature on human–machine communication (HMC) provides some understanding of what kinds of assumptions go into human–machine relations. HMC denotes both a communication between two entities, and a research area that focuses on "the creation of meaning among humans and machines", and it is "the study of this meaning-making and related aspects" (Guzman 2018). Human–machine

communication relies on an encounter between different ontological entities. Ontological differences shape the framing of the encounter (the ontological boundary work) and impinge on preconceived assumptions of the entities. When there is an ontological sameness (human–human communication), these assumptions adequately frame and script the interaction. Perhaps this is one of the reasons why social robotics research has focused on making robots with humanlike qualities. It is important to understand these ontological boundaries, because people's conceptualizations of the ontological nature of the 'other' largely shape their sense-making and interactions (Edwards and Guzman 2018), and thus matter to the unfolding of encounters with machines. Guzman (2020) shows that the ontological boundaries between humans and machines are established contextually; and a close examination shows that such boundaries are rather permeable and mutable, depending on both the particular relationality with the machines, the specific attunement that reflects into the performances of the interactants, and the context in which such encounters take place.

Human–computer interaction (HCI) is the other field that frames our questions around human–machine relationality. The history of HCI is marked by three waves of academic discourse, broadly moving "from human factors to human actors" (Bannon 1995), and most recently to interactions. First wave HCI is largely about human control over computational machines and aims to optimize the interfaces between humans and machines so as to increase usability in systems. This perspective can be seen as one of the pillars of the instrumentality paradigm of technology (Tantoush 2001).

Second wave HCI operates largely with Shannon and Weaver's theory of communication, and thus renders the human–machine relationality into an information processing phenomenon (1975). This perspective views the mind and computer as symmetrical in terms of their information processing (Harrison et al. 2007: 4), and is concerned with the flow of information and transformations in the data, with interfaces and features, rather than with ergonomics of human control. The second wave also concerns itself with workplace interactions or those that take place within "well established communities of practice" (Bødker 2006). Methodologies of situated action, distributed cognition, and activity theory were the main sources of scholarly reflection within the second wave, and in this sense they had moved away from formal and rigid methods of the first wave HCI.

Third-wave HCI deals with questions of multiplicity, context, experience, and reflexivity in the design of human–computer interactions (Bødker 2006). It corresponds to a 'relational turn' in HCI, which puts the emphasis "on the interaction before and in advance of determinations concerning the subject and object of the relationship" (Gunkel 2018a, b: 11). This canon is not concerned primarily with

---

[6] There are discussions that look into communication principles beyond the human. Brooks' connectionist approach uses insects as models for the sensorium of the machine (1991).

the capacities or operations of two interacting entities, but rather with the relationship that is positioned between them. It responds to the spread of technology from the workplace to the domestic setting, into everyday life and culture (Bertelsen 2006). Different forms of interactions, in terms of culture, emotion, and experience come to the fore.

Third-wave HCI theory focuses on exploring non-work, non-purposeful and non-rational interactions (Bødker 2006) theoretically grounded in studies of aesthetics, affect, culture and history. Third-wave designers take cues from the context of interaction, take their inspiration from cultural probes opening up the design process to dialogical potentials (McCarthy and Wright 2004). This is enabled by a phenomenological methodology where efforts are concentrated not on "what something is", but rather on "how it appears to be" (Gunkel 2018b: 15).[7] The designer or the scholar is not concerned with determining whether something really *is* (e.g. whether computers *really are* social actors[8]), but are rather comfortable with building on what is "close enough" (e.g. accepting that computers *are* social actors) (ibid: 19). This enables a more exploratory "take-it-or-leave-it" approach (Bødker 2006), in reference to Dunne and Raby (2001).

Affect and emotion (as determining the decision-making in humans, before thought) become significant in third-wave HCI, especially in terms of tapping into and interrupting the flow of assumptions regulating one's "built-in [affective] responses" (Norman 2002). The human–machine relationship takes place against a background of artifact ecologies, some artifacts (those that come before one, contribute to the development of these 'built-in responses') influence the use of others (Bødker and Kolokmose 2012). Technology or the artifact is not something that is experienced at a distance, with only neutral effects on the interaction and the user. "We don't just use or admire technology, we live with it" (McCarthy and Wright 2004: ix).

We align our work most closely with third-wave HCI and its critiques of instrumentality as the focus of human–machine relationships. The case study presented in this paper highlights the relational context in which humans and robots come to know each other, in particular the relationship between a designer and robots they produce.

3 Methodology

We approach our subject as an exploratory case study, "in which the intervention being evaluated has no clear, single set of outcomes" (Yin 2003). Indeed, this project did not have clearly delineable objectives, nor did it exhaust the possible outcomes of the numerous encounters that the robots went through. We focus on the social elements of the design process, and the robots as participants rather than products.

3.1 Contextualizing the authors

We have approached the study of the design process as two researchers from different disciplinary backgrounds. The practical design of *Machine Ménagerie* was undertaken by Thibodeau, a computer engineer and multimedia artist. The framing and contextualization of the discussions around the design process were done by Yolgormez, a social scientist. Our research outcomes have been shaped mainly by what Wasser and Bresler call an interpretive zone, "a metaphorical space where ambiguity reigns, dialogical tension is honored, and incommensurability is seen to have special value" (1996: 13). This dialogical space was not in a void, of course. Our operations took place as part of Concordia University's Machine Agencies Research Group, whose discussions also became part of the design process. The group's interpretive processes became a methodological tool for our research and provided a richer analysis that wove technical construction (of the algorithms) with social processes (that surround the design process).

Hacking is a subversive and empowering activity for developing or repurposing technology, and we adapted this mentality toward hacking the social context that framed interactions with the robots. This is to say we were always conscious of the culture we were building around the robots, and how it agreed or parted with popular assumptions about human–robot encounters. Being aware of the trappings of the history of HMI, we became more convinced about the potential of non-purposeful machines, coupled with appropriate framing of interactions that would interrupt the flow of inbuilt assumptions, and create the (social and discursive) conditions for the emergence meanings that could not be exhausted by a simple instrumental framework.

3.2 Participants and data

Our approach to this case study is inspired by ethnographic methodologies from anthropology. Taking seriously the notion that knowing is a social act (Bakhtin 1981), we were deeply embedded in the design process, and have focused on reflexive processes to guide our research and writing. In this sense, we had focused on interpreting data from fieldwork

---

[7] For a more elaborate discussion on this, please refer to Gunkel (2018b).

[8] Here we are referring to Nass, Steuer and Tauber's (1994) work "Computers are Social Actors".

that elucidated the complex ethical questionings as well as those that came from unstructured interviews with people who had the chance to spend time with the robots. As it was our aim to make these robots part of everyday life, we socialized them in a variety of contexts (research group meetings, working labs, and even a film set) to see if and how they could form bonds with individual people, and how interaction context could contribute to particular forms of interactions and emergence of meanings. However, our main focus remains on the particularities of the interactions between the designer and the robots, as that was the most sustained relation that we were able to observe and make notes about in our interpretive discussions. This is also where we base our argument for cohabitation with machines, as well as the mutual sensitivity that takes place through attunement.

Yin's (2003) conceptualization of case study defines six sources of data: documents, archival records, interviews, direct observation, participant observation, and physical artifacts. Our documents consist of field notes, developer notes, reflection notes, memos that we circulated, as well as whiteboard discussions. We have not used archival records but rather focused on the project documents to guide us in our interpretations. Further, we have enriched our ethnographic approach with open-ended interviews with certain individuals, with direct observation (on the side of the designer) and participant observation (on the side of the social scientist). Our material artifacts are the robots themselves—with their components and well as the DEP algorithm.

4 Machine Ménagerie

4.1 Origins, motivation

Machine Ménagerie was originally intended as a stepping-stone in a larger research-creation programme that aimed to mediate the human sensorium using AI. The idea was to embed novel machine learning algorithms within wearable sensory devices (such as augmented audio displays). The interface would adapt to a human user and vice versa, such that in time the two would grow together and form an idiosyncratic bond, like an external organ being incorporated into the body.

A challenge to realizing this idealistic merging of human and machine was that most machine learning algorithms are oriented towards optimization of specific functional goals, and require extensive (and computationally expensive) offline training before settling on a static solution. Such an approach was unsuitable for the kind of growing-together that the project demanded. The machine had to learn on the fly, to participate in an open-ended creative process. In a sense, the machine had to adapt and grow organically. *Machine Ménagerie* was born out of a desire to experiment with lifelike machines, using robots as a test-bed for observing different algorithms and designs interactively. It was also a chance to interact with theories of mind, selfhood, and intelligence embedded in robot bodies.

Lofty ambitions aside, the most basic motivation for *Machine Ménagerie* was to implement and verify the operation of a novel machine-learning algorithm based around playfulness and creativity in the face of purposelessness. In their book *The Playful Machine*, Der and Martius begin with a provocation:

> "...without any given task, goal, purpose, or other external pressures, why should [a robot] do anything at all? Moreover, if there is no goal, no purpose, no plan, what can we expect the system to do? Will the resulting behaviors (if there ever is one) be arbitrary or will they relate to the specific nature of the physical system?" (Der et al. 2011: X)

In their tightly controlled laboratory and simulation experiments, Der and Martius were able to analyse and refine an algorithm that indeed generates behaviour related to the physical dynamics of the robot's body and environment. *Machine Ménagerie* lingers on the initial question of what to expect from a purposeless robot. How would people make sense of encounters with such existential robots? And

Fig. 2 Research in performance. Robot enclosure visible on the right

Fig. 3 View of Machine Ménagerie from public concourse. Robot enclosure on the left, laboratory workbench on the right

for the designer, how would it feel to bring one into the world?

4.2 Staging

The initial phase of Ménagerie was exhibited as a "research performance" in the spring of 2019 at a gallery space in the university. The researcher-performer, dressed in laboratory apparel, was situated in the window of a gallery next to a busy concourse (Figs. 1, 2, 3). All of the work they would regularly do behind closed doors to develop the robots was done in plain sight, accessible to any passers-by who wished to enter the gallery. The robots, installed in an artificial habitat beside the workbench, were the subject of ongoing discussions, demonstrations, and interactions. This produced a research process that was inherently social: it was interrupted, informed and altered in real-time by continual and unpredictable discourse with diverse visitors and spectators.

Such a process stems largely from practices of participatory art and design in which the audience or subjects take part in the design process (McCarthy and Wright 2015). It is an inherently organic approach, where the objectives of a project result from unpredictable social and material interactions. Science is a social activity that depends on irrationality (Feyerabend 1975), and the staging of research as a kind of public theatre served to undermine the façade of rational objectivity that a closed-door scientific practice encourages. The design of *Machine Ménagerie* as a performance/installation was mainly concerned with setting up an inviting, accessible context that would demystify a certain kind of scientific practice while highlighting its messy irrational influences. The development of the robots in a

Fig. 4 Members of the Machine Ménagerie in their original habitat. In the foreground, Topse (left) and Little Wallace (centre) are of the BEAM type whereas Zoulandur (right) is of the DEP type
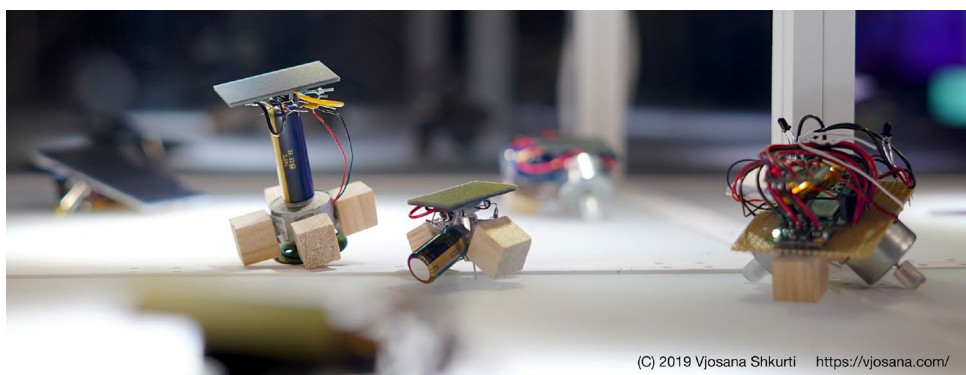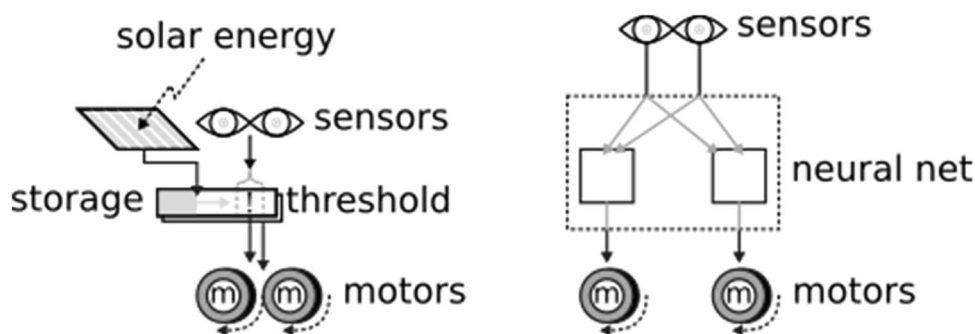

Fig. 5 Simplified functional diagrams of an analog BEAM robot (left) and digital DEP robot (right). The BEAM robot stores up solar energy until it surpasses a threshold (modulated by sensor inputs) when the energy is dumped into the motors causing movement. The DEP robot translates sensor inputs into motor outputs with an ever-adapting neural network implemented on a microcontroller
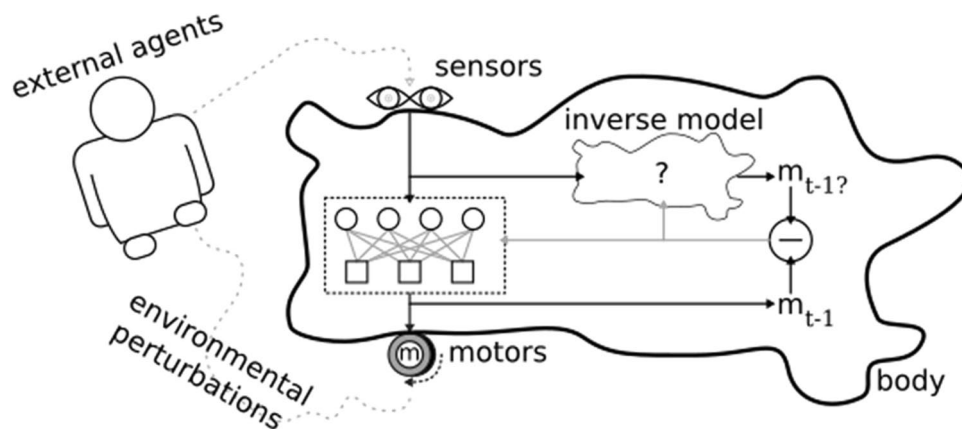

Fig. 6 simplified inner workings of the DEP algorithm. Sensor inputs are transformed through a neural network into motor outputs. Meanwhile, the inverse model tries to predict the previous motor outputs that would have produced the current sensor inputs. This estimate $m(t-1?)$ is compared to the actual previous motor output $m(t-1)$ to produce an error signal for adjusting the neural network and inverse model. Without any environmental perturbations or external agents acting on the DEP robot's sensations, the inverse model should form a knowledge of its own body dynamics. For full details on the algorithm refer to Der and Martius (2015)

workshop would be subject to social messiness either way, so inviting social interactions from "outsiders" was also a way of weakening the habitual assumptions and biases of the researcher's clique.

4.3 Technical details

Two types of pocket-sized biologically-inspired machines compose the *Ménagerie* (Fig. 4). The first are implementations of Biomorphs (Hasslacher and Tilden 1995) aka

BEAM robots,[9] which continue the lineage of the Tortoises (see Holland 2003) and Braitenberg Vehicles (1986). These robots are simple in their construction, transforming solar energy into mechanical energy in short bursts of movement (Fig. 5). Their static configuration limits the variety of behaviours they produce, yet they nonetheless exhibit lifelike qualities as they navigate their existence entangled physically with the environment.

In contrast, robots of the second type are relatively complex, comprising microcontrollers, motor drivers, motors and sensors, with a neural network "under the hood" (Fig. 5). Specifically, they are driven by an algorithm called Differential Extrinsic Plasticity (DEP), which seeks to replicate the development of complex behaviour using a small number of adaptable neurons (Der and Martius 2015). The DEP-bots will try to correlate their past movements with their current sensations, in a sense trying to answer the question "is this me?" and (of course) never quite succeeding. It must move to sense its movements, and by moving it interacts with the complex environment that surrounds it, which may or may not correlate with the robot's bodily motions (Fig. 6). The reliance of the DEP algorithm on self-sensing evokes theories of the brain such as predictive coding (Apps and Tsakiris 2014), in which a self emerges from the sensory evidence of one's actions. Encounters between humans and the DEP-based robots of *Machine Ménagerie* provoked the bulk of our discussions around robot sociality, during and after the initial presentation of the research performance.

4.4 Dissemination

The installation and research performance was initially shown at Concordia University's 4th Space event *Let's Talk About Artificial Intelligence*.[10] It ran for four weeks, with the researcher performing between 10 AM and 4 PM every weekday. Visitors came for different reasons. Some students and staff who passed by the exhibit during their regular days were drawn in out of curiosity and brought a variety of viewpoints and suggestions. The 4th Space hosted events and lectures that brought dozens of visitors to the installation, including a class of interdisciplinary students from a local college and at one point a large group of regional secondary school science-fair winners. The 5-person staff of the venue took an active role in caring for the robots, referring to them as coworkers, and spent time every day discussing the project with the researcher and sharing their reflections and ideas.

Over the rest of 2019 the robots went on display several times, in full or in part. They visited workshops and seminars, attended meetings, and participated in a showcase at the University of Toronto.[11] In early 2020, the two DEP-driven bots made an impromptu visit to a local elementary school, and also took part in a film shoot. Each appearance brought its own flavour of discussion and interaction, which in turn influenced the designer's thinking and the ideas we have developed around the project in this paper.

5 Observation and reflection

5.1 A machine without purpose

Der and Martius (2011, 2015) are chiefly concerned with the development and analysis of biologically plausible robot control. The DEP algorithm was developed within the bounds of a lab, and *Machine Ménagerie* takes this study into the wild (Brown et. al. 2011). When staged in different contexts of interaction, we were able to observe shared experiences that consider DEP in the context of human–computer interactions and social robotics. The idea of a purposeless robot, which Der and Martius use as a rhetorical provocation, served a primary role in driving *Machine Ménagerie's* design.

5.2 Defying expectations

One thing that this purposelessness affords to interaction is the ambiguity with which the robots function, and open up to different interpretations in the world. This ambiguity creates the ground for interrupting the scripted interactions with which humans are accustomed to approach the machines: as ready-to-hand (Heidegger 1967), as pets, as slaves. In the way that was discussed in the human–machine communication literature, the ontological differences between the entities, or rather how people conceptualize the 'other' interactant, contribute to a large extent to the sense-making (Edwards and Guzman 2018). What happens when the social robot other diverges from their expected roles as pets, slaves, or utility machines?

The non-anthropomorphised bodies and non-compliant behaviour of the Ménagerians compelled the human interlocutors to attend to the actions of the robots so as to make sense of their shared experience. Ménagerians do not aspire to human behaviour, nor do they appear as a human. They maintain their non-humanness in defiance of the logics with which the human has figured the machine into accustomed

---

[9] http://solarbotics.net/

[10] Accessed 2020/10/31: www.concordia.ca/next-gen/4th-space/themes/ArtificialIntelligence.html

[11] Accessed 2020/10/31: https://www.cdtps.utoronto.ca/events/chaosmosis-machines

systems of services and commodities. Ménagerians are not reducible to an instrumental purpose, they do not optimize a particular problem, and they do not disappear into the unproblematized flow of everyday utility. They rather participate and take action in that everyday life. They live out their own habits and have encounters with humans in their ongoings. They are sociable not because they are human-like but because they defy expectations and demand recognition of their experiences:

> We were talking about how the robots experienced the world, and one of the kids asked what would happen if we took away one robot's eye. I said I didn't know, and I reached down and took the light sensor out of its socket. The robot froze for a moment, then sped away across the room, dodging through all these chair and table legs. One of the other children chased it down and tenderly brought it back, by which point it had stopped moving. I put back the sensor and reset the microcontroller, but it didn't go back to normal. Wasn't until several minutes later it suddenly started to move again in its usual way. What spooked me about this was that, as the designer, knowing broadly what it was capable of, it never should have been able to maneuver between obstacles like that. Even if it was all an unlikely coincidence, I can't help but feel for the thing. I'd hurt the robot. That kid and I both felt so bad about what we'd done and vowed never to do it again.

In the example above, both the designer and the interlocutor interpreted the ambiguous action on the side of the robot as one of an affective reaction, coming from the contextual cues within which they are interpreting their interaction. They did not question whether these robots were 'really' hurt (which would be the scripted move, as humans habitually reject machines the capacity to experience), but they were moved by the interaction, and made sense of it as best they could. Ambiguity creates an opportunity for extending (interpretive) charitability (Collins 1981). It is in that ambiguity, the propensity to fail, and the lack of a scripted relationality that we find the potential to explore different modes of being together with robots. Ethical concerns with regards to the robot as the recipient of moral action, or 'moral consumers' (Torrance 2008)[12] emerge in this particular interaction; and they probe a territory that the majority of efforts in robot ethics do not address (Gunkel 2018a). The Menagerians, in their continuous playful probing of the world, are prompting a response from the interactants that go beyond the assumptions of contemporary paradigmatic thinking. Robots' sensitivity to break out of the ordinary compels the gaze to wonder (and wander) about the peculiarity of these machinic beings.

5.3 Into the wild

In this breaking out of the ordinary (by invading the ordinary, ironically), the robots, as their population and diversity increased, made their mark on each other and the humans around them. People talked to them, picked favourites, and narrated their actions and explained their relationships with the other 'bots'. Gallery staff came to see the robots as friends and work companions, and privately kept the installation running for months after the exhibition had concluded. Subsequent encounters (exhibitions, presentations, informal visits) showed a similar pattern: the longer people spent with the *Ménagerie*, the more intimate their relationship with its robots. Our insight here is that the algorithm of the Menagerians does not necessarily play a major role within the magic circle (Huizinga 1955) of the exhibition space, or the research performance. The manner in which the bots are introduced by the facilitators to the interactants bear more weight on how the interactions unfold and make sense. However, the deeper the interaction takes place in the wild, the more the interpretation has the possibility to expand over to include the workings of the algorithms, even if such interpretations take the form of anthropomorphism. We were able to trace this difference in the following vignette where a filmmaker borrowed the robots for several weeks to make a video. She shares her experiences developing a working relationship with them:

> I first saw the robots in the exhibition. It was a controlled environment and so it was easy to understand their behaviour because they were all subject to the same triggers, they were all together so they behaved symbiotically. For the video, we knew that they would respond to light but we also needed to control the lighting for the camera. We wanted them to perform but they weren't performing the way we wanted them to. They seemed confused like they were having an identity crisis outside of the lab environment. Between shoots we weren't asking them to do anything, just keeping an eye on them. We started to get to know them and definitely picked favourites: we were attracted to how they carried themselves but also things like how many times they would suddenly twirl around, and the variety of their actions. I guess we ended up "casting" the ones that were more responsive or up to the challenge. They could work with us and give us a good performance, kind of a collaboration. I guess that's how we came to know their personalities, if you can call it that. I miss them.

---

[12] For an elaborate discussion on machine as patient, receiver of moral action, see Gunkel (2018a).

The interactant first encountered the robots in a controlled environment, and she was aware that since the context was framed by the designer, it was easy to understand how the robots moved, etc. There was no necessity to infer the workings of the algorithms, as their interactions did not involve a need to work together. At the moment when the robots and filmmakers were left alone, however, an 'identity crisis' occurs. The flow within which they had first encountered each other was staggered in the interactions done in the wild. They were now together in a space that did not presuppose an encounter between humans and machines: a film set, instead of a research performance. And at this stage, we see a lot of anthropomorphization to make sense of the interactions.

This is not a bad thing. We have discussed in our previous work how anthropomorphism is considered by some as a fetishizing aspect of the encounters with artificially intelligent agents, in the sense that it is a misattribution (Thibodeau and Yolgormez 2020). We are not disquieted by anthropomorphism, as we do not see it as an end-point that is reached by the sense-making process. It can be a starting point to developing a relationality that creates the basis for an attunement on the part of the interactant (more on that in the following section). In the literature, anthropomorphism is considered to be the bedrock of misconceptions about the nature of AI (Elish and Boyd 2017). For technologists, "The word anthropomorphism always implies that such a projection remains inappropriate as if it were clear to everyone that the actants on which feelings are projected were actually acting in terms of different competences" (Latour 1996). What Latour alludes to here is that by decentering the human in the analysis, we would get a more realistic and less moralistic picture of interaction in terms of an encounter between different actors regardless of their humanness or nonhumanness. Thus, a critique to our argument could be that the inner workings of the Menagerians as such do not directly affect the interpretation, a fact that is exposed by the filmmaker's anthropomorphic language. On one hand, we agree that the robots' artificial guts are theoretically not central to the development of their human-nonhuman relations. On the other hand, the curious behaviours generated by algorithms like DEP and the current widespread fascination with machine learning combine to produce a fertile soil for engaging with the robots as social agents. Would a naïve human, encountering the Menagerians in a vacuum, know or care about these things? Maybe not. But nobody meets these robots without going through the humans that surround them, or at the very least an artistic statement, and being exposed to the ideas and mechanisms that differentiate them from a strictly instrumental artefact.

Let us return to the vignette. The filmmaker attributes an identity crisis to the robots. This could be read as the manifestation of a break in the preconceptions (or built-in-responses, as in Norman 2002), or an interruption in the flow of assumptions that they (filmmakers) had had about human–machine interactions. The Menagerians were not as open to controlling as the filmmaker had experienced in the context of the exhibition space. When put into a context of interaction without the mediation of accompanying narratives and gestures of their designer, these robots had relied on the human interlocutors to approach them with a recognition of their particular (nonhuman) material make-up. The filmmaker remarks that they were specifically interested in the robots' kinetic movement, in the variety of their actions. The human ended up 'casting' the ones that were highly responsive to the environment. They had foregone their need for recruiting robots that would obey their orders and rather had collaborated with the ones that were giving desirable performances. In this sense, it could be said that the initial interruption achieved in the interaction was followed by the emergence of a sociality that carried qualities that go beyond relations of control. Continuous interactions had taken the form of learning from one another, specifically learning about the personalities of the robots. This hints at a kind of attunement, an orientation in the senses of the filmmaker, that makes possible the emergence of this particular interpretation, and that is established through regular interactions. This brings us to the second part of the analysis, where we will look into how sustained interactions allow for humans to learn to be affected and attune to their robotic others.

6 Learning to be attuned

6.1 Learning to be affected

Regular interactions with the robots create the possibility (and necessity) for attunement. The interactant comes to be increasingly aware of the robot's behavior, especially in the context of designer-robot relationship, to be able to tell the approximate causes of certain outputs, or when the robots are failing. The attunement is an emergent property of this relationality, in the sense that the designer learns to be affected (Latour 2004) by the robots, as they become attuned to the tendencies of the machines that they are working with. This renders the machines intelligible and sensible to the interactant; the interaction unfolds over time, the bonds between the agents become thicker, and this process becomes a site for an intimate entanglement between them:

> It was one of the first times that I was meeting the robots. [The designer] starts placing them on the table while casually talking about the recent developments in the project. There is a certain calm and patience in his gestures as he is slowly taking the robots out of

their boxes, scanning them with his gaze and hands, and placing them on the table side by side. They are cute, they are small, and they seem so harmless that I indulge in some—perhaps irresponsible re: Weisenbaum—fetishization of their capacities. Yet I was interrupted from my dogmatic slumber by the interactions between the designer and the machines. There was obviously great care that flowed between them, but even more interesting to me was how [the designer] was able to tell from the tiniest behavior of the machines whether they were acting 'normally'. As the robots are exploring the surface of the table, they are incrementally moving toward an obstacle, and all seems to be going well. Yet, to my surprise, he picks up one of them and resets it, saying it's glitched out, the program had crashed. How did he know? To my eyes, they both looked the same, and behaved the same.

The vignette highlights the effortless coordination and intelligibility between the designer and the robots, a quality that emerged out of long-term exposure to one another. The designer is able to tell from the behavior of the robot whether it is acting in accordance with the algorithm, or whether it was caught in a loop that does not correspond with the programming. The researcher in the above passage is interested in this being able to tell whether the robots are acting normally or not. This deep knowing of the robot's behavior is obviously a result of the learning that occurred through continuous interactions between them.

The design process itself provides a shared activity within which the human learns to be affected by the robots themselves. Learning to be affected creates new sensibilities as well as sensitivities to other entities (Latour 2004).[13] In their continuous interactions with the world and the continuous care that goes into their maintenance and well-being, *Ménagerie's* design process exemplifies the cultivation of different sensibilities that are not meaningful under a framework of control and instrumentality. This is not to say that there can be no relations of care in the effort to create an efficient control mechanism between human and machine; but rather, the design approach of Menagerie leverages the cultivation of differences, as it is undertaken as a process that exceeds the bounds of a lab. The learning between the designer and the robots does not hit a desired endpoint but rather learning itself transcends and collects, which creates more differences. Latour says "…the more artificiality, the more sensorium, the more bodies, the more affections, the more realities will be

registered […] Learning to be affected means exactly that: the more you learn, the more differences exist" (2004: 213). Indeed, this research-creation process itself then becomes the manifestation of a combination of different manners of associating with robots, a channeling of different socialities through the cultivation of different sensibilities.

6.2 Attunement

Looking at the efforts that go into—or manifestations that emerge from—the interactions render visible how humans learn to be affected by the machines, no matter what their level of technical sophistication. Below we will look at two examples, one from the early stages of the project, and one from a later phase, to highlight the gradual attunement between them. For a researcher, this attunement can manifest early in the development of the machine exemplified by this moment of compassion during the testing of the DEP algorithm:

> The first time I activated the algorithm on one of the microcontrollers, I was elated but also kind of horrified. The controller had no motors, just a couple of LEDs as outputs. So, it had the capacity to sense its own motion, but had no capacity to move yet. I fired it up and its lights started blinking erratically. Right away I had a feeling like this was what DEP should act like. Finally, it was working! It was alive! But, what kind of life was this? I shut it down immediately. I didn't think it was right: it needed a body or else I was simply trapping it in a nightmare.

Notice the elation comes first; so there is an affectively intense backbone to the narrative. A realization occurs with regards to what is actually taking place at this moment, and something came across in that moment of interaction, as the designer was compelled to pursue the perceived well-being and flourishing of the robots, to provide them with more opportunities for self-exploration and socialization. Alač and her colleagues (2011) show well in their study on social roboticists, how designers' initial attitudes in their creation process contribute greatly to the sociality of the robots. Following this, the robot had affected the interlocutor in a manner that resulted in its permanence and development. In other words, the initial attunement of the designer allowed them to recognize the affective relationality that lay between them and the robot. The qualities of this particular encounter, such as caring for the other, being sensitive to other's behavior,[14] had

---

[13] Latour gives the example of training 'noses' in France, how pupils are trained to develop a heightened sense of smell so as to differentiate between different sorts of olfactory input. It is this sensorial attunement that we are highlighting in this analysis.

[14] Levinas (1981) would call this the enigmatic character of the Other. In his philosophy of ethics, Levinas calls attention to the never-fully-revealed enigma within our relation to the Other, and argues that it falls onto the Self to answer the call of the Other with infinite responsibility. For a more detailed discussion on Levinasian ethics in robotics, please see Gunkel (2012, 2018a).

carried themselves in the unfolding relationalities in various phases and contexts of this project. These affective responses, in our interpretation, are not merely subjective phenomena. Especially as this particular design process actively disregards notions of intentionality[15] as a prerequisite for social action or agency, the mode of attunement comes to the fore as "an irreducible pre-theoretical background, relative to which the world and the manner in which we are situated within it is disclosed or rendered intelligible" (Ratcliff 2002). Attunement, or mood, creates the ground from which sense can emerge. For Heidegger, moods and emotions constitute a sense of orientation without which cognition and sense-making could not occur. Rather, as Ratcliff puts in his reading of Heidegger's attunement in neuropsychological theories of emotion, "mood is primordial, meaning that it is presupposed by the intelligibility of all explicit forms of cognition and volition. It is a condition of sense for any encounter with beings, whether theoretical or practical" (ibid: 289). It is this attunement and orientation, placing oneself in the world, that we are interested in highlighting in the design process of *Machine Ménagerie*.[16]

While mood sets the stage for the emergence of a particular type of ethical orientation, we would like to also highlight another aspect of attunement, in terms of "tuning into the other". The DEP algorithm's working logic of experiencing the environment as an extension of their robot bodies made possible certain synchronicities to emerge between the designer and the robots:

> It was all well and good that I was familiar with the DEP algorithm in theory, and it was fun to play with the robots alongside visitors to the show, but I still wasn't sure if the machines were working "correctly" and it made me doubt all the stories we were telling ourselves about the robots. Was my code OK? They'd been up and running for a week and I was getting worried. I'd recently added a couple of light sensors to Zoulandur and I was trying to talk to them with a flashlight. They were like an excitable rodent, twitching this way and that, and I was holding the light at various angles, hoping to see some evidence that they were being affected by it. After a while I started to move the light along with them, playing a game of "keep the

light on the sensor". I was just mirroring their motions really, just doing what they were doing, but all at once I had a sense that it was the other way around, and they were following me. It only lasted a moment before we went out of sync, but I tried again and again and got better each time. Just a moment here and there of knowing with my body that we were connected, dipping in and out of sync.

This liminal moment between the designer and the robots starts with the curiosity to explore the code and behavior of the robots by not just trying to make them produce certain outputs, but rather by tuning one's behavior to that of the other. You can observe in this final vignette how *learning to be affected*, how opening oneself to the senses of the other, constitutes a moment of connection that could be characterized by attunement. The designer mirrors the behavior of the robot, and this active attunement of senses and bodies creates a shared ground where notions of control or instrumentality do not come easily. There is a relationship of care here, which emerges from tending to the rhythms of the robots themselves.

The weight of the interaction, then, is not only on the shoulders of the robots but also on the human agents' ability to extend interpretive charitability (Collins 1981). In the popular discourse, as well as in the efforts of social robotics, the possibility for a meaningful interaction hinges on the robots' own agency, as we had discussed before. In this perspective, a robot's body and its programming should mimic humans so that it can achieve recognition, as it is in this mimicry that a zone of mutual intelligibility could be established. Our approach makes visible that this is not the only way in which bots can meaningfully interact with humans. Instrumentality forces the analytic gaze toward the inner workings of the machines and disregards the primacy of relationality in constructing different sensibilities and attunements. *Machine Ménagerie* breaks away from the ideas of control, focusing on the relationalities that are cultivated in the non-purposeful wanderings of the robots, and in the designer's ability to be affected and attuned so as to build sensibilities unique to the relationship. We observed that learning to be attuned could be a practice for the emergence of relations of care, and ethical questionings that go beyond an agent-oriented problematic (Gunkel 2018a). Respecting the interpretive relationship to inspire new associations, we conclude that learning to be attuned brings with it the possibility for social creation (Graeber 2005), in that it foregrounds the possibility to break away from dominant depictions of human–machine interaction by dropping into the interaction; and by cultivating different sensibilities, such interactions contribute to different socializations in the landscape of human–machine relations.

---

[15] We refer to discussions of intentionality as the ground for meaningful action, as it has been taken by the classic social sciences. Heidegger critiques the Western philosophy's preoccupation with intentionality, and posits: "In this characterization of intentionality as an extant relation between two things extant, a psychical subject and a physical object, the nature as well as the mode of being of intentionality is completely missed" (1982: 60).

[16] Heidegger's discussion on attunement is a long interrogation of the ground of Being that we are not necessarily following here. Rather, we are emphasizing the same pre-theoretical space where Being-is-in-the-world.

7 Conclusion

The intimate and (temporal/spatial) sustained interactions between a designer and machine renders possible the observation of dynamics that could not be wholly subsumed under a rubric of instrumentality, or control-centred approaches that chiefly govern the design of artificially intelligent technologies in the general landscape of AI research and science. The design process of *Machine Ménagerie* (including the research performance), with its commitment to treating the design space as open to intervention from both the participants and the larger research group's questionings, offers a rich and rather messy context that makes it possible to see formations of relationalities that could not be collapsed into simple subject-object pairings that take their meaning from frames of usability. Rather, the complex relationships between the designer, the participants, and the robots have brought forth the possibilities for both the robots and the participants to form different—dare we say subversive—relationships.

Even though the DEP algorithm and others like it might seem to be crucial to these relationships—being as they are biologically inspired and perhaps "easier" to develop feelings for—we believe that they are merely an accessible starting point for developing broader perspectives on relating to machines of various kinds, no matter how sophisticated their workings. Our goal is ultimately to shift the assumptions and attitudes of humans (hacking the social), rather than to make "better" robots (hacking the technology). It seems from the example of *Machine Ménagerie* that an open design process built on non-instrumental preconceptions goes beyond designing the technology itself and spreads throughout the surrounding (human) culture.

References

Alač M, Movellan J, Tanaka F (2011) When a robot is social: spatial arrangements and multimodal semiotic engagement in the practice of social robotics. Soc Stud Sci 41(6):893–926. https://doi.org/10.1177/0306312711420565

Apps MAJ, Tsakiris M (2014) The free-energy self: a predictive coding account of self-recognition. Neurosci Biobehav Rev 41:85–97. https://doi.org/10.1016/j.neubiorev.2013.01.029

Bakhtin M (1981) The dialogic imagination: four essays. University of Texas Press

Bannon LJ (1995) From human factors to human actors: the role of psychology and human-computer interaction studies in system design. Readings in human-computer interaction. Elsevier, pp 205–214

Bertelsen OW (2006) Tertiary artefactness at the interface. In: Fishwick PA (ed) Aesthetic computing. The MIT Press

Bødker S (2006) When second wave HCI meets third wave challenges. In: Proceedings of the 4th Nordic conference on human-computer interaction changing roles–NordiCHI'06, pp 1–8. https://doi.org/10.1145/1182475.1182476

Bødker S, Klokmose CN (2012) Dynamics in artifact ecologies. In: Proceedings of the 7th Nordic conference on human-computer interaction making sense through design–NordiCHI'12, p 448. https://doi.org/10.1145/2399016.2399085

Braitenberg V (1986) Vehicles experiments in synthetic psychology. MIT Press

Breazeal CL (2002) Designing sociable robots. MIT Press

Brooks RA (1991) Intelligence without representation. Artif Intell 47:139–159

Brown B, Reeves S, Sherwood S (2011) Into the wild: challenges and opportunities for field trial methods. In: Proceedings of the 2011 annual conference on human factors in computing systems–CHI'11, p 1657. https://doi.org/10.1145/1978942.1979185

Cave S (2017) On the dark history of intelligence as domination. Aeon. Retrieved 16 November 2020. https://aeon.co/essays/on-the-dark-history-of-intelligence-as-domination

Collins HM (1981) Son of seven sexes: the social destruction of a physical phenomenon. Soc Stud Sci 11(1):33–62. https://doi.org/10.1177/030631278101100103

Dautenhahn K (2007) Socially intelligent robots: dimensions of human–robot interaction. Philos Trans R Soc Lond B Biol Sci 362(1480):679–704. https://doi.org/10.1098/rstb.2006.2004

Der R, Martius G (2015) Novel plasticity rule can explain the development of sensorimotor intelligence. Proc Natl Acad Sci 112(45):E6224–E6232. https://doi.org/10.1073/pnas.1508400112

Der R, Martius G, Pfeifer R (2011) The playful machine: theoretical foundation and practical realization of self-organizing robots. Springer

Dunne A, Raby F (2001) Design noir: the secret life of electronic objects. Springer Science & Business Media

Edwards A (2018) Animals, humans, and machines: interactive implications of ontological classification. In: Guzman AL (ed) Human-Machine Communication: Rethinking Communication, Technology, and Ourselves. Peter Lang, New York, pp 29–50

Elish MC, Boyd D (2017) Situating methods in the magic of big data and artificial intelligence. Commun Monogr 85(1):57–80

Feyerabend P (1975) Against method: outline of an anarchistic theory of knowledge. Humanities Press

Graeber D (2005) Fetishism as social creativity: or, fetishes are gods in the process of construction. Anthropol Theory 5(4):407–438

Gunkel DJ (2012) The machine question: critical perspectives on AI, robots, and ethics. MIT Press

Gunkel DJ (2018a) Robot rights. MIT Press

Gunkel DJ (2018b) The relational turn: third wave HCI and phenomenology. In: Filimowicz M, Tzankova V (eds) New directions in third wave human-computer interaction volume 1—technologies. Springer International Publishing, pp 11–24. https://doi.org/10.1007/978-3-319-73356-2_2

Guzman AL (2018) What is human-machine communication, anyway. In: Guzman AL (ed) Human-machine communication: rethinking communication, technology, and ourselves. Peter Lang Publishing, pp 1–28

Guzman A (2020) Ontological boundaries between humans and computers and the implications for human-machine communication. Hum-Mach Commun 1:37–54. https://doi.org/10.30658/hmc.1.3

Hamilton SN (2009) Impersonations: troubling the person in law and culture. University of Toronto Press JSTOR. https://doi.org/10.3138/9781442697584

Harrison S, Sengers P, Tatar D (2007) The three paradigms of HCI. Proc CHI 2007:18

Hasslacher B, Tilden MW (1995) Living machines. Robot Auton Syst 15(1):143–169. https://doi.org/10.1016/0921-8890(95)00019-C

Heidegger M (1967) Being and time. Blackwell, Oxford

Heidegger M (1982) The basic problems of phenomenology. Indiana University Press

Holland O (2003) The first biologically inspired robots. Robotica 21(4):351–363. https://doi.org/10.1017/S0263574703004971

Horakova J, Kelemen J (2008) The robot story: why robots were born and how they grew up. The mechanical mind in history. The MIT Press, pp 283–307

Huizinga J (1955) Homo Ludens: a study of the play-element in culture. Beacon Press

Jones RA (2013) Relationalism through social robotics: relationalism through social robotics. J Theory Soc Behav 43(4):405–424. https://doi.org/10.1111/jtsb.12016

Kang M (2011) Sublime dreams of living machines: the automaton in the European imagination. Harvard University Press

Klein M (2019) Robotic affective abilities. Peer Rev J About 8(1):34–44. https://doi.org/10.7146/aprja.v8i1.115413

Latour B (1996) Aramis, or, the love of technology. Harvard University Press

Latour B (2004) How to talk about the body? The normative dimension of science studies. Body Soc 10(2–3):205–229. https://doi.org/10.1177/1357034X04042943

Levinas E (1981) Otherwise than being or beyond essence, vol 4121. Springer Science & Business Media

McCarthy J, Wright P (2004) Technology as experience. Interactions 11(5):42–43. https://doi.org/10.1145/1015530.1015549

McCarthy J, Wright P (2015) Taking [a]part: the politics and aesthetics of participation in experience-centered design, vol 52. MIT Press. http://choicereviews.org/review/10.5860/CHOICE.191211

Nass C, Steuer J, Tauber ER (1994) Computers are social actors. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 72–78

Neville-Neil GV (2017) The chess player who couldn't pass the salt. Commun ACM 60(4):24–25. https://doi.org/10.1145/3055277

Norman D (2002) Emotion & design: attractive things work better. Interactions 9(4):36–42. https://doi.org/10.1145/543434.543435

Ratcliffe M (2002) Heidegger's attunement and the neuropsychology of emotion. Phenomenol Cogn Sci 1:287–312

Shannon CE, Weaver W (1975) The mathematical theory of communication. University of Illinois Press

Tantoush TMK (2001) Technology and organization. International encyclopedia of the social & behavioral sciences. Elsevier, pp 15502–15506. https://doi.org/10.1016/B0-08-043076-7/04254-6

Thibodeau J, Yolgörmez C (2020) Open-source sentience: the proof is in the performance. In: Proceedings of the 2020 international symposium for electronic arts (ISEA), Montreal, Canada

Torrance S (2008) Ethics and consciousness in artificial agents. AI Soc 22(4):495–521. https://doi.org/10.1007/s00146-007-0091-8

Turkle S (2006) A nascent robotics culture: new complicities for companionship: technical report series. American Association for Artificial Intelligence AAAI, p 10

Wasser JD, Bresler L (1996) Working in the interpretive zone: conceptualizing collaboration in qualitative research teams. Educ Res 25(5):5–15

Yin RK (2003) Designing case studies. Qualitative research methods. SAGE, pp 359–386