Designing Better Human-Robot Interactions Through Enactment, Engagement, and Reflection

Hannah Pelikan
hannah.pelikan@liu.se
Linköping University
Linköping, Sweden

David Porfirio
david.porfirio.ctr@nrl.navy.mil
NRC Postdoctoral Research Associate
Naval Research Laboratory
Washington, DC, USA

Katie Winkle
katie.winkle@it.uu.se
Uppsala University
Stockholm, Sweden

ABSTRACT

Designing the *interaction* in human-robot interaction (HRI) is difficult, with coherent design guidelines still lacking. Other fields such as conversational user interfaces (CUI) and human-computer interaction (HCI), on the other hand, have faced similar challenges and have consequently created guidelines and frameworks for designing improved interaction experiences. We present a novel design method for HRI, which we call *StEER*, which helps bridge the design gap between HRI and other fields. Our method is motivated by the goal of supporting inexperienced or resource-constrained designers in designing with their whole body, focusing on concrete, video-taped interaction and critical reflection and documentation of design choices. We discuss our experiences of working through this method as we iterated and refined it to substantiate our claims that *StEER* (a) is a substantial contribution to HRI design methods, (b) facilitates design for (conversational) interaction, and (c) supports feminist HRI.

CCS CONCEPTS

• **Human-centered computing** → **Interaction design process and methods**; HCI design and evaluation methods; • **Computer systems organization** → *Robotics*.

KEYWORDS

interaction design, situated interaction, video analysis, feminist HRI

ACM Reference Format:
Hannah Pelikan, David Porfirio, and Katie Winkle. 2023. Designing Better Human-Robot Interactions Through Enactment, Engagement, and Reflection. In *Proceedings of the CUI@HRI Workshop at the 2023 ACM/IEEE International Conference on Human-Robot Interaction (HRI '23), March 13, 2023, Stockholm, Sweden.* ACM, New York, NY, USA, 7 pages.

1 INTRODUCTION

Designing robots to interact with humans (the *interaction* component of human-robot interaction) is especially difficult for resource-constrained designers, developers, and decision makers. Human-robot interaction (HRI) designers often lack formal expertise in understanding interaction norms, sufficient time to develop and nuance their designs, or opportunities to reflect on how design choices may impact interactions with target users. Thus, what a robot says and does is today often determined in an *ad hoc* manner.

Interface and user experience (UX) designers facing similar challenges in the 1980s and 1990s resulted in the development of guidelines and tools for the design of human-computer interaction (HCI). These guidelines offer partial relief to resource-constrained designers. The conversational user interface (CUI) community has similarly begun to call for and develop similar guidelines specifically for voice interfaces [23, 25, 26]. While HCI and CUI interfaces are usually immobile and static, robots, by contrast, exist in a physical medium and incorporate a range of different interaction modalities. The embodiment of robots poses many design challenges that are unique to HRI, which prevents the ability to directly, and only, draw from those rigorous design approaches that are established in HCI and emerging for CUI. We therefore see a need for the development of HRI design methods while taking inspiration from the important lessons that have been learned in HCI and CUI.

We see two main strands of HRI work that focus on detailed robot *interaction* design, rather than the more abstract design of robot applications and functionality: The first strand encompasses work that applies rigorous design methodology to design robot movement [*e.g.,* 2, 44, 49] or sound [22, 36], to name a few examples. Such work is time costly and requires in-depth training, which is often restricted to the labs and schools in which such methods have been developed. Furthermore, recent work has highlighted that HRI design publications are difficult to generalize from, particularly when trying to translate insights to other robots [19]. The second strand of prior work in HRI interaction design draws inspiration from human-human interaction, specifically when designing speech-based interaction. Some of this work closely mimics specific elements of human behavior on a robot [*e.g.,* 27, 45], or formulates design suggestions by observing human interaction in detail [e.g., 18]. Such work is often carried out without involving formal design methods and iterations, making it difficult to translate it to contexts beyond the study setting.

Despite the existence of these strands of work, HRI design is often an *ad hoc* task, in which individual practitioners, such as PhD students or individual contractors tasked with managing a robot, make decisions about how a robot should interact. In contrast to design professionals, these *lay designers* have limited resources and may not have access to experts to consult during the design process. Lay designers themselves are unlikely to be familiar with interaction design techniques such as bodystorming, improvisation, animation,

HRI '23, March 13, 2023, Stockholm, Sweden

video prototyping and so forth. Furthermore, mimicking human-human interaction in all its detail is neither practically feasible nor necessarily appropriate for many HRI use cases. Consequently, lacking guidance, lay designers draw on preconceived intuitions about how humans act without further exploring or enacting them, developing implementations 'on-the-fly' rather than practicing any formal design process. Such *ad hoc* design practice typically results in a failure to document any specific choices, assumptions, or prerequisites, which hence are "built in" (unknowingly or otherwise) to the final interaction. Additionally, practicing *ad hoc* design contrasts with the generally agreed upon working standards of many professional communities, such as the software development community, which advocates *commenting* and *committing* code for traceability and documentation. We observe a lack of readily available design methods, standards, and tools that support lay designers in designing multimodal conversational interactions with robots, including in the documentation of their specific design choices.

To address these issues, we developed an interaction prototyping method with lay designers in mind, but potentially more widely applicable to others. Our method focuses on designing for embodied *conversational* human-robot interactions, which uses the whole body and encourages critical reflection on the design process. Our method consist of four main stages, **St**op, **E**nact, **E**ngage, and **R**eflect (*StEER*). In what follows, we motivate our approach, present our method, and discuss our envisioned next steps.

2 DESIGNING A DESIGN METHOD

We believe that HRI can learn from developments in CUI to focus more specifically on interaction design, rather than, for instance, the design of a robot's form factor. This motivates the question, *how can we support lay designers and developers in engaging more concretely with the interaction component of HRI*? To what extent, if any, do the methods we are using to design and prototype human-robot interactions already do so?

In this paper, we present in-progress work on our design method, *StEER*, with the goal of eventually facilitating this method with a software tool or collection of tools. To design our method, we began with reflections on what we believe to be the core knowledge gap faced by lay HRI designers: *what makes for a good human-robot interaction?* We take a position similar to Marc Hassenzahl and his conception of HCI as *experience design* [15]; that is, HRI is fundamentally about delivering or supporting *experiences*, and HRI design should start from desired experience rather than technological capabilities of interaction.

2.1 Taking a First Person Perspective

Our work in developing *StEER* began with speculation of a design method that is simultaneously sensitive to user *and* designer embodiment, self-disclosure and ecology, and how this might influence resultant HRI design. Our initial ideation for *StEER* thereby began an exercise to satisfy these criteria in addition to supporting the transparency and documentation that we advocate should be present within design.

Taking inspiration from service robots that are beginning to be deployed in a range of different stores or offices in countries such as Japan, the United States, and Sweden, we imagined ourselves as

Figure 1: We enacted a *barista* scenario using a variety of different tools and techniques. One author performed the enactment step with a virtual background of a coffee shop (left); another author enacted using their phone camera (center); and another author enacted while sitting at their desktop computer (right).

lay designers and catalyzed our ideation with a toy *barista* scenario of a service robot deployed in a coffee shop. In realizing that lay designers may be tempted to begin their design task by mapping the robot's functionality to its task, we wondered how we might instead facilitate the engagement of lay designers' own first person experiences as customers and visitors of coffee shops. To begin, each of use chose a moment in memory to engage with further in addressing the toy scenario. Engaging with our own experiences, we learnt three key lessons about the barista scenario that would not have been evident had we jumped to prematurely coding the robot without any reflection.

First, drawing on an experience in which the customer had not felt "seen" by the barista, we realized that going to a coffee shop is usually not about coffee alone. Instead, ordering a drink and a meal at a café is a social experience beyond simply feeding oneself. Motivations and expectations may center about the social interaction with the barista, about being seen and treating oneself to a break. In the coffee shop setting, a robot that makes ordering faster and more efficient might for instance collide with the very reasons for choosing a café over a coffee machine.

Second, engaging with a different moment within our past experiences, we also reflected on the social expectations around paying for things. Even if coffee may be ordered in a group sometimes, this may not mean that everyone is necessarily paying together. Negotiating and displaying who is with the group and who is going to pay are subtle and embodied skills that may be easily missed when taking a simplified one-user-one-robot perspective.

Finally, we reflected on our experiences as foreigners drinking coffee in other countries. We observed that while negotiating what language to order in may be tricky—a source of embarrassment and offense or an opportunity to put your *Duolingo* skills to the test, depending on how the interaction goes— the steps in which a visit to a café is accomplished and how food is ordered are very similar across cultures. There are aspects that seem quite independent of the language, such as waiting until the barista is acknowledging the next customer, then making an order, paying, and then later picking up or serving the food.

Working with these concrete first-hand experiences has been a crucial element in our workflow and has helped us to reflect on social expectations and stimulated discussion about alternative barista designs. We realized that our approach to the barista robot culminated in a few distinct steps. First, we *stopped* ourselves

from jumping to an ill-informed robot implementation without due consideration. To catalyze our contemplation, we *enacted*, or role-played, ourselves within the barista scenario. Figure 1 displays still images of videos that we captured during our enactment process, with the intent on sharing these videos with each other to foster ideation of what makes an effective barista interaction. Following our enactment, we *engaged* with our design artifacts—the videos—by re-watching them and taking notes. Finally, we extracted design decisions from these artifacts through *reflection*. In the following sections we describe our stop-enact-engage-reflect *StEER* method, focusing on how our method can be reproduced by others.

3 THE *STEER* METHOD

Our method involves four key steps: Stop, Enact, Engage and Reflect (*StEER*, shown in Figure 2). In the following, we describe our vision of formalizing each of these steps. We highlight, however, that while we identified each step as a standalone component within the process, lay designers may find themselves iterating back and forth between these, sometimes blurring the boundaries between the start and finish of each step.

3.1 Stop

Rather than diving immediately into the premature design or programming of specific robot behaviors, *StEER* encourages lay designers to pause before making any design decisions. We encourage the initial contemplation on designers' own experiences that are relevant for the design case to situate themselves in the concrete setting being designed for. At this stage, different experiences may come to mind, both good and bad. We encourage the recall of memories and concrete experiences that may even appear as mundane. In cases where the designer has no experience at all, this stage is important for becoming aware of a possible discrepancy between the lay designer's identity and the identities of the people being designed for, such as a non-menstruating lay designer designing a robot for interacting with menstruating people. Becoming aware of these discrepancies can encourage lay designers to seek for others' experiences, such as by conversing with people in their surroundings and reading material online. In the case of the coffee shop, paying particular attention when visiting local coffee shops was a good way to complete this step. In observing others and oneself, this stage can be seen to share similarities with existing ethnographic and autoethnographic design approaches [14]. It encourages reflexivity, but does not necessarily aim to provide full-fledged ethnographic descriptions of the setting.

3.2 Enact

As a second step, the lay designer repeatedly enacts memorizable moments within an imagined target design context, ideally while video recording oneself. In our case, we reenacted moments that we had experienced in coffee shops, discovering different paths that the barista interaction can take. Some of the authors preferred taking notes prior to enactment and then acting out a loose script, making modifications as they re-enacted it repeatedly. Others enacted entirely from memory, modifying the scene during repeated trials. Enacting our pre-written notes typically led to changes in this particular script, as we spoke in more colloquial ways than initially written down, and added prosody and gestures, which were absent from the initial written script.

As we performed these exercises individually, we also found different solutions for enacting different people, using props such as a hat or jacket or differing virtual backgrounds to designate different people in the interaction (*e.g.,* Figure 1, left), standing in different locations, or recording different roles separately. What was important is that we video recorded all re-enactments, giving us the opportunity to re-watch what we did. We captured recordings on laptops, desktops, and mobile phones, generally finding that handheld recordings lead to more natural, rapid enactments, perhaps due to the ubiquity of taking selfies and recording with handheld devices.

Enactments are intended to draw on the whole body, forcing the designer to engage with the interaction in real-time. Thereby, on the one hand, they share parallels with improvisation [43] and bodystorming [28], which are common methods in interaction design that encourage the designer to engage with the interaction that is to be designed through their own body. Additionally, re-enactments of transcribed interaction snippets are also sometimes used in fields studying human interaction such as ethnomethodology, helping the analyst to develop a deeper understanding of specific aspects of an interaction by performing them [46]. More generally, when acting as a man or woman, a young or old person, or people from different nationalities, our goal with the enactment stage is for lay designers to be immediately and concretely confronted with the fact that their own body may be quite different from the envisioned users. We started by recording situations involving human-human interaction, but with a Wizard-of-Oz paradigm [7], one may of course transition to enact a robot sooner or later as well.

3.3 Engage with the material

The third step involves re-watching the video recordings of the enacted scenes, taking notes on one's observations. In our case, we rewatched the videos together, discussing interaction paths that others had enacted and starting to formulate key considerations for designing a robot barista. At this stage, we advise lay designers to start producing a transcript of the interaction, either by correcting and adjusting the initial script or by producing written, script-like descriptions of what people are saying and doing in the interaction.

The engagement step shares similarities with and can be informed by transcription practices from fields such as conversation analysis [21, 50], which encourage the video analyst to write down exactly what they hear and see. Engagement therefore is an exercise in becoming aware of how the interaction evolves on a moment-by-moment basis, exploring how one action leads to the next, and how each utterance influences what responses should follow in the interaction. We would like to stress that this stage is *not* the same as video coding, since the goal is not to quantify or count any behaviors but rather to foster a deeper understanding of what is happening in a particular instance and how one interactional move builds on a previous one and leads to the next [41]. Engaging with how interaction progresses in real-time can help lay designers to become aware of the specific interactional moves and resources that contribute to this progress: a gesture, a certain word, or a particular prosody to an utterance.

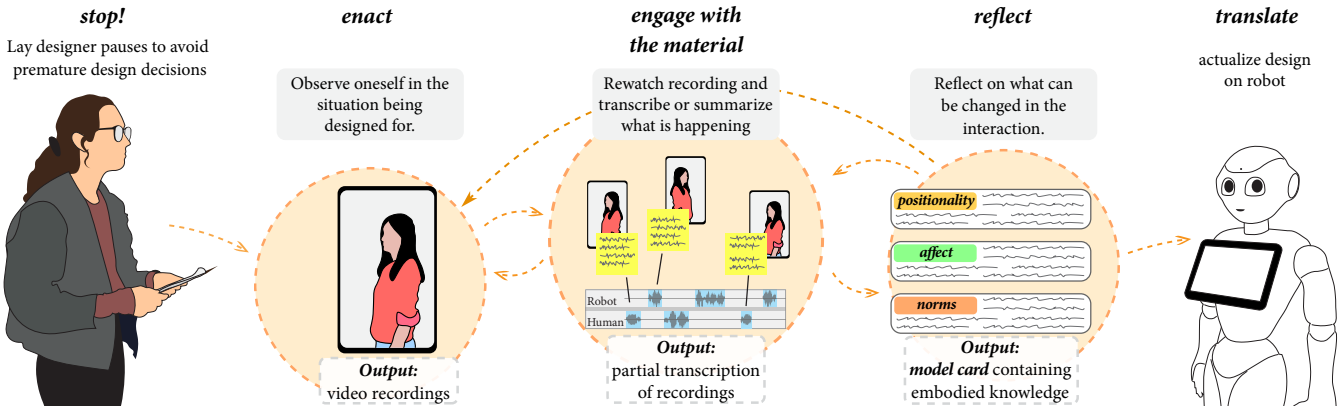

Figure 2: The pipeline that describes how the *StEER* method is used.

Rewatching and discussing our videos, we then started to develop alternative interaction trajectories, gradually working out what is important in a barista encounter. The transcribed snippets can later be used as first drafts of what the robot should say, or serve as scripts for further performance of the enactment step of *StEER* with alternative trajectories.

3.4 Reflect

The final step of *StEER* involves thorough reflection. Motivated by increasing calls for technology designers to reflect on and make explicit their biases, expectations, assumptions, and positionality, we note the need to support designers in identifying and documenting their design process in an explicit way that can be revisited later and could be used and exported for transparency.

Whilst we do recommend that designers take time at the end of each design iteration to engage solely in reflection on the aforementioned topics, reflection is likely to naturally emerge and be documented during engagement with the material. For example, an author who had that experience of not being "seen" (Section 2.1) noted, on entering the cafe, that the barista was engaged in an intense conversation in Italian with some customers in the back of the store. The barista did not take note of her arrival and continued their conversation with the patrons even as she approached the counter hoping to order. The barista eventually acknowledged her by saying "what do you want?" Later, reflecting on the experience, the author was able to explicize her expectations: to be attended to a little more closely and politely, in line with her expectations and previous experiences grounded within the cozy British tea shops and the hipster coffee shops in Stockholm that she typically frequented. She found herself wondering whether, if she was instead an Italian-speaking woman used to dashing between stores in the busy streets of Milan or Rome, she would still have felt slighted and unseen in this interaction, or whether she might have been quicker and more effective at getting the barista's attention and may even have found a pleasant reminder of 'home in being surrounded by familiar language and behavior.

In the context of the design of a barista robot, our discussions prompted reflection and further discussion on the fact that, whether explicitly or not, humans design interactions according to a specific set of cultural norms, most likely reflecting on our own experiences and positionality. Our biased interaction designs will (a) not result in the same interaction experience for all users, and (b) propagate and reflect particular norms as being the "correct" or "expected" way to behave. We propose that the *StEER* method can incorporate recent reflexive questions put forward in the context of Feminist HRI design/development practice [53]. The reflection step can encourage designers to work through these questions and document their answers in a format designed to accompany/be released alongside their final design, akin to the *model cards* being used as a way to document (amongst other things) *"the context in which [machine learning] models are intended to be used, details of the performance evaluation procedures, and other relevant information"* (emphasis our own) [20].

4 FUTURE DIRECTIONS

In order to further refine and iterate on *StEER*, we plan to run a number of small design workshops at HRI/social robotics labs where we invite HRI researchers to work through *StEER* to address a set of HRI design briefs. We, as a project team, are diverse in our backgrounds, ranging from interaction analysis to HRI systems engineering. We recognize how our current version and imagined application of *StEER* reflects our unique motivations and skillsets that we each brought to this project, but we want to ensure that we are designing a method that works for (would be) HRI designers of other backgrounds, too.

4.1 Designing Better Human-Robot Interactions

We see several ways in which *StEER* can address current design concerns, which we highlight below.

4.1.1 Making HRI More Inclusive. We posit that *StEER* can make HRI more inclusive. As we highlighted in the introduction (§1), current HRI design methodology often requires co-present team members and/or experience in engaging with formal improvisation and roleplay techniques. Engaging in bodystorming, improvisation or sketching storyboards can be intimidating or unapproachable for those who have no experience with such methods, and possibly even inaccessible to researchers who are not part of labs where

these methods are developed. We have particularly focused on lay designers, thinking of ways that are intuitive and can be rapidly applied to a design task, combining enactment with sitting solitarily at one's own desk. Lay designers typically lack formal training in interaction design, drawing instead on intuition to make impromptu design decisions. Despite the existence of a vast body of design methodology within related fields such as HCI, which has partly inspired design methodology in HRI, lay designers likely do not leverage these potentially useful methods to their fullest extent due to lack of familiarity and/or confidence. Perhaps lay designers do not even think of themselves as *designers*—a recent survey of HRI researchers indicates that the overwhelming majority have educational backgrounds not in design, but in computer science and engineering [52]. Yet, taking Herbert Simon's concept of design as devising 'courses of action aimed at changing existing situations into preferred ones' [42], anyone deploying a robot which will result in HRI experiences, and/or working to improve particular HRI experiences, is indeed a designer.

Being applicable both to one person and to groups, we believe that *StEER* can provide more democratic access to good design. By contributing a way to document design choices, our method provides concrete guidelines on *how* to design, not just showing design results [19]. Similarly, it encourages designers to reflect their understanding of interaction through encouraging enactment and analysis.

4.1.2 Designing for Interaction. We further argue that our method is interactional in nature in that it encourages designers to build interactions that are better adapted to the contingency and situatedness of the real world [39, 48]. HRI work has traditionally focused on designing robots, focusing on individual aspects such as form factor, movement, or sound [10, 24, 35], rather than explicitly designing full-fledged interactions that incorporate these aspects [31]. Social interaction is only one of several metrics on which robots can be evaluated [47], and until now it is often measured in abstract concepts such as trust [40] or expressivity [49]. Evaluations of actual moment-by-moment interaction with robots remain rare [29, 37, 51].

Lay designers, often with a background in computer science and engineering, are trained to create technology, but typically do not learn much about human communication. As interaction analytic work has pointed out repeatedly, people do not speak like one might imagine them to do so [9, 38]. It is therefore particularly problematic when people who are not trained in studying interaction construct imagined interactional scripts with no further design exploration. Although *StEER* does not replace studying interactions with robots in real world settings as a complementary option, it encourages designers to engage with how people actually speak, using the whole body, applying prosody and silences, which is quite different from written texts. We believe that this aspect of our method can be particularly interesting for the conversational HRI community working on spoken interaction.

In designating interaction as the central goal, *StEER* shares parallels with interactional streams of work within CUI, particularly Ethnomethodologic Conversation Analysis [13, 16, 39]. Such work has repeatedly highlighted how the current design of voice interfaces fails to keep up with interactional dynamics [1, 30, 32]. We believe that HRI can learn from HCI and CUI to view robots as interfaces that can be designed to stimulate interaction [4]. Rather than taking an input-output model with discrete moments of engagement, such as providing input to the robot by answering a question, our method encourages lay designers to focus on how interactions progress over time. Fischer et al. [11] have previously argued how CUIs can be designed for this interactional *progressivity*. We believe that our method can be seen as a step towards encouraging such a focus on interactional progress also in HRI, since it encourages designers to think of the whole interaction, including how users may react to what a robot says and does.

4.1.3 Supporting Feminist HRI. Finally, we posit that our method directly supports Feminist HRI practice, as recently put forward as a power-sensitive approach to HRI design, yielding better, and more ethical HRI regardless of the specific application context [53]. Drawing on those qualities and principles of feminist interaction put forward by Winkle et al., which in turn draw from Data Feminism [8] and Feminist HCI [3] we suggest our method particularly supports:

- a focus on embodiment and ecology,
- explicit reflection on and capture of designer positionality,
- explicit reflection on and capture of norms (un)knowingly designed into (or explicitly challenged) in resultant design,
- pluralistic thinking, as the designer is invited to roleplay as multiple interactants.

4.2 Imagining A Software Design Suite for *StEER*

We view our method as generally technology agnostic—beyond the hardware requirement for video capture and re-play, *StEER* can be practiced without a dedicated end-user development environment. Still, we believe that technology has the potential to help support and facilitate the *StEER* method, especially for designers new to *StEER*. Therefore, we see the potential to embody *StEER* within a software end-user development suite.

Although existing state-of-the-art robot programming tools have proven effective for robot end users and designers who lack programming expertise (*e.g., Interaction Composer* [12]) or expertise in robot norms (*e.g., RoVer* [34]), no existing software environment could be easily adapted to facilitate the *StEER* method. In imagining our own software suite, we draw inspiration from existing tools that embody other design methods, such as *Synthé*'s use of the *bodystorming* design approach [33] and *KIWIS* use of an I-know-it-when-I-see-it design approach [17]. Many other existing tools utilize a similar *in-situ* approach to *KIWIS*, such as *GhostAR* [5] and *V.Ra* [6], which utilize augmented reality to situate design decisions within the robot's physical environment.

We envision our own software suite to facilitate the *enact*, *engage*, and *reflect* components of *StEER* shown in Figure 2. As an example of a potential software suite, the *enact* phase might occur on an app within a handheld device with a camera, such as a mobile phone. The camera can be used to self-record videos of lay designers enacting hypothetical interactions, and the app can allow simple annotations of key points to revisit during the *engage* phase. The videos can then be uploaded to a desktop *engagement* application, enabling

designers to scrub through, transcribe, annotate, and ultimately analyze their videos. Finally, the same desktop application can be used for the *reflect* phase, which automatically produces a model card of design decisions and justifications based on the designer's reflections.

5 CONCLUSION

We presented a design method that consists of four main steps: (1) stopping and becoming aware of one's own experiences before starting to design, (2) enacting relevant situations from memory and exploring alternative interaction trajectories that could occur with a robot, (3) engaging with the video-recorded enactments through, for instance, transcribing them, and (4) reflecting on one's own positionality and documenting design choices. We call our method *StEER*.

We argue that *StEER* can be applied by individuals without formal design training, focusing on interaction as embodied and situated accomplishment. Being grounded in the phenomenological presence of the body, rather than claiming objectivity, our method can be regarded as a concrete example of feminist HRI. Our design method contributes to ongoing conversations in HRI on documenting design processes [19], and to CUI work that explores the design of progressivity in interactions [11].

ACKNOWLEDGMENTS

We would like to thank the organizers of the KTH RPL Summer School, which stimulated our collaboration on *StEER*. HP is supported by the Swedish Research Council (2016-00827). KW is partly funded by the Wallenberg AI, Autonomous Systems and Software Program—Humanities and Society (WASP-HS) funded by the Marianne and Marcus Wallenberg Foundation and the Marcus and Amalia Wallenberg Foundation. DP's involvement with this work occurred while being affiliated with the University of Wisconsin–Madison.

REFERENCES

[1] Saul Albert and Magnus Hamann. 2021. Putting wake words to bed: We speak wake words with systematically varied prosody, but CUIs don't listen. In *CUI 2021 - 3rd Conference on Conversational User Interfaces* (Bilbao (online) Spain). ACM, 1–5. https://doi.org/10.1145/3469595.3469608

[2] Lucy Anderson-Bashan, Benny Megidish, Hadas Erel, Iddo Wald, Guy Hoffman, Oren Zuckerman, and Andrey Grishko. 2018. The greeting machine: An abstract robotic object for opening encounters. In *2018 27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 595–602. https://doi.org/10.1109/ROMAN.2018.8525516 ISSN: 1944-9437.

[3] 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 (CHI '10)*. Association for Computing Machinery, New York, NY, USA, 1301–1310. https://doi.org/10.1145/1753326.1753521

[4] Michel Beaudouin-Lafon. 2004. Designing Interaction, Not Interfaces. *Proceedings of the Working Conference on Advanced Visual Interfaces*, 15–22. https://doi.org/10.1145/989863.989865

[5] Yuanzhi Cao, Tianyi Wang, Xun Qian, Pawan S. Rao, Manav Wadhawan, Ke Huo, and Karthik Ramani. 2019. GhostAR: A Time-Space Editor for Embodied Authoring of Human-Robot Collaborative Task with Augmented Reality. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 521–534. https://doi.org/10.1145/3332165.3347902

[6] Yuanzhi Cao, Zhuangying Xu, Fan Li, Wentao Zhong, Ke Huo, and Karthik Ramani. 2019. V.Ra: An In-Situ Visual Authoring System for Robot-IoT Task Planning with Augmented Reality. In *Proceedings of the 2019 on Designing Interactive Systems Conference* (San Diego, CA, USA) *(DIS '19)*. Association for Computing Machinery, New York, NY, USA, 1059–1070. https://doi.org/10.1145/3322276.3322278

[7] Nils Dahlbäck, Arne Jönsson, and Lars Ahrenberg. 1993. Wizard of Oz Studies: Why and How. *Proceedings of the 1st International Conference on Intelligent User Interfaces*, 193–200. https://doi.org/10.1145/169891.169968

[8] Catherine D'Ignazio and Lauren F. Klein. 2020. *Data Feminism.* MIT Press. Google-Books-ID: x5nSDwAAQBAJ.

[9] N J Enfield. 2013. *Relationship Thinking: Agency, Enchrony, and Human Sociality.* OUP.

[10] Friederike Eyssel, Dieta Kuchenbrandt, Simon Bobinger, Laura de Ruiter, and Frank Hegel. 2012. 'If You Sound like Me, You Must Be More Human': On the Interplay of Robot and User Features on Human-Robot Acceptance and Anthropomorphism. In *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction* (Boston, Massachusetts, USA) *(HRI '12)*. Association for Computing Machinery, New York, NY, USA, 125–126. https://doi.org/10.1145/2157689.2157717

[11] Joel E Fischer, Stuart Reeves, Martin Porcheron, and Rein Ove Sikveland. 2019. Progressivity for Voice Interface Design. *Proceedings of the 1st International Conference on Conversational User Interfaces.* https://doi.org/10.1145/3342775.3342788

[12] Dylan F Glas, Satoru Satake, Takayuki Kanda, and Norihiro Hagita. 2012. An Interaction Design Framework for Social Robots. *Robotics: Science and Systems VII* (2012).

[13] Charles Goodwin. 2018. *Co-operative action.* Cambridge University Press.

[14] Cathrine Hasse, Stine Trentemøller, and Jessica Sorenson. 2019. Special Issue on Ethnography in Human-Robot Interaction Research. *Paladyn, Journal of Behavioral Robotics* 10, 1 (2019), 180–181. https://doi.org/doi:10.1515/pjbr-2019-0015

[15] Marc Hassenzahl. 2010. *Experience design: Technology for all the right reasons.* Morgan & Claypool Publishers.

[16] Ian Hutchby and Robin Wooffitt. 2008. *Conversation Analysis* (sse ed.). Polity Press.

[17] Heather Knight and Reid Simmons. 2017. An intelligent design interface for dancers to teach robots. In *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. 1344–1350. https://doi.org/10.1109/ROMAN.2017.8172479

[18] Dimosthenis Kontogiorgos and Joakim Gustafson. 2021. Measuring Collaboration Load With Pupillary Responses - Implications for the Design of Instructions in Task-Oriented HRI. *Frontiers in Psychology* 12 (2021), 2355. https://doi.org/10.3389/fpsyg.2021.623657

[19] Maria Luce Lupetti, Cristina Zaga, and Nazli Cila. 2021. Designerly Ways of Knowing in HRI: Broadening the Scope of Design-Oriented HRI Through the Concept of Intermediate-Level Knowledge. *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, 389–398. https://doi.org/10.1145/3434073.3444668

[20] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. 2018. Model Cards for Model Reporting. *CoRR* abs/1810.03993 (2018). arXiv:1810.03993 http://arxiv.org/abs/1810.03993

[21] Lorenza Mondada. 2018. Multiple Temporalities of Language and Body in Interaction: Challenges for Transcribing Multimodality. *Research on Language and Social Interaction* 51, 1 (Jan. 2018), 85–106. https://doi.org/10.1080/08351813.2018.1413878

[22] Dylan Moore, Rebecca Currano, G Ella Strack, and David Sirkin. 2019. The Case for Implicit External Human-Machine Interfaces for Autonomous Vehicles. *Proceedings of the 11th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, 295–307. https://doi.org/10.1145/3342197.3345320

[23] Robert J Moore and Raphael Arar. 2019. *Conversational UX Design: A Practitioner's Guide to the Natural Conversation Framework.* Association for Computing Machinery.

[24] Jonathan Mumm and Bilge Mutlu. 2011. Human-Robot Proxemics: Physical and Psychological Distancing in Human-Robot Interaction. In *Proceedings of the 6th International Conference on Human-Robot Interaction* (Lausanne, Switzerland) *(HRI '11)*. Association for Computing Machinery, New York, NY, USA, 331–338. https://doi.org/10.1145/1957656.1957786

[25] Christine Murad and Cosmin Munteanu. 2019. "I Don't Know What You're Talking about, HALexa": The Case for Voice User Interface Guidelines. In *Proceedings of the 1st International Conference on Conversational User Interfaces* (Dublin, Ireland) *(CUI '19)*. Association for Computing Machinery, New York, NY, USA, Article 9, 3 pages. https://doi.org/10.1145/3342775.3342795

[26] Christine Murad, Cosmin Munteanu, Benjamin R. Cowan, and Leigh Clark. 2019. Revolution or Evolution? Speech Interaction and HCI Design Guidelines. *IEEE Pervasive Computing* 18, 2 (2019), 33–45. https://doi.org/10.1109/MPRV.2019.2906991

[27] Bilge Mutlu, Takayuki Kanda, Jodi Forlizzi, Jessica Hodgins, and Hiroshi Ishiguro. 2012. Conversational Gaze Mechanisms for Humanlike Robots. *ACM Trans. Interact. Intell. Syst.* 1 (1 2012), 12:1–12:33. Issue 2. https://doi.org/10.1145/2070719.2070725

[28] Antti Oulasvirta, Esko Kurvinen, and Tomi Kankainen. 2003. Understanding contexts by being there: case studies in bodystorming. *Personal and ubiquitous computing* 7 (2003), 125–134. https://doi.org/10.1007/s00779-003-0238-7

[29] Hannah Pelikan, Leelo Keevallik, and Mathias Broth. 2020. Människa-robot-interaktion. Studentlitteratur, 395–415.

[30] Hannah R.M. Pelikan and Mathias Broth. 2016. Why That Nao?: How Humans Adapt to a Conventional Humanoid Robot in Taking Turns-at-Talk. *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, 4921–4932. https://doi.org/10.1145/2858036.2858478

[31] Hannah R.M. Pelikan and Malte Jung. 2023. Designing robot sound-in-interaction: The case of autonomous public transport shuttle buses. In *Proceedings of the 2023 International Conference on Human-Robot Interaction* (Stockholm, Sweden) *(HRI '23)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3568162.3576979

[32] Martin Porcheron, Joel E Fischer, Stuart Reeves, and Sarah Sharples. 2018. Voice Interfaces in Everyday Life. *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*, 640:1–640:12. https://doi.org/10.1145/3173574.3174214

[33] David Porfirio, Evan Fisher, Allison Sauppé, Aws Albarghouthi, and Bilge Mutlu. 2019. Bodystorming Human-Robot Interactions. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 479–491. https://doi.org/10.1145/3332165.3347957

[34] David Porfirio, Allison Sauppé, Aws Albarghouthi, and Bilge Mutlu. 2018. Authoring and Verifying Human-Robot Interactions. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (Berlin, Germany) *(UIST '18)*. Association for Computing Machinery, New York, NY, USA, 75–86. https://doi.org/10.1145/3242587.3242634

[35] Laurel D. Riek, Tal-Chen Rabinowitch, Bhismadev Chakrabarti, and Peter Robinson. 2009. How anthropomorphism affects empathy toward robots. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*. ACM, La Jolla California USA, 245–246. https://doi.org/10.1145/1514095.1514158

[36] Frederic Anthony Robinson, Mari Velonaki, and Oliver Bown. 2021. Smooth Operator: Tuning Robot Perception Through Artificial Movement Sound. *Proceedings of the 2021 ACM/IEEE International Conference on Human-Robot Interaction*, 53–62. https://doi.org/10.1145/3434073.3444658

[37] Damien Rudaz, Karen Tatarian, Rebecca Stower, and Christian Licoppe. 2023. From Inanimate Object to Agent: Impact of Pre-Beginnings on the Emergence of Greetings with a Robot. *J. Hum.-Robot Interact.* (jan 2023). https://doi.org/10.1145/3575806 Just Accepted.

[38] Harvey Sacks. 1992. *Lectures on conversation*. Blackwell. OCLC: 254229422.

[39] Harvey Sacks, Emanuel A Schegloff, and Gail Jefferson. 1974. A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language* 50 (12 1974), 696. Issue 4. https://doi.org/10.2307/412243

[40] Maha Salem, Gabriella Lakatos, Farshid Amirabdollahian, and Kerstin Dautenhahn. 2015. Would You Trust a (Faulty) Robot?: Effects of Error, Task Type and Personality on Human-Robot Cooperation and Trust. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction (HRI '15)*. ACM, New York, NY, USA, 141–148. https://doi.org/10.1145/2696454.2696497

[41] Emanuel A Schegloff. 2007. *Sequence Organization in Interaction*. Cambridge University Press. 1–300 pages. https://doi.org/10.1017/CBO9780511791208

[42] Herbert A Simon. 2019. *The Sciences of the Artificial, reissue of the third edition with a new introduction by John Laird*. MIT press.

[43] David Sirkin and Wendy Ju. 2015. *Embodied Design Improvisation: A Method to Make Tacit Design Knowledge Explicit and Usable*. Springer International Publishing, Cham, 195–209. https://doi.org/10.1007/978-3-319-06823-7_11

[44] David Sirkin, Brian Mok, Stephen Yang, and Wendy Ju. 2015. Mechanical Ottoman: How Robotic Furniture Offers and Withdraws Support. *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, 11–18. https://doi.org/10.1145/2696454.2696461

[45] Gabriel Skantze, Anna Hjalmarsson, and Catharine Oertel. 2014. Turn-taking, feedback and joint attention in situated human–robot interaction. *Speech Communication* 65 (2014), 50–66. https://doi.org/10.1016/j.specom.2014.05.005

[46] Philippe Sormani. 2020. Re-enactment as a Research Strategy: From Performance Art to Video Analysis and Back Again. In *Dialogues between artistic research and science and technology studies*, Henk Borgdorff, Peter Peters, and Trevor Pinch (Eds.). Routledge, Taylor & Francis Group, New York, 184–200.

[47] Aaron Steinfeld, Terrence Fong, David Kaber, Michael Lewis, Jean Scholtz, Alan Schultz, and Michael Goodrich. 2006. Common Metrics for Human-Robot Interaction. In *Proceedings of the 1st ACM SIGCHI/SIGART Conference on Human-Robot Interaction* (Salt Lake City, Utah, USA) *(HRI '06)*. Association for Computing Machinery, New York, NY, USA, 33–40. https://doi.org/10.1145/1121241.1121249

[48] Lucy A Suchman. 1987. *Plans and Situated Actions: The Problem of Human-Machine Communication*. Cambridge University Press.

[49] Leila Takayama, Doug Dooley, and Wendy Ju. 2011. Expressing Thought: Improving Robot Readability with Animation Principles. *Proceedings of the 6th International Conference on Human-Robot Interaction*, 69–76. https://doi.org/10.1145/1957656.1957674

[50] Paul ten Have. 2002. Reflections on transcription. *Cahiers de praxématique* (1 2002), 21–43. Issue 39. https://doi.org/10.4000/praxematique.1833

[51] Sylvaine Tuncer, Sarah Gillet, and Iolanda Leite. 2022. Robot-Mediated Inclusive Processes in Groups of Children: From Gaze Aversion to Mutual Smiling Gaze. *Frontiers in Robotics and AI* 9 (3 2022). https://doi.org/10.3389/frobt.2022.729146

[52] Katie Winkle, Erik Lagerstedt, Ilaria Torre, and Anna Offenwanger. 2022. 15 Years of (Who)man Robot Interaction: Reviewing the H in Human-Robot Interaction. *ACM Transactions on Human-Robot Interaction* (Nov. 2022). https://doi.org/10.1145/3571718

[53] Katie Winkle, Donald McMillan, Maria Arnelid, Madeline Balaam, Katherine Harrison, Ericka Johnson, and Iolanda Leite. 2023. Feminist Human-Robot Interaction: Disentangling Power, Principles and Practice for Better, More Ethical HRI. In *Proceedings of the 18th International Conference on Human-Robot Interaction* (Stockholm, Sweden) *(HRI '23)*. Association for Computing Machinery, New York, NY, USA. https://doi.org/10.1145/3568162.3576973