Borrowing, Poking and Entangling. In Search of Shared Spaces Between Science and Technology Studies and Human-Robot Interaction

Anna Dobrosovestnova
Jesse de Pagter
Astrid Weiss
anna.dobrosovestova@tuwien.ac.at
jesse.de.pagter@tuwien.ac.at
astrid.weiss@tuwien.ac.at
TU Wien
Vienna, Austria

ABSTRACT

In this paper, we reflect on the disciplinary foundations and dominant practices in the field of Human-Robot Interaction (HRI) from the perspective of our own experience of working interdisciplinarily and drawing on colleagues' ongoing work that transcends disciplinary boundaries. As a part of this reflection, we explore possibilities for the field's theoretical and methodological expansion, which we contend is needed, given the rapid expansion of robotic technologies in the real world settings. We argue the field of science and technology studies (STS) can be a valuable collaborator and contributor in the process of negotiating disciplinary boundaries of HRI and advancing the field beyond common narratives of technological solutionism and determinism. We frame STS as a field with a strong tradition of studying social and political embeddedness of science and technology, and how these are co-constitutive and co-emergent. STS also investigates the roles and responsibility different actors share in this process. To further explore how the interfacing between STS and HRI can be enacted, we sketch out three modes of interdisciplinary collaboration we call i) Borrowing, ii) Poking and iii) Entangling. We argue that each of these modes comes with advantages, disadvantages and challenges. In the conclusion, we engage the notions of "thinking with care" and disciplinary reflexivity, as an invitation to fellow scholars to consider which disciplinary assumptions are brought to the table when enacting different modes of interfacing between HRI and STS, and how these are entangled with the goals and (desired) outcomes of research practices.

CCS CONCEPTS

• **Social and professional topics** → **Socio-technical systems**; • **Computing methodologies** → **Philosophical/theoretical foundations of artificial intelligence**.

This work is licensed under a Creative Commons Attribution International 4.0 License.

HRI '23 Companion, March 13–16, 2023, Stockholm, Sweden
© 2023 Copyright held by the owner/author(s).
ACM ISBN 978-1-4503-9970-8/23/03.
https://doi.org/10.1145/3568294.3580033

KEYWORDS

Science and technology studies, future of HRI, interdisciplinary collaborations, social embeddedness of HRI

ACM Reference Format:
Anna Dobrosovestnova, Jesse de Pagter, and Astrid Weiss. 2023. Borrowing, Poking and Entangling. In Search of Shared Spaces Between Science and Technology Studies and Human-Robot Interaction. In *Companion of the 2023 ACM/IEEE International Conference on Human-Robot Interaction (HRI '23 Companion), March 13–16, 2023, Stockholm, Sweden.* ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3568294.3580033

1 INTRODUCTION

This paper takes the form of a critical essay in which we reflect on how "matters of concern" [40] and "matters of care" [18] are framed within the pluridisciplinary field of Human-Robot Interaction (HRI); in it, we consider possibilities for the theoretical and methodological expansion of HRI. To accomplish this, we turn to the field of Science and Technology Studies (STS) as a valuable collaborator in the process of renegotiating the disciplinary and onto-epistemological boundaries of HRI.

Our motivation for inviting further discussion of the possibilities of expanding theoretical and methodological boundaries of HRI, and to explore what role(s) STS could play in this process, stems from the personal experience of working across disciplines, and from our search for (multi-)disciplinary identity. Despite the differences in the authors' disciplinary backgrounds and individual research paths – the first author has a background in culture studies and cognitive science, and is now formally affiliated with HRI, the second author has a background in history of technology, philosophy of technology and STS, and the third author is a sociologist by training and has been contributing to the HRI community for over 15 years – we all share a certain discomfort when identifying with any given research field. While this paper targets a HRI audience, we feel both at home, but also somewhat alien to this community, as we feel about the STS community likewise. This occasionally ambivalent experience of disciplinary in-betweenness is, to no small extent, caused by our perception of both implicit and explicit incompatibilities between the types of research that HRI and STS communities encourage and produce, and the goals the fields pursue. All three authors have experienced tensions (though differently manifested) that stem from holding both, in the words

of Ruth Müller "interest in the natural or technical sciences" and the "desire for critical social science inquiry" [49, p.87]. Bearing in mind our somewhat ambiguous positionality, the observations and suggestions we formulate in this paper stem both from inside and outside perspectives on each of the fields.

From the perspective of HRI as a field, rather than from our personal experiences, we argue the need for further reflection on the present and future of HRI is motivated by the rapid expansion of technologies into real-world environments coupled with a growing awareness – also within the field – concerning the dynamic and complex nature of the socio-technical entanglements, and the ethical and political dimensions involved in HRI. Evolving policies and regulations with regard to AI and robotic technologies also contribute to the shift towards increased recognition of the social embeddedness of technologies in general. Developments in feminist technoscience and posthumanism that informed the third and, some even argue, the fourth wave in the field of Human-Computer Interaction (HCI) [4, e.g.], also offer new discourses and theoretical frameworks that push HRI towards rethinking relations between human, nonhuman and technological actors. Pragmatically speaking, STS – itself an interdisciplinary field in a very particular way [36] – has concerned itself with the social and political embeddedness of science and technology for several decades, as well as with questions about the impact of technologies and practices of how to reflect about the responsibility of different actors for how technology and society co-constitute each other. Thus, our key premise here is that STS can offer HRI a set of powerful conceptual tools and methodological approaches that would be of value for HRI-scholars engaged in the practice of developing and evaluating robotic technologies outside the widespread narratives of technological solutionism and determinism.

To facilitate the conciliation of HRI and STS, we ask: what does it mean for STS and HRI to intersect or collaborate? How can an interfacing between these very dynamic fields be enacted? And *where*? In other words, what could be the place(s) for STS in HRI? Here, we understand *place* as a metaphor to capture the idea of the possible roles STS can play for/within HRI; that is, we are concerned with the question how STS can be "translated, applied, or otherwise made useful" [48, p.101] for accommodation within HRI [1]? Place, in this context, can also mean the actual site(s), such as academic journals and conferences, where STS and HRI intersect. We are aware that asking these questions also means that we must further ask what this interfacing brings to STS, or how STS is transformed in the process of being brought into relation with HRI. Though we touch upon the idea occasionally throughout the paper, an extended discussion, though important, remains outside the scope of current work, in which we focus primarily on the HRI (perspective).

To address how STS can be made useful to HRI, we sketch out three modes of possible interdisciplinary interfacing. We refer to these modes as: i) *Borrowing*, ii) *Poking* and iii) *Entangling*. We differentiate the three modes based on i) whether they assume that some form of disciplinary boundaries exist at all, ii) the agency that the mode affords each field in the encounter, and iii) the extent of the modes' commitment to the renegotiation of the conceptual and methodological assumptions that it encounters. We do not recommend treating these modes as a complete representation of the interdisciplinary reality "out there". The reality, by necessity, is likely to be much more messy [41] and complex. Despite the authors' shared bias towards Entangling as our preferred mode of interdisciplinary interfacing, for reasons we will address below, we do not wish to suggest that one of the modes is *inherently* better than others at accomplishing this interfacing; what constitutes 'better' depends too much on the particular goals and perspective of each individual and project. That said, we hope that framing the discussion on interdisciplinarity through the lens of the three modes we have proposed is a helpful analytical exercise to advance discussions about the goals of HRI, its dominant methodological approaches, and where they may be lacking responsibility and disciplinary reflexivity.

In Sect. 2, we set the context for the paper by sketching out the trajectory HRI has taken as a pluridisciplinary field, and we present a more detailed argument for why the current discussion of theoretical and methodological diversification is a timely one. In Sect. 3, we proceed to introduce STS, some of its main lines of research, and the components of its methodological toolkit that we believe can enrich HRI in the process of its theoretical and methodological expansion. In Sect. 4, we present the three modes of disciplinary interfacing, give some examples of ongoing work to substantiate these, and point to some challenges and opportunities that come with each of them.

2 HRI. EMERGING PERSPECTIVES

2.1 Multidisciplinary Field of Human-Robot Interaction

HRI has its origins in the study of Human-Machine Interaction (HMI) or human factors. The initial goal of HMI research was to optimize machines for use by humans. The way for HRI was paved by the development of machines from passive devices dependent on user input to autonomous and pro-active systems. In the industrial context, however, roboticists historically considered only experts, such as robot programmers and trained maintenance personnel, as a target group for their systems. Their objective was to design efficient automated machines. The vision of co-existing intelligent service robots [55] surrounded by people in unstructured environments, as well as the narrative of the multi-functional humanoid social robot, broadened the envisioned target audience to include inexperienced users.

A constitutive effect of this development was that the main objective shifted towards the so-called 'intuitive' and 'natural' interaction, a notion that fellow researchers have since called into question [16]. To achieve 'intuitive', 'natural' interactions, the first interdisciplinary research projects on human-robot interaction were set up in the early 2000s, comprising representatives of various relevant disciplines such as AI, robotics, computer science, sociology, developmental psychology, art/design, and Human-Computer Interaction (HCI). This appeared to be a special endeavour, and Kiesler and Hinds noted in their *Introduction to the Special Issue of Human-Robot Interaction* in the Human-Computer Interaction journal that "people seem to perceive autonomous robots differently than they do most other computer technologies" [38, p. 3]. Similarly,

[1] Under the caveat that, in the process, HRI may need to redefine what it itself considers 'useful'

Dautenhahn argued that "human-robot interaction is very different from human-human interaction, human-computer interaction, and human-animal interaction" [17, p.2].

What was then primarily an assumption, has since been substantiated with empirical evidence; as it turns out, studying how people interact with robots raises a variety of unique wicked problems, i.e., problems that resist complete definition and resolution [56]. This circumstance requires that we acknowledge that every prototypical robot design is inextricably related to its problem space, the technological readiness level, its design, and to people, including the researchers involved in the development. In other words, problems and solution are necessarily highly interdependent in HRI, subjective, and fluid. In their work, Kiesler and Hinds pointed out that "designing these robots appropriately will require a deep understanding of the context of use and of the ethical and social considerations surrounding this context" [39, p. 4]. Similarly, Dautenhahn argued that the "exploring of the design space is likely never to be completed" [15, p. 21]. In other words, building a thorough understanding of how different design choices affect the complex variety of human responses towards robots, cognitive, affective, emotional, and relational, is a broader challenge than initially expected by the research community [66].

Despite the diverse disciplinary backgrounds of the founding figures of the field and the subsequent acceptance of the complexity of HRI, quantitative psychology still takes the lead in defining the standards of quality and acceptability for HRI publications, based on its own well-defined experimental designs, limited stimuli and preferred statistical methods [17]. The reasons for this can be traced, among other factors, to the fact that most seminal HRI studies, as well as much of today's work, are conducted with robotic prototypes in controlled laboratory settings according to standards and practices inherited from quantitative psychology. As such, the activities of HRI researchers are often well integrated into the work of robotic engineers, not in the least because they often share work spaces. In this environment both robotic engineering and most HRI research conforms neatly to the spirit of what Theodore M. Porter famously described as a "trust in numbers" [52]. This dominating positivist orientation has contributed decisively to the current situation in which (social) robotics and HRI researchers focus on what robots can do [58], while robo-ethicists, philosophers of technology and sociologists deliberate post factum what robots and, for that matter, their developers and researchers, *should do*.

2.2 Emerging Perspectives in HRI

Aware of the limitations that the dominant positivist orientation comes with, nonetheless we accept there can be some merit in attempts to translate questions related to the social embeddedness of robotic technologies into positivist, empirical research designs. HRI research on topics like transparency, explainability and trustworthiness of robots offers some good examples where fields like ethics, policy-making and HRI have collaborated successfully to develop empirically informed notions that have aided the process of ethical standardization in turn [3, 29, 69, e.g.]. When it comes to successfully integrating robots into society, however "success" may be defined, HRI is certainly a field that can add substantial empirical evidence to inform decisions about what meaningful strategies there are to accomplish this. We already see HRI, also in its current form, contributing to the development of regulations and laws that accompany this process.

Despite the demonstrated value of positivist research in HRI, we join other scholars in calling for a more critical "re-configuration" of HRI [7, 66, e.g.]. These voices primarily advocate further theoretical and methodological diversification to accommodate the increasing complexity of the social embeddedness of robots at all stages of their development and implementation [58]; their criticisms are coupled with an increasing awareness among HRI researchers of the blurred boundaries between HRI design and the social context of its use [7], as well as a growing recognition of the diversity of users, and the ways they engage with technologies beyond the interaction scripts as construed by developers [34, 44]. In line with these shifts, more inquiries into the social impact of robotic technologies are taking place alongside common areas of research in HRI such as usability or acceptance, as robots move from laboratories to public and private spaces [71].

Being sensitive to these developments, HRI conferences are gradually opening up to workshops and other types of sessions that engage in discussions of how the technologies humans build reshape society in general and specific practices of care, or service labor, for example. Journals have also proposed new Special Issues that are open to a wider range of authors from other disciplines. Theoretical and methodological developments within HRI are also *already* taking place to address these paradigmatic shifts. Prominent examples of new approaches that have already established themselves in HRI include value-based approaches, such as Value Sensitive Design (VSD), which are strongly focused on the way values can be implemented in the design process [13, 63]. Similarly, participatory approaches in HRI explore a way of integrating perspectives and values of non-experts [45, 68]. However, these approaches still exist at the margins of HRI; they also incur their own limitations [30]. This means the true scope of the complexity of situated interactions in the real world contexts still leaves much to be explored.

2.2.1 Further Arguments for Theoretical and Methodological Expansion. A number of other arguments reinforce the need for further theoretical and methodological expansion, and continue to strengthen the case that it can be enacted productively in collaboration with scholars from social sciences and humanities.

First, critical points have already emerged in the broader context of HCI, where researchers have advocated for approaches that explicitly engage a much more outspoken and critical stance vis-à-vis the sociopolitical impact of technology research and design. This explicitly ethical and critical stance has also informed the paradigmatic shifts in HCI that are commonly described as different waves in HCI. For instance, Blevis et al. argue for what they call the 'fourth wave' in interaction design where the primary focus ought to be on values, ethics and politics [8]. In this case, the goal is not just to *recognize* how potential issues can emerge or need to be addressed during the design phase. Rather, the aim of these approaches is to go a step further and to view the development of interactive technologies as a way to proactively embody and represent values and ethical and political stance in design of artifacts and interactions.

Second, increased attention to the perspectives that emphasize voices of marginalized communities also challenge the established conceptual and methodological practices within HRI. Insights from gender studies and efforts toward decolonialization of technology [1, 43, e.g.] contribute to the public and academic discourses that shape the trajectory of HRI as a field by nudging towards further renegotiation of the established narratives and approaches. Again, substantial work on integrating these perspectives with the topics concerning technology design has already been undertaken within HCI community [5, 37, e.g.] and can be helpful in paving path for similar efforts within HRI.

Third, efforts in governance and policy-making are both becoming increasingly concerned with the impact and regulation of robots in new social contexts. New laws and regulatory frameworks are being developed as governments are becoming increasingly focused on the potential systemic risks that the real-world implementation of types of (autonomous) robots might incur [28]. At the same time, worries over the effect of interactive robots and concerns about their impact are also prominent in public discussions. The heightened attention to potentially harmful consequences of robotic technologies adds urgency to calls for increased collaborations with trained ethicists and social sciences scholars in order to address these emerging concerns.

3 SCIENCE AND TECHNOLOGY STUDIES: A PERFECT FIT TO HRI?

STS is an interdisciplinary field that emerged in the 1970's when the socially embedded character of science and technology became an increasingly important topic of research and theory [59]. Early STS research maintained a strong focus on demonstrating how the production of science and technology are socially embedded; that is, how scientific institutions and the practice of scientific knowledge production, as well as technology design, are entangled with society, people's lives, situated practices, values and politics. From the perspective of early STS-scholars, technology emerges from and is constitutive of the social worlds [62]. Taken broadly, "STS views science and technology as historical products of human labor, investments, choices, and designs" [25]. In the process of doing science and making technologies, people also negotiate their practices, identities, societies, their bodies and material surroundings (ibid.). At a risk of over-simplification, today contemporary STS can characterized by "its engagement with various publics and decision makers, its influence on intellectual directions in cognate fields, its ambivalence about conceptual categories and dichotomies, and its attention to places, practices, and things" [54, p.1].

From a methodological perspective, both seminal STS research as well as current research rely heavily on ethnographic approaches. These were initially brought into the field from sociology and anthropology to study the mundane practices and relational and technological arrangements of the day-to-day work of scientists and engineers [42]. At their methodological core, these methodologies are still 'thick' qualitative descriptions. Even though they have since developed differently due to their different enactment and different political agendas pursued in various STS branches and spin-offs [31, e.g. postcolonial and feminist STS]), in all of these forms, they have remained sensitive to the plurality of human and non-human actors performing socio-technical entanglement. Not only do such studies provide rich insights into gathering empirical data, they also foster interdisciplinary collaboration with scientists, engineers and technicians, since they require a close engagement with the practices, routines and dynamics in science and technology [2, 60].

Because of their direct and critical engagement with the practices of knowledge generation, insights from STS are useful when addressing the ambivalence and complexity of the practices of robots design, on the one hand, and the social contexts where such robots are situated on the other. For instance, STS-based approaches can be engaged when it comes to examining, mapping and subverting stereotypes and biases that researchers and developers have with respect to the (potential) 'users' of technology, or even to problematize the very notion of the idealized 'user' [6, 62, e.g.]. This allows a more nuanced understanding of the diversity of people and their ways of engaging with technology. STS can also offer theoretical and methodological insights with respect to how (dominant) social codes, like problematic gender and racial narratives, shape the appearances and behaviors of robots in specific areas of application, such as healthcare or service industries [9, 24, 53].

Because of its focus on scrutinizing relations between science, technology, society and policy-making, STS also has a strong track record on issues like transparency and the accessibility of science and technology development. As a field, STS has a long history of engaging with the ways technologies and techno-scientific development are perceived by the wider public, and how this perception changes under different circumstances [35]. Since its inception, research in STS has also occupied itself with the public uptake of emerging technologies such as robots [32], and with how such topics are communicated about by different actors. This makes STS an ideal ally for understanding what happens when robots enter real-world contexts, where their interpretations and acceptance or rejection by the public is shaped by diverse expressions and representations of hopes, fears and speculations about the (near) future.

At its core, STS has been concerned with theorizing socio-technical change on the basis of empirical social research. This research has been conducted with a range of methods and skills that are focused on continuously reframing the understanding of technological artifacts as a part of their socio-technical context. If these skills can be brought to bear in a fertile collaborative context, STS holds considerable promises for the way issues related to the socio-political dimensions of robotic technologies can become actionable topics for HRI research beyond the dyadic paradigm [34]. In particular, STS' focus on the entanglement of technology and society entails tangible efforts towards defining and studying robots as socio-technical systems. By establishing solid definitions of robotic artifacts as socio-technical systems, robots can also be studied as such. Socio-technical systems approaches are generally recognized as useful perspectives for understanding the complexities of technologies in their social context, since they emphasize the intrinsic entanglement of the technical, social and institutional dimensions that lead to the specific ways of using technological artifacts [64].

4 MODES OF INTERFACING BETWEEN STS AND HRI

But what does it mean for STS and HRI to come together? How can such coming together be performed? Drawing on the literature on interdisciplinary collaborations and selected examples from HRI research that already enacts some forms of bridging between these fields, we outline three modes of interfacing between disciplines, which we refer to as *Borrowing*, *Poking* and *Entangling*. We differentiate these modes based on i) the assumptions they make about the idea of disciplinary boundaries, disciplinary agency; ii) the modes' commitment to re-configuring theoretical and methodological assumptions; and iii) the disciplinary goals each mode requires. Following Sheila Jasanoff, this means making explicit the trained and often taken for granted forms of reasoning and research practices that come with being situated within a discipline [36].

4.1 Borrowing

As Borrowing we refer to a mode of interfacing wherein scholars situated within HRI introduce selected concepts from the theoretical packages of STS (or another field) into what one would otherwise consider a standard HRI study set-up. [2] Borrowing can also be the use of STS methodologies (e.g., ethnography-inspired approaches) in a slightly adjusted or simplified form in pursuit of the standard HRI goals, like improving the capacities of a robot or people's 'user experiences'. In such cases, the qualitative strengths and focus on 'thick' data, which are characteristic of STS, offer a basis for exploratory studies that otherwise do not adhere to the critical, deconstructive analytical stance that is fundamental to STS research. In this case, the researchers in the Borrowing mode implement conceptual and methodological vignettes from STS, while at the same time preserving a conventional focus on the development of robots that are more efficient, likeable or easier for people to accept [11, 47].

A study by Fortunati and colleagues [27] exemplifies Borrowing. In this case, the theoretical framework of social representations and a selective deployment of the concept of "imaginaries" is used to empirically investigate how children and pre-teens develop their lay-understanding of robots. According to the authors, these understandings are important to address, because they will shape individuals' future attitudes towards robots and the practices of use. In a somewhat similar study, [Weiss et al.] draw on studies in STS that exemplify how expectations about robots are performative and consequential for how the robots are perceived. Bearing on this theoretical core, the authors carry out an exploratory investigation with 52 user study participants, who interacted with a humanoid robot for the first time, and six experts from different industries, with the goal of comparing the way they understand and talk about future human-robot relationships.

The particular downside to Borrowing as a mode of disciplinary interfacing is that in many cases it is likely to result in what Lynch refers to as a "surfeit model", where the agency shifts from STS to HRI [48]. In practice, this means that the original complexity – and sometimes even the very meaning of a concept – gets either stripped of its nuance or lost altogether. How *constructivism* has been interpreted in many studies in educational robotics offers a salient example. Setting aside the disputes regarding the definitions of the various forms of constructivism in STS and philosophy of science more broadly [51, c.f.], the noteworthy point here is that – regardless of the relevant definition of constructivism – more often that not, empirical work in educational robotics is conducted according to a straightforward behaviorist paradigm that translates constructivist principles into study designs with which it is fundamentally incompatible [20, e.g.]. Concepts like culture and gender, widely theorized in STS, have been 'lost in translation' in a similar way. In much current HRI research, culture is typically operationalized as (hegemonic) national culture, and then further reduced to just another input within the standard positivist epistemological framework. Not only does this approach disregard many alternative understandings of what people consider constitutive to their own culture, it also neglects other, later, and more widely accepted conceptualizations of culture available among experts on the subject in STS and culture studies. These include, for example, more local notions of cultures, or culture as a form of epistemic practices and ways of knowing that shape a community [12]. With regard to gender, excellent initiatives within HRI have begun to address the notion more critically, specifically how gender biases are operationalized in robot designs or inscribed by people based on specific design features and dominant cultural narratives [50, e.g.]. That said, many of the discourses and research efforts on the subject of gender and robots still operate with a cis-binary understanding of gender and ableist paradigm of embodiment [30].

Despite these limitations, Borrowing, as a somewhat casual form of interdisciplinarity, can generate fruitful encounters nevertheless. The main promise of borrowing concepts and methodologies is that it helps to familiarize HRI researchers with perspectives and concepts that have the potential to become useful in establishing new kinds of expertise about the socio-cultural embeddedness of robots and robots' design in future. From the perspective of STS, being borrowed from can prepare the ground for a more extended access to the HRI community for fieldwork through which STS scholars can study the processes of robot development and integration. As such, borrowing sets the stage for interdisciplinary collaboration wherein HRI can function both as a research site where HRI and STS can come together, and as an object of study for STS.

4.2 Poking

Poking refers to the situation where scholars who are either situated within STS, or scholars who explicitly adopt an STS perspective, rely on conceptual and methodological toolkits from STS to analyze a phenomenon or practice in HRI. More often than not, these analyses are deconstructive in their nature: that is, the goal that is being pursued need not have anything to do with an intent to design or improve a specific technology. Rather, it is to unpack or unveil, or critically reflect on the chosen practices and/or assumptions in HRI, or robotics more broadly. In many aspects, poking is similar to the notion of *critique* as described by [Fitzgerald and Callard]. When discussing common ways how social sciences often position themselves vis-à-vis cognitive neuroscience, they mentioned critique as

[2] Given how heterogeneous and dynamic the HRI is, it is not at all trivial to define what being situated within the field means. Here, we simply mean that scholars either self-identify with this field and/or publish in established HRI venues, such as HRI Conference Proceedings, Ro-Man, Journal of Social Robotics, Transactions in Human-Robot Interactions and others.

a mode that uses tools of "historical, social and cultural analysis as external methods to either: (1) uncover unconscious or hidden biases within the new brain sciences, and to locate nefarious social, political, economic and epistemic agendas within them [...]; or (2) deflate particular neuroscientific trends or claims that have found favour within the humanities or social sciences" [26, p.9].

Importantly, Poking operates on a strong assumption about disciplinary boundaries. In this mode, these boundaries are maintained both through the separation of and respect for disciplinary goals and methodologies, but also through an intentional separation in terms of the sites where such works are published and which audiences they primarily target. It has been our observation and experience of publishing our own work that studies that we would classify as 'Poking' mode, are more commonly published outside of specialised HRI venues. Consistent with their roots in STS, such studies do not (directly) aim 'to build better technologies'; editor and reviewers comments to 'Poking' studies indicate that, from an HRI perspective, such work is too general, or too critical, for HRI scholars to translate it into actionable solutions. If submitted to HRI venues, such papers are met with curiosity, at best, but also a degree of confusion regarding their practical value for HRI. In the worst case, the research is considered completely irrelevant, as if it were merely "adding unnecessary and unwelcome complexity" [65]. In such cases, authors are curtly recommended to submit elsewhere. That means it remains rather unclear what effect, if any, Poking has on HRI practices.

Burema published an example of this kind of work in the AI & Society Journal [10]. The target article builds on the premise that there is a need to critically address the bias in how older adults are represented in HRI. The author engages the social construction of technology framework from an STS perspective, coupled with a qualitative content analysis of 96 HRI publications. Together these components make the case that biases have the potential to reinforce problematic and unreflected narratives about older adults as "burdensome", "weak" and in need of technology-mediated "fixing". In turn, such representations of older adults reinforce ageist and neoliberal narratives about aging and the elderly. The first author's work, also published in AI & Society Journal, is another example of Poking. Specifically, that analysis drew on the literature in sociology of work, feminist technoscience and STS to critically address the questions concerning affective labor in service industries in which robots and humans are envisioned to share work tasks [21].

Another example of Poking from the first author's work led to the separation of the same fieldwork into two strands at two different venues – one for HRI and another for a more philosophy of technology and STS-oriented audience. With collaborators, the first author drew on an ethnographic stance to study whether, when and how people assist commercial delivery robots on the streets of a city [23]. A separate publication, and third constellation of co-authors, presented reflections on the meaning of such help, also from the perspective of the labor involved [22]. Separating these dimensions of research into different publications and presenting them at different venues for different audiences, preserved an implicit disciplinary divide.

The apparent disadvantage of Poking is that it can be perceived by the HRI community as in-actionable, hostile or even arrogant. Since Poking is commonly situated in venues outside mainstream HRI, such studies risk remaining inconsequential (for HRI), because they never reach a (broad) HRI audience. As a methodology of interfacing, Poking is often deliberately based on an outsider's perspective, that rules out immediately integrating the new insights into the research practice facing critique. As such, Poking is not so much an interdisciplinary integration, as a confrontation between disciplines.

Poking can nevertheless stimulate debates on the different futures of robots in society; for example, it can lead to debates about investments to establish increasing overlaps between audiences for scholarly work in HRI and STS. Practical examples of such overlaps can be found in interdisciplinary work where researchers are actively participating in both HRI and STS communities, in conference sessions, Special Issues or similar shared venues. Moreover, Poking can contribute to the demystification of robotics and HRI, and cast a more critical eye at the hubris and buzzwords commonly found in the imagery of robotics as a field of technological development [61]. In the long run, these kinds of interactions and debates can help to draw a more balanced picture of robotics that is open to alternative interpretations and different imaginaries of human-robot futures. From the perspective of STS, these types of debates have a potential to establish an environment where STS scholars can get involved in a more constructive manner in robotics development, should they want to, while promoting design practices that engage with the critical values that STS- perspectives popularise.

4.3 Entangling

Entangling, as we conceptualize it here, assumes a shared space where HRI and STS participate in the ongoing renegotiation of (disciplinary) boundaries, as well as the boundaries of what constitutes an object of a study. In contrast to Borrowing and Poking, which operate based on the assumption of (more or less) fixed boundaries, Entangling moves away from the premise of a clear-cut disciplinary separation; instead the Entangling mode suggest that if there were a boundary, it would be rather porous and fluid. In that sense Entangling challenges the notion of disciplinary division itself, and in-so-doing captures the "counter inter- regime" inclination discussed by [Fitzgerald and Callard] in the article on the dynamics between fields.

Concerning theories and methodologies, Entangling means that the assumptions that scholars bring to the table will be put to the test, reconfigured or even set aside entirely. Although HRI scholars may continue to focus on designing robots, becoming entangled with scholars from social sciences means that these robots, and the purpose they are to fulfil, are envisioned in radically different ways, and with a deepened sensitivity towards how they are constitutive of new socio-technical arrangements. These new concerns may include explicit considerations for how robots can also function as sites of intervention and subversion of established dominant orders [57]. Entangling also means that the separation between the 'social' and the 'technical' in how robots are defined is no longer feasible. Instead, robots are operationalized as socio-technical systems throughout all stages of their development and implementation; this perspective also considers how robots reshape the processes used to design and implement them over long(er) periods of time. In the Entangling mode, the way research problems and goals are

framed is renegotiated with due consideration for the political and historical contexts that shape the very situation a robot is supposed to enter;discursive, historical and political dimensions behind the research and development of new types of robots are explicitly taken into account. For example, when working on a project where the goal is to design a robot for the service sector, enacting entangling means, among other things, explicitly considering the topics of gendered, racialized and often 'invisible' labor that service industries rely upon today.

Entangling most strongly raises issues of the responsibility that designers and researchers share for contributing their part to either reinforcing, or challenging established practices and dominant discourses about robots in and for society. In this regard, Entangling means abiding by the premise, widely shared within STS and philosophy of technology communities, that technologies are not *neutral* [70, e.g.]. This also means that the role of ethical and political dimensions in developing robotic technologies must be explicitly addressed throughout *all stages* of robot development: from conceptualization to design and implementation in the context of deployment. A substantial effort in this mode has been carried out by the philosopher Johanna Seibt and colleagues; with The 5 Principles of Integrative Social Robotics, they seek out to propose an approach for generating social robotics solutions that are culturally sustainable [58].

[Hornecker et al.] perform an Entangling effort in their engagement of feminist STS thinking to explore, through fieldwork in care facilities, how the functioning of assistive technology, even in the case of seemingly straightforward 'mechanistic' tasks such as lifting, is deeply intertwined with emotional labor. Grounded in the STS perspectives on care, on the one hand, and on the empirical insights from fieldwork on the other hand, the authors derive a triadic interaction framework for design of future robotic technologies. Within the proposed framework, the notion of interaction is shifted from the dyadic paradigm of HRI to encompass residents (of care facilities), caregivers and technology itself [33]. [Lee et al.] is another example of entangled work, in which the scholars engaged the notion of robots as 'boundary objects' and applied Artifact Analysis to explore, with expert and non-expert participants alike, the different aspects of technology that come to fore for them. According to the authors, becoming aware of these differences in how robots are framed by different people allows the designers to situate themselves on the boundary between researchers and users. In a similar vein, [Šabanović and Chang] argued that robot sociality is an emergent relational property of interactions, and illustrated this premise with an empirical, multi-sited case study of the commercial care robot PARO used in eldercare. Drawing on the analysis of a series of observational studies at different sites that included the laboratory and the care facility, Sabanovic and colleagues brought forth how the robot is constructed through sense making and practices of different actors in different institutional contexts (ibid).

In our view, this kind of work sets an excellent example of how HRI related research can go beyond the dyadic interaction paradigm to embed the social context as a crucial component of successful robot design and implementation. Such work also calls for new types of reflexive practices that incorporate considerations for how research sites are constructed , how methods are performed, and how human-robot interactions – and different dimensions pertaining to these (e.g., robot sociality) – are constructed and socially and culturally, and how they are politically situated. Ethnographic technology studies have dedicated extensive attention to such reflexivity, articulating the role of the researcher in crafting the research site, for example, and examining how the researchers' presence (re)shapes it [14].

To conclude, as a mode of interfacing between HRI and STS, Entangling means engaging: i) a different understanding of what a robot is beyond a mere engineering artifact, ii) frequently, mixed methods approach and/or expansion of the standard HRI toolbox to include methods from social sciences, and iii) extension of the research questions and goals beyond the standard focus on usability and tasks performance to include socio-relational dimensions of interactions. Another characteristic feature of Entangling is how it extends the pool of participants' relevant for HRI studies beyond the so called primary intended users. Entangled work considers how other human and nonhuman actors are involved in shaping and sustaining interactions and *relations* in particular enactments of socio-technical networks. Entangling, taken together, though maybe a rather laborious mode of collaboration, promises deeper insights and solutions that are more socially and ethically sound. Importantly, such solutions explicitly address the issues of power hierarchies and topics like access, i.e. who is included and who is excluded when a particular robot enters the social arena? To us the questions of power and access in HRI also call forth more explicit considerations for who benefits from robotic technologies, and who gets to decide not only which technologies are introduced into real-world settings, but also which technologies are developed and purchased in academic settings. We recognize how such questions are most likely to be coupled with conflicting roles and affiliations. While it is not our goal to equate research and activism, we nevertheless believe that it is important not to turn a blind eye to such questions, no matter how controversial they may be. One way to tackle them as a community would be, for example, to discuss potential 'exit strategies' for researchers in the situations where conflict of values between different stakeholders, including researchers and developers, is impossible to reconcile without substantially compromising scientific and ethical rigor.

5 CONCLUSIONS

In this paper, we discussed what role(s) STS, as a field that has been long concerned both with examining how science, technology and society co-constitute each other, as well as how scientific practices are also socially and politically embedded, can play in/for HRI in response to the increasing need for conceptual and methodological expansion of our field and 're-configuring' of its goals and affinities. We framed our discussion of why STS can be a good companion on this journey by sketching out three modes of interfacing between these fields, which we called Borrowing, Poking and Entangling. By drawing on those examples of existing studies in HRI that already bridge HRI and STS in some way, we exemplified how each of these modes of enacting interdisciplinarity comes with its own advantages and disadvantages, as well as particular underlying assumptions about disciplinary boundaries. As we mentioned in the Sect. 1, while we do hold a bias towards entangling because it

seems like the most enriching mode to us – though also the most challenging and resource-intensive – we think there is merit in all three modes of interfacing. What one must consider is how the (explicitly or implicitly) chosen mode relates to the research goals at hand, what the mode brings in, and also what is left out in the process of translation from one discipline to another.

By inviting fellow HRI researchers to reflect on their disciplinary and interdisciplinary practices, our intention is not to stir confrontation. Rather, it is to delver deeper into what María Puig de la Bellacasa coined as "thinking with care", when she wrote: "thinking in the world involves acknowledging our own involvements in perpetuating dominant values, rather than retreating into the secure position of an enlightened outsider who knows better." [19, p.197]. Such non-idealized thinking with care, both as an epistemic, but also as a relational practice, is not only an affective or ethical stance. It involves the ongoing labor of reflecting on the assumptions that are ingrained in our knowledge-production practices, as well as the labor of keeping an open mind to the perspective of others, and the willingness to change in relation to them and the world that our epistemic practices co-constitute.

ACKNOWLEDGMENTS

The first author is a recipient of a DOC Fellowship of the Austrian Academy of Sciences. This work was also partly funded by a FWF ConnectingMinds grant to the project Caring Robots // Robotic Care (CM 100-N). The first author would also like to thank Jane Calvert, Doris Allhutter, Edward J. Hackett, Fredy Mora Gamez, and her peers at the Department of Science and Technology Studies, University of Vienna, for all the conversations that shaped her in ways that go far beyond the modes of disciplinary interfacing described in this paper. We also thank the anonymous reviewers for their thoughtful comments to the first draft of this paper.

REFERENCES

[1] Boris Abramovic, Grisha Coleman, Marco Donnarumma, Elizabeth Jochum, and Christina Schoux Casey. 2021. Decolonizing the Machine: Race, Gender and Disability in Robots and Algorithmic Art. *Proceedings of Polititcs of the machines-Rogue Research 2021 3* (2021), 3–13.

[2] Madeleine Akrich. 1992. The De-Scription of Technical Objects. In *Shaping Technology/Building Society : Studies in Sociotechnical Change*, Wiebe Bijker and John Law (Eds.). MIT Press, Cambridge, MA, 205–224.

[3] Thomas Arnold and Matthias Scheutz. 2018. The "big red button" is too late: an alternative model for the ethical evaluation of AI systems. *Ethics and Information Technology* 20, 1 (2018), 59–69.

[4] Simone Ashby, Julian Hanna, Sónia Matos, Callum Nash, and Alexis Faria. 2019. Fourth-Wave HCI Meets the 21st Century Manifesto. In *Proceedings of the Halfway to the Future Symposium 2019* (Nottingham, United Kingdom) *(HTTF 2019)*. Association for Computing Machinery, New York, NY, USA, Article 23, 11 pages. https://doi.org/10.1145/3363384.3363467

[5] Shaowen Bardzell. 2010. Feminist HCI: Taking Stock and Outlining an Agenda for Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 1301–1310. https://doi.org/10.1145/1753326.1753521

[6] Eric P. S. Baumer and Jed R. Brubaker. 2017. Post-userism. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 6291–6303. https://doi.org/10.1145/3025453.3025740

[7] Andreas Bischof, Eva Hornecker, Antonia Lina Krummheuer, and Matthias Rehm. 2022. Re-Configuring Human-Robot Interaction.. In *HRI*. 1234–1236.

[8] Eli Blevis, Kenny Chow, Ilpo Koskinen, Sharon Poggenpohl, and Christine Tsin. 2014. Billions of interaction designers. *Interactions* 21, 6 (2014), 34–41.

[9] Raquel Boso Perez. 2020. Atanasoski, N. and Vora, K. Surrogate Humanity: Race, Robots, and the Politics of Technological Futures, London and Durham: Duke University Press. 2019. 240pp $99.95 (cloth) $25.95 (pbk) $25.95 (ebk). ISBN 978-1-4780-0386-1. *Sociology of Health and Illness* 42, 6 (2020), 1489–1490. Publisher: Wiley.

[10] Dafna Burema. 2022. A critical analysis of the representations of older adults in the field of human–robot interaction. *AI & SOCIETY* 37, 2 (2022), 455–465.

[11] Praminda Caleb-Solly, Sanja Dogramadzi, David Ellender, Tina Fear, and Herjan van den Heuvel. 2014. A mixed-method approach to evoke creative and holistic thinking about robots in a home environment. In *Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction*. 374–381.

[12] Karin Knorr Cetina. 1999. *Epistemic Cultures: How the Sciences Make Knowledge*. Harvard University Press, Cambridge, MA.

[13] EunJeong Cheon and Norman Makoto Su. 2016. Integrating roboticist values into a Value Sensitive Design framework for humanoid robots. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 375–382.

[14] Bohkyung Chun. 2019. Doing Autoethnography of Social Robots: Ethnographic Reflexivity in HRI. *Paladyn, Journal of Behavioral Robotics* 10, 1 (Jan. 2019), 228–236. https://doi.org/10.1515/pjbr-2019-0019

[15] Kerstin Dautenhahn. 2004. Robots we like to live with?!-A developmental perspective on a personalized, life-long robot companion. In *RO-MAN 2004. 13th International Conference on Robot and Human Interactive Communication*. IEEE, 17–22.

[16] Kerstin Dautenhahn. 2013. Human-robot interaction. *The Encyclopedia of Human-Computer Interaction, 2nd ed.* (2013).

[17] Kerstin Dautenhahn. 2018. Some Brief Thoughts on the Past and Future of Human-Robot Interaction. *ACM Transactions on Human-Robot Interaction* 7, 1 (May 2018), 1–3. https://doi.org/10.1145/3209769

[18] Maria Puig de la Bellacasa. 2011. Matters of care in technoscience: Assembling neglected things. *Social Studies of Science* 41, 1 (Feb. 2011), 85–106. https://doi.org/10.1177/0306312710380301 Publisher: SAGE Publications Ltd.

[19] Maria Puig De La Bellacasa. 2012. 'Nothing comes without its world': thinking with care. *The Sociological Review* 60, 2 (2012), 197–216. Publisher: SAGE Publications Sage UK: London, England.

[20] Anna Dobrosovestnova. 2019. MASTERARBEIT/MASTER'S THESIS. (2019).

[21] Anna Dobrosovestnova, Glenda Hannibal, and Tim Reinboth. 2021. Service robots for affective labor: a sociology of labor perspective. *AI & SOCIETY* (April 2021). https://doi.org/10.1007/s00146-021-01208-x

[22] Anna Dobrosovestnova and Tim Reinboth. 2023. Helping-as-Work and Helping-as-Care: Mapping Ambiguities of Helping Commercial Delivery Robots. *Social Robots in Social Institutions* (2023), 239–248. https://doi.org/10.3233/FAIA220623 Publisher: IOS Press.

[23] Anna Dobrosovestnova, Isabel Schwaninger, and Astrid Weiss. 2022. With a Little Help of Humans. An Exploratory Study of Delivery Robots Stuck in Snow. In *2022 31st IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*. 1023–1029. https://doi.org/10.1109/RO-MAN53752.2022.9900588

[24] Dion N. Farquhar. 2022. Power and Domination in Emerging Technologies.

[25] Ulrike Felt, Rayvon Fouché, Clark A. Miller, and Laurel Smith-Doerr. 2016. *The handbook of science and technology studies*. Mit Press.

[26] Des Fitzgerald and Felicity Callard. 2015. Social science and neuroscience beyond interdisciplinarity: Experimental entanglements. *Theory, Culture & Society* 32, 1 (2015), 3–32.

[27] Leopoldina Fortunati, Anna Esposito, Mauro Sarrica, and Giovanni Ferrin. 2015. Children's knowledge and imaginary about robots. *International journal of social robotics* 7, 5 (2015), 685–695.

[28] Eduard Fosch Villaronga. 2019. "I Love You," Said the Robot: Boundaries of the Use of Emotions in Human-Robot Interactions. In *Emotional Design in Human-Robot Interaction*. Springer, 93–110.

[29] Eduard Fosch-Villaronga, Christoph Lutz, and Aurelia Tamò-Larrieux. 2020. Gathering expert opinions for social robots' ethical, legal, and societal concerns: Findings from four international workshops. *International Journal of Social Robotics* 12, 2 (2020), 441–458.

[30] Kathrin Gerling, Kay Kender, Katta Spiel, Saskia Van der Oord, Dieter Bayens, Arno Depoortere, and Maria Aufheimer. 2022. Reflections on Ableism in Participatory Technology Design. In *Mensch und Computer 2022 - Workshopband*, K. Marky, U. Grünefeld, and T. Kosch (Eds.). Gesellschaft für Informatik e.V., 3. https://doi.org/10.18420/muc2022-mci-ws02-224

[31] Sandra Harding. 2009. Postcolonial and feminist philosophies of science and technology: convergences and dissonances. *Postcolonial Studies* 12, 4 (Dec. 2009), 401–421. https://doi.org/10.1080/13688790903350658 Publisher: Routledge _eprint: https://doi.org/10.1080/13688790903350658.

[32] Stephen Hilgartner. 1990. The dominant view of popularization: Conceptual problems, political uses. *Social studies of science* 20, 3 (1990), 519–539.

[33] Eva Hornecker, Andreas Bischof, Philipp Graf, Lena Franzkowiak, and Norbert Krüger. 2020. The Interactive Enactment of Care Technologies and Its Implications for Human-Robot-Interaction in Care. In *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society* (Tallinn, Estonia) *(NordiCHI '20)*. Association for Computing Machinery, New York, NY, USA, Article 78, 11 pages. https://doi.org/10.1145/3419249.3420103

[34] Eva Hornecker, Antonia Krummheuer, Andreas Bischof, and Matthias Rehm. 2022. Beyond dyadic HRI: Building robots for society. *interactions* 29, 3 (2022), 48–53.

[35] Alan Irwin and Brian Wynne. 1996. Misunderstanding science?: the public reconstruction of science and technology. (1996).

[36] Sheila Jasanoff. 2017. A Field of Its Own. *The Oxford Handbook of Interdisciplinarity, Oxford* (2017), 173–187.

[37] Cayla Key, Cally Gatehouse, and Nick Taylor. 2022. Feminist Care in the Anthropocene: Packing and Unpacking Tensions in Posthumanist HCI. In *Designing Interactive Systems Conference* (Virtual Event, Australia) *(DIS '22)*. Association for Computing Machinery, New York, NY, USA, 677–692. https://doi.org/10.1145/3532106.3533540

[38] Sara Kiesler and Pamela Hinds. 2004. Introduction to This Special Issue on Human-Robot Interaction. *Human-Computer Interaction* 19, 1 (June 2004), 1–8. https://doi.org/10.1207/s15327051hci1901&2_1

[39] Sara Kiesler and Pamela Hinds. 2004. Introduction to this special issue on human-robot interaction. *Human-Computer Interaction* 19, 1-2 (2004), 1–8.

[40] Bruno Latour. 2004. Why Has Critique Run out of Steam? From Matters of Fact to Matters of Concern. *Critical Inquiry* 30, 2 (Jan. 2004), 225–248. https://doi.org/10.1086/421123 Publisher: The University of Chicago Press.

[41] John Law. 2004. *After method: Mess in social science research*. Routledge.

[42] John Law. 2016. 1 STS as Method. *The handbook of science and technology studies* (2016), 31.

[43] Shaimaa Lazem, Danilo Giglitto, Makuochi Samuel Nkwo, Hafeni Mthoko, Jessica Upani, and Anicia Peters. 2022. Challenges and paradoxes in decolonising HCI: A critical discussion. *Computer Supported Cooperative Work (CSCW)* 31, 2 (2022), 159–196.

[44] Hee Rin Lee, EunJeong Cheon, Chaeyun Lim, and Kerstin Fischer. 2022. Configuring Humans: What Roles Humans Play in HRI Research. In *Proceedings of the 2022 ACM/IEEE International Conference on Human-Robot Interaction (HRI '22)*. IEEE Press, Sapporo, Hokkaido, Japan, 478–492.

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

[46] Hee Rin Lee, Selma Šabanovic, and Erik Stolterman. 2014. Stay on the Boundary: Artifact Analysis Exploring Researcher and User Framing of Robot Design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1471–1474. https://doi.org/10.1145/2556288.2557395

[47] Deirdre E Logan, Cynthia Breazeal, Matthew S Goodwin, Sooyeon Jeong, Brianna O'Connell, Duncan Smith-Freedman, James Heathers, and Peter Weinstock. 2019. Social robots for hospitalized children. *Pediatrics* 144, 1 (2019).

[48] Michael Lynch. 2009. Science as a vacation: Deficits, surfeits, PUSS, and doing your own job. *Organization* 16, 1 (2009), 101–119.

[49] Ruth Müller. 2017. Crafting a career in STS: meaning making, assessment, and interdisciplinary engagement. *Engaging Science, Technology, and Society* 3 (2017), 84–91.

[50] Giulia Perugia, Stefano Guidi, Margherita Bicchi, and Oronzo Parlangeli. 2022. *The Shape of Our Bias: Perceived Age and Gender in the Humanoid Robots of the ABOT Database*. https://doi.org/10.1109/HRI53351.2022.9889366

[51] D. C. Phillips. 1995. The Good, the Bad, and the Ugly: The Many Faces of Constructivism. *Educational Researcher* 24, 7 (Oct. 1995), 5–12. https://doi.org/10.3102/0013189X024007005 Publisher: American Educational Research Association.

[52] Theodore M Porter. 1996. Trust in numbers. In *Trust in Numbers*. Princeton University Press.

[53] Matthias Rehm, Antonia L Krummheuer, Kasper Rodil, Mai Nguyen, and Bjørn Thorlacius. 2016. From Social Practices to Social Robots–User-Driven Robot Development in Elder Care. In *International Conference on Social Robotics*. Springer, 692–701.

[54] Jesse Richmond. 2008. Edward J. Hackett, Olga Amsterdamska, Michael Lynch and Judy Wajcman (eds.) The Handbook of Science and Technology Studies, Cambridge, MA and London: MIT Press, 2007. Pp. xiv+ 1065. ISBN 978-0-262-08364-5.£ 35.95 (hardback). *The British Journal for the History of Science* 41, 4 (2008), 628–629.

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

[56] Horst WJ Rittel and Melvin M. Webber. 1973. Dilemmas in a general theory of planning. *Policy sciences* 4, 2 (1973), 155–169. Publisher: Springer.

[57] Bojana Romic. 2021. Negotiating Anthropomorphism in the Ai-Da Robot. *International Journal of Social Robotics* (Jan. 2021).

[58] Johanna Seibt, Malene Damholdt, and Christina Vestergaard. 2018. *Five Principles of Integrative Social Robotics*. https://doi.org/10.3233/978-1-61499-931-7-28

[59] Sergio Sismondo. 2011. *An introduction to science and technology studies*. John Wiley & Sons.

[60] Lucy Suchman. 1993. Working relations of technology production and use. *Computer supported cooperative work* 2, 1 (1993), 21–39.

[61] Lucy Suchman. 2019. Demystifying the intelligent machine. In *Cyborg Futures*. Springer, 35–61.

[62] LK Van den Scott, Carrie B Sanders, and Anthony J Puddephatt. 2017. Reconceptualizing users through enriching ethnography. *The Handbook of Science and Technology Studies,* (2017), 501–527.

[63] Aimee Van Wynsberghe. 2020. Designing robots for care: Care centered value-sensitive design. In *Machine Ethics and Robot Ethics*. Routledge, 185–211.

[64] Pieter Vermaas, Peter Kroes, Ibo Van de Poel, Maarten Franssen, and Wybo Houkes. 2011. A philosophy of technology: from technical artefacts to sociotechnical systems. *Synthesis Lectures on Engineers, Technology, and Society* 6, 1 (2011), 1–134.

[65] Ana Viseu. 2015. Caring for nanotechnology? Being an integrated social scientist. *Social Studies of Science* 45, 5 (2015), 642–664.

[66] Astrid Weiss. 2022. *One cannot know everything — On the Need of Epistemological Diversity in Human-centered HRI Research*. Habilitation Treatise. Universität Salzburg.

[67] Astrid Weiss, Judith Igelsböck, Daniela Wurhofer, and Manfred Tscheligi. 2011. Looking forward to a "robotic society"? *International Journal of Social Robotics* 3, 2 (2011), 111–123.

[68] Astrid Weiss and Katta Spiel. 2022. Robots beyond Science Fiction: mutual learning in human–robot interaction on the way to participatory approaches. *AI & SOCIETY* 37, 2 (2022), 501–515.

[69] Alan FT Winfield, Anouk van Maris, Katie Winkle, Marina Jirotka, Pericle Salvini, Helena Webb, Arianna Schuler Scott, Jaimie Lee Freeman, Lars Kunze, Petr Slovak, et al. 2022. Ethical Risk Assessment for Social Robots: Case Studies in Smart Robot Toys. In *Towards Trustworthy Artificial Intelligent Systems*. Springer, 61–76.

[70] Langdon Winner. 1980. Do artifacts have politics? *Daedalus* 109, 1 (1980), 121–36.

[71] Ricarda Wullenkord and Friederike Eyssel. 2020. Societal and ethical issues in HRI. *Current Robotics Reports* 1, 3 (2020), 85–96.

[72] Selma Šabanović and Wan-Ling Chang. 2016. Socializing robots: constructing robotic sociality in the design and use of the assistive robot PARO. *AI & SOCIETY* 31, 4 (Nov. 2016), 537–551. https://doi.org/10.1007/s00146-015-0636-1