



# Recipient design in human–robot interaction: the emergent assessment of a robot’s competence

Sylvaine Tuncer<sup>1</sup> · Christian Licoppe<sup>2</sup> · Paul Luff<sup>1</sup> · Christian Heath<sup>1</sup>

Received: 6 April 2022 / Accepted: 21 November 2022  
© The Author(s) 2023

## Abstract

People meeting a robot for the first time do not know what it is capable of and therefore how to interact with it—what actions to produce, and how to produce them. Despite social robotics’ long-standing interest in the effects of robots’ appearance and conduct on users, and efforts to identify factors likely to improve human–robot interaction, little attention has been paid to how participants evaluate their robotic partner in the unfolding of actual interactions. This paper draws from qualitative analyses of video-recorded interactions between a robot and groups of participants, in the framework of ethnomethodology and conversation analysis. We analyse the particular ways in which participants shape their embodied actions, how they can reproduce a prior action that failed to obtain a response from the robot; and how they explore the robot’s embodied nature. We find a set of recurrent methods or practices, showing that robot-recipient design displays not only participants’ initial assumptions about the robot’s competences, but also more importantly perhaps their continuous assessment of the robot’s behaviour, and their attempts to adapt to it. Participants locally produce and constantly revise their understanding of the robot as a more or less competent co-participant, drawing from its past, current, and projected conduct and responsiveness. We discuss the implications of these findings for research in robotics and human–robot interactions, and the value of the approach to shed new light on old questions by paying attention to the quality of gesture and the sequential organisation of interaction.

**Keywords** Human–robot interaction · Recipient design · Conversation analysis · Embodied action · Interactional competence

## 1 Introduction

When they encounter a robot for the first time, human participants rarely know what it is capable of, especially what its interactional skills are. This unknown poses a range of problems, not the least as to what actions the robot is likely to recognise<sup>1</sup> and respond to, and how these actions should be designed for the robot to recognise them and for it to be able respond. In this paper, we explore some of the ways in which human participants attempt to interact with robots in first encounters, with a focus on the quality of embodied actions and sequentiality.

Studies in human–robot interaction (HRI) and social robotics have focused on robots’ conduct to identify components which enable smooth interactions, such as arm and head movements (Baddoura and Venture 2015), the timing to initiate talk (Gehle et al. 2017), or the level of embodiment (Kontogiorgos et al. 2020). Most studies rely on questionnaires to elicit participants’ experience and perceptions of the robot’s conduct and social skills (e.g., Baddoura and Venture 2013; Ruijten et al. 2019). In addition to questionnaires, video recordings have been utilised to extract metrics, to propose and test new programmes and suggest recommendations for design. Studies in human–computer interaction have taken a more interdisciplinary and qualitative

---

✉ Sylvaine Tuncer  
Sylvaine.tuncer@kcl.ac.uk

<sup>1</sup> King’s Business School, King’s College, Bush House, 30 Aldwych, London WC2B 4BG, UK

<sup>2</sup> Télécom Paris, Palaiseau Cedex F-91123, France

<sup>1</sup> By ‘recognise’, we mean the robot subsequently acting in a way which human participants treat as responding to their own previous conduct and meaningful, notwithstanding the possibility that the action be fortuitous, and its appropriateness as a response a coincidence. Whether or not robots can understand actions the way humans do is far beyond our considerations (on human–machine conversations as a simulacrum of conversation, see Button and Sharrock 1995).

approach. Some of them explore how best to coordinate a robot's talk and embodied actions to create an appropriate spatial formation, engage participants, or select a recipient (Kuzuoka et al. 2010; Yamazaki et al. 2009, 2012). Others, with less direct implications for design (Dourish 2006), explore how people talk to machines (e.g., Fischer 2010; Fischer et al. 2012; Pelikan and Broth 2016; Porcheron et al. 2018), addressing aspects and phenomena such as semantics, utterance flow, or repair sequences. Besides, these studies shift the attention from perceptions of interactions—typically accessed through questionnaires—to interactions themselves, as they unfold and are accomplished.

In this paper, in an endeavour to further our understanding of interactions and sociality in the making between humans and robots, we take this perspective. We leave perceptions and metrics in the background to provide qualitative insights on the kind of work done by participants when they interact with robots. This is in line with the approach proposed by Fischer (2021), but whereas Fischer aims to track anthropomorphising behaviour, we consider more specifically phenomena whereby interactional competence is attributed to the robot. Human participants' behaviour and the design of their actions exhibit both the indeterminacy and some assumptions of robot's competences, they also show some of the emergent methods and resources to make the interaction work.

## 2 Background

People can take a considerable effort to interact with robots, through language but also through embodied conduct. They do so especially with human-like robots, whose appearance and bodily behaviour tend to create high expectations in terms of social skills (Fink 2012). First encounters (Bergmann et al. 2012) are particularly revealing because, as participants have no reliable knowledge of the robot's interactional competencies, they rely on their own assumptions as well as what they discover about those competencies in the course of their interaction with the robot. In fact, the way participants produce and shape their embodied actions *exhibits* their assumptions and online evaluations of the robot's interactional competence. In doing so, it offers a resource for analysis. In this paper, we present the findings from in-depth, systematic analyses of video-recorded interactions between human participants and a robot (see Methods section below), in the framework of ethnomethodology and conversation analysis (Garfinkel 1967; Sidnell and Stivers 2013). We pay particular attention to both the quality of those actions, and how they fit in and contribute to the sequential organisation of interaction. The notion of sequential organisation, or sequentiality, refers to the way actions build on each other, each projecting a number of

possible next ones, and each responding in one among many possible ways. Courses of actions are thereby seen as collaboratively shaped in a stepwise and emergent fashion. Sequentiality also provides a vehicle for foregrounding the participants' perspective or standpoint, that is, to prioritise how they respond to (and thereby are seen to interpret) the action of the robot, and indeed, modify, repair or alter that response in the light of the robot's response.

In this study, we are primarily interested in the practices that people rely upon and deploy in responding and seeking to interact with robots. We focus on a set of gestures which are recurrent in the initial phase of interactions: hand waves, offers and attempts to shake hands, and embodied referencing. These gestures have a typical and simple design; they also form short sequences of actions amenable to systematic analysis. Additionally, because they are critical in the early stages of encounters, robot designers have put a lot of effort in making those introductory gestures look natural, and they have recurrently used them in experiments (e.g., Baddoura and Venture 2015 for greeting gestures; Mead and Mataric 2016 for pointing gestures).

Thus, our study, with its approach, method and focus, is a response to Fischer's call for studies that “describe the dynamics of people's behaviour over the course of an interaction and in response to robot behaviour” (Fischer 2021: 4:1). Besides, we argue that focusing on participants' continuous assessment of robots' interactional competence re-specifies Fischer's broader interest in “the progression of the expression of anthropomorphism over time” (Fischer 2021: 4:3).

### 2.1 Recipient design and repair in conversation analysis

An overview of our data revealed that participants tend to shape their embodied actions in remarkable, atypical fashions, which display an orientation to the robot's unknown, assumed, and discoverable competencies. This phenomenon builds on a long-standing topic in conversation analytic research on human–human interaction: recipient design, which was first described as “a multitude of respects in which the talk by a party in a conversation is constructed or designed in ways which display an orientation and sensitivity to the particular other(s) who are the co-participants” (Sacks et al. 1974: 727). Since then, recipient design has become a core concept in conversation analysis, mainly studied in talk-in-interaction. Deppermann (2015) specifies that “recipient design of turns is informed by prior knowledge about and shared experience with recipients” (*Ibid.*, 63). Thus, orientation and sensitivity to the recipient are necessary, but not sufficient to produce recipient-designed actions: one also needs minimal acquaintance or familiarity with the co-participant. Not that strangers have no equivalent

resources to interact—they rely on the other being socially competent, and they can make assumptions about each other's basic social characteristics, be they frail and misguided. However, in first encounters with robots, people cannot rely on such common-sense knowledge and assumptions. They need to rely on much more uncertain a priori conceptions or assumptions about the robot; and to create new, empirical knowledge about this particular robot as they interact with it, or at least attempt to. To the best of our knowledge, no detailed analysis has been undertaken so far of recipient design in human–robot interaction with a focus on embodied conduct.

Our study also contributes to research on repair in interaction. Conversation Analysis has extensively studied how participants manage trouble in talk-in-interaction through repair. For example, repair can be initiated in different places with respect to the source of the trouble (Schegloff 1992): a speaker can embed self-correction as they are producing their turn at talk, thereby anticipating a problem in its intelligibility by the hearer, or the recipient can initiate repair in second position with an open-class repair initiator (Drew 1997) or an alternative question seeking clarification (Koshik 2005). While the role of embodied conduct in verbal repair has been studied (e.g., Oloff 2018), the organisation of repair of embodied actions remains largely understudied (except Lerner and Raymond 2017).

## 2.2 The relevance of studying robot-recipient design

Current robots' capacities are still relatively limited compared to the general public's expectations (Malle et al. 2021), which are often influenced by unrealistic representations conveyed in the media (Weiss and Spiel 2021). Thus, it is not uncommon for participants to be disappointed when they meet a robot for the first time (de Graaf et al. 2017). This is all the more unfortunate for robot designers and promoters that first encounters are seen as critical to the subsequent acceptance and adoption of social robots (e.g., Cafaro et al. 2016), as first impressions tend to endure and be difficult to change (Paetzel et al. 2020). To avoid deception and/or disappointment, the principle of transparency was recently proposed in robotics, according to which robots' appearance and behaviour should display their actual capacities—no more, no less (Baillie et al. 2019; Malle et al. 2021; Złotowski et al. 2020). Following such a design principle would seem likely to provide users with resources to adjust their expectations as they meet the robot, and with means to interact with it.

However, despite being an appropriate and relevant objective, designing for transparency is not straightforward. There are few guidelines, propositions or recommendations on how this could be done. This is understandable since how

to display the capacities of a robot, how to make them concretely apparent and available, in the robot's appearance and conduct, remains an open question. It is even unclear how and whether a robot's capacities can even be defined and/or described *in abstracto*, extracted from the local situation and context in which an action is produced. Furthermore, the principle of transparency might conflict with other more pragmatic concerns of the developers and promoters of social robotics. To encourage people to engage with robots, and ultimately attract buyers, roboticists can rightfully be tempted to overstate or exaggerate their robots' capacities (Parviainen and Coeckelbergh 2021). They may want to emphasise particular features which, in turn, convey such key qualities as interactional competence and likeability, conducive to bonding. This seems to be the case with humanoid robots, which are human-like in their physical appearance, movements and/or voice.<sup>2</sup> However, this is very likely to create expectations of human-like capacities (Eyssel et al. 2011), which in turn might occasion disappointment and/or deception. In other words, while transparency seems a sensible principle for design, neither how it can be applied in practice, nor how roboticists would actually want to apply it, are evident.

To sum up, human participants in their first encounters are left to their own devices to discover how to interact with a particular robot. To avoid deception and disappointment, robot designers recommend that robots' interactional capacities be perceptible in their appearance and conduct; however practical solutions and evidence of engagement in this direction are still lacking. In this paper, we identify some of the methods through which human participants display their expectations, and how they discover a robot's competencies in the course of first encounters. As we move from shaping initial actions to re-producing and transforming actions, we also show the inferences participants make on a moment-to-moment basis, building on the robot's responses, and how participants continuously revise their perception of the robot's competencies. To identify and unpack such emergent and elusive processes, detailed analyses of interactions—the temporal and sequential unfolding of actions, and the quality of embodied conduct—are necessary, which we undertake with systematic analyses of video recordings.

The contribution we aim to make is twofold. First, while the pervasiveness of technologies in all domains of social life have been accompanied by a growing number of studies in research on social interaction (e.g., Arminen et al. 2016), interactions with robots in the wild remain under-investigated in this field. The qualitative insights we provide on embodied interactions with robots build on earlier interests in conversations with machines (e.g., Luff et al. 1990), the 'embodied turn' in conversation analysis (Nevile

<sup>2</sup> But not too human-like either, to avoid the uncanny valley problem, c.f. Mori et al. (2012); Rosenthal-von der Pütten et al. (2014).

2015), and the recent interest in social robotics for qualitative approaches of interactions as they unfold (Fischer 2021). Second, our findings regarding human participants' methods to make the interaction proceed have implications for design regarding current efforts to deploy robots in the wild (Hyuk Park et al. 2020); and they have practical implications for the principle of transparency discussed above. Combining these lines of interests and concerns in the analyses, we hope to encourage an interdisciplinary research agenda that explores and questions the effects of introducing humanoid, interactive robots in various social settings, as well as the roles and responsibilities they might be given.

### 3 Data and methods

Our data are qualitative naturalistic experiments (Heath and Luff 2018), which do not involve experimental conditions, dependent variables, hypotheses, or measurements, as conventional experiments used in most user studies or evaluations of robots typically do. Naturalistic experiments are a convenient method to study human–robot interaction in the wild as the occasion for the human–robot encounter is occasioned by the researchers. More common in the disciplines of human–computer interaction and computer-supported cooperative work than in human–robot interaction, they consist of observing how participants use the technology and how they make sense of its capabilities. They are exploratory, as they seek to explore and identify some un-prespecified foundational issues in these forms of encounters. Whilst the encounter is based on a scenario or script, little or no guidance is given to participants as to what they should do and how they should do it. In this way, the investigations bring to the fore participants' perspectives, the problems they can be seen to encounter and how they seek to resolve them through the sequential unfolding of the interaction. Participants are often in pairs or groups so that they are more likely to make their actions and understandings apparent to their peers. Data can be collected through semi-structured or unstructured interviews, observations, and/or recordings of participants' behaviour. For this study, we collected and analysed video recordings of the interactions only. Since our aim is to focus on how the interactions unfolded, the study does not include interviews with participants or questionnaires.

The recordings give access to phenomena which are otherwise taken for granted and therefore remain unnoticed, or not remarked upon (Tolmie 2011) with a focus on in-depth analysis of instances of behaviour rather than broad categorisation of activities. Detailed, qualitative analysis of naturalistic experiments yields in-depth understanding of participants' practical reasoning as they try to use a system and interpret its various capabilities. Revealing these

seen-but-unnoticed methods, and analysing their implications, not only sheds new light on social organisation and processes, but also has practical value to inform practices and technology design. When the analysis is presented, the focus is not on general summarisation of findings but on particular instances or fragments of activity, often discussed in fine detail.

Our analytic orientation for the video data draws upon ethnomethodology and conversation analysis (Garfinkel 1967; Sidnell and Stivers 2012). This qualitative approach seeks to unpack and understand participants' methods as they proceed in collaborative courses of action and interactions, focusing on language and embodied conduct as the topics of inquiry.

For this study, we draw on two types of naturalistic experiments, and two datasets. The first and principal dataset consists of encounters between a humanoid robot and passers-by in a university hallway (Ben-Youssef et al. 2019). On detecting people approaching, the robot attempted to initiate interaction by greeting them with a handwave and verbal greetings. Passers-by were free to ignore the robot or engage with it. Markings on the floor indicated the robot's proximal space; and a poster on the wall informed passers-by that they were being video recorded for research purposes. The poster also provided a brief description of the study, but it mentioned neither its goal, nor details about the robot itself. After the participants had been involved in the interaction, the robot asked them for their consent to be recorded and for the recordings to be used for research. Participants varied in number, in age, and in reason for walking in the hallway: the data include students, university staff, visitors and families with young children; alone, in pairs or in groups. We extracted from this dataset a collection of instances where participants approached the robot and engaged with it, initially through gestures and hand movements of different kinds.

Our second dataset, which we used as a complement, consists of interactions re-producing a museum visit between a humanoid robot guide called 'TalkTorque' (Yamazaki et al. 2014) and groups of three to four participants. The participants were recruited among the general public, as English speakers in Japan. In a room prepared for the quasi-naturalistic experiment, the robot would comment on objects displayed on a table for the 'visitors'. The robot's behaviour combined talk, body, arm and head movements and followed a script. A wizard-of-oz system allowed for minor variations in the script, such as postponing an action or inserting a turn-at-talk (for example, "Please come closer" if the participants were standing too far away from the objects). The robot mainly not only talked about the objects, but also asked questions such as "Which would you prefer to have in your home?".

Once we had identified robot-recipient design as both a pervasive phenomenon in our data and an understudied topic, we extracted instances of waving gestures and handshakes from our primary dataset. From our second dataset, we extracted embodied references to objects, or pointing

they approach while looking at it. The man is walking ahead of the woman, followed by the child. As they approach, the robot begins to raise its arm, it initiates a verbal greeting (“welcome”, line 1, image 1.1), and waves its arm above its head.

#### Fragment 1

```

01 R bienvenue, ×
      welcome,
      ×Im1.1
02 (0.6) × (1.0) ((mother starts waving, then father))
03 ×Im1.2
04 W ↑cou↓cou:
      ↑hi:↓
05 (1.1) ((Mother walks to robot))
06 W ×bienvenue:
      welco:me
      ×Im1.3
      (0.5)
07 W merci: ×
      thank you:
      ×Im1.4

```



Image 1.1



Image 1.2



Image 1.3



Image 1.4

gestures (Hindmarsh and Heath 2000), which usefully complemented the handwaves and offers to shake hands from the first dataset. We did not categorise the participants in groups. The four authors met regularly to analyse the fragments qualitatively, instance by instance, an established method in the social sciences (e.g., Sidnell 2013). We present the results in this paper, which includes a few representative instances in the form of snippets combining transcriptions of the talk and frame grabs from the videos. The language used in the museum guide robot data is English, and French in the University hallway data, with translations provided on an additional line in the transcripts.

## 4 On some typical gestures in the openings of interactions

Before we present the main findings, it is worth illustrating how some typical gestures can enable human–robot interaction. Fragment 1 is an example of the workings of waving gestures during openings. Waving gestures not only have a typical, recognisable shape and form of movement, they can also accomplish a decisive move in the openings of face-to-face interactions in distant greetings when participants see each other from afar (Kendon and Ferber 1973). In Fragment 1, a family of four—a man and a woman with a young child and a baby carried by the woman—have seen the robot from far away, and

After the robot’s verbal greeting and while it is still waving, the man starts waving in turn, shortly followed by the woman. Their returning gestures potentially complete the greeting sequence. Shortly after, the woman responds verbally with a singsong *coucou*: (“hi”, line 4), a greeting in French which is recurrent in, and particularly suited to, slightly unanticipated mutual visual ‘appearances’ (Licoppe 2017). Then, the woman walks towards the robot, self-selecting as its main co-participant, while the man steps aside. While walking, she produces two additional responses to the robot’s greeting turn (“welcome”, “thank you”, lines 6–7), and then stops in front of the robot, waiting for it to proceed with a next action.

The exchange of waving gestures plays a central part in opening this encounter. As they approach, and even though they know nothing about either the robot’s capacities or what kind of activity it may propose, the two participants treat the waving gesture as a relevant, timely and meaningful greeting action, and an invitation to engage in an interaction. They align to the robot’s conduct firstly by reciprocating the waving gesture, then by engaging in the interaction. The robot’s waving is both a simple gesture and the anchor point from which the participants elaborate to progress the interaction.

Several characteristics—shared with gestures like hand proffers inviting a handshake or pointing gestures—make waving gestures critical resources for human participants in their first encounters with robots. Firstly, with their typical appearance and movement, they can be unproblematically



produced by humanoid robots, as well as recognised by human participants and responded to. Secondly, they make a limited range of responses expectable, and those responses can be similarly straightforward embodied actions. Thirdly, the initiating action and its response form a full sequence which achieves a decisive step in an emerging or ongoing interaction—for instance greetings, or a pointing gesture occasioning a turn of the head toward an object, both allowing the interaction to proceed. Lastly, by completing a sequence, the robot and the participant(s) engage further in the particular course of action in which the sequence is embedded, and which it progresses.

### Fragment 2

```

01 N hel↑lo:↓
02 (2.5)
03 N #il nous en[tend?
    can/does it/he hear us?
    r #raises arm-->
04 R [bonjour,
    hello,
05 (1.2) ((Robot waves))
06 N *[bon↑jour
    he↑llo
    n *waves-->
07 P [bonjour.^
    hello.
    p ^smiles
08 (0.4) ^×(0.2)×(0.2)×(0.6)^(0.6)
    ×Im2.1×Im2.2×Im2.3
    p ^-----^waves

```



Image 2.1



Image 2.2



Image 2.3

recipient design. Fragment 2 involves three participants, of whom we'll consider only two,<sup>3</sup> Ned and Phil, respectively, left and right on the images. The analysis focuses on Phil's actions. As the group approaches from the robot's left, the robot turns its head to them. Phil stops at a short distance, and greets with *hello*: (line 1). For 2.5 s, the robot doesn't respond in any way, which leads Ned to question its capacity to hear: "can/does it/he hear us?"<sup>4</sup> (line 3). While Ned is still speaking, the robot starts to move its arm, says "hello" (line 4), and waves with its hand above its head.

## 5 Robot-recipient design in the production of action

In our data, we observed that participants shaped their gestures in ways that exhibit their expectations regarding the robot's interactional capacities. There seemed to be two methods that participants utilised: they either emphasised a gesture in a particular way or sought to follow or align with the robot's emergent conduct. These seemed markedly different from how one might engage with 'ordinary' co-participants.

### 5.1 Emphasising gesture

An additional characteristic of the type of simple gestures we focus on is that their standard shape lends itself well to variations, slight transformations of the gestures through which participants can adapt to various contingencies. Such variations in the quality of the gesture can be revealing in terms of

Ned waves, and then he and Phil answer verbally at the same time with "bonjour" (hello, lines 6–7). After his verbal greeting, Phil reciprocates the waving gesture in turn. With respect to both the robot's initial gesture and Ned's responding gesture, Phil's waving comes late, and it seems encouraged by Ned's prior waving. The delay seems to question the very need to reciprocate the waving, thus displaying uncertainty as to whether the robot can perceive embodied actions at all, in line with Phil questioning its ability to hear (line 3). More importantly, the conduct is shaped in a particular way. Phil makes his hand larger and particularly visible by spreading his fingers; he makes two broad movements from left to right; and he produces these movements in the space

<sup>3</sup> The third participant walks before Ned and Phil, she stands on Phil's right, out of the frame grabs.

<sup>4</sup> The utterance in French can be translated in several ways, modal verb or not, and the pronoun "il" can be neutral or male.

in front of him (Images 2.1, 2.2 and 2.3) instead of above or next to his head, where most waving gestures are produced (Kendon and Ferber 1973). The broad movement, the spatial positioning, and the shape of the hand, seem to aim to make the gesture more prominent, to emphasise or highlight it so that it is more perceptible for the robot than an ordinary

### Fragment 3

```
01 R bonjour,
    hello,
02   (0.3)
03 B °bonjour.°
    hello.
04   (1.2)^(0.6)^(0.2)×(0.2)×(0.2)×
05 r   *raises arm-->
06 b   ^looks at robot's moving arm
07   ^raises arm-->
      ×Im3.1×Im3.2×Im3.3
```



Image 3.1

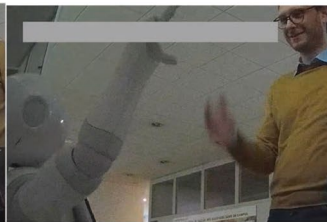


Image 3.2

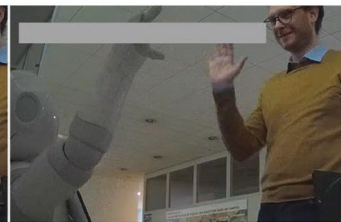


Image 3.3

waving gesture. These qualities exhibit Phil's expectation or assumption that the robot has a limited ability to perceive a waving gesture, perhaps even gestures in general. Such ways of emphasising gestures for robots—waving gestures and others—are pervasive in our data.

This fragment also demonstrates that participants' expectations and understandings of the robot's capacities are particularly influenced by what happened just before; and that they evolve in time, even in short periods of time such as these very few seconds. They seem to infer from the robot's delay in responding to Ned's greeting at the very beginning as evidence that it has limited capacities (as Ned's question on line 3 suggests). They make a subsequent move after the robot has produced a response shortly after, and yet Phil's emphatic waving gesture still addresses a recipient with limited capacities. In other words, participants' understanding of what actions can be produced for the robot evolves with every single step of the interaction, each move the robot makes or does not make, and when exactly. They appear particularly sensitive to what happened just before in the interaction, a kind of phenomenon that is not likely to arise in post hoc interviews with participants.

## 5.2 Following and aligning to the robot's emerging conduct

Matters of timeliness and sequentiality become particularly apparent when participants adapt their actions to the robot's actions on a moment-by-moment level by closely monitoring

its actions. In the following fragment, while Ben is quietly standing in front of the robot, the robot initiates greetings with *bonjour* ("hello", line 1), and Ben reciprocates with a similar *bonjour* (line 2).

A silence follows the exchange of verbal greetings. Ben would reasonably expect the robot to make the next move as he does not have any information about its capacities and the activity the robot may propose. After a 1.2-s silence, the robot starts moving its arm. Ben immediately turns his head to look at the moving hand, and starts moving his right hand in turn. The shape of Ben's hand at this point projects a handshake: the main fingers are held together, the thumb separated, the palm open, and the hand slightly thrust towards the robot (Image 3.1). Within the next tenths of seconds, as the robot continues to raise its arm bringing its hand above its head, Ben changes the shape of his hand by opening up his fingers; and he changes the trajectory of his arm from in front of him to above him (Image 3.2). Thus, his projected handshake pivots (Lerner and Raymond 2017) into a different gesture: as the robot starts waving, Ben has taken his hand on his side and holds it as a static greeting gesture (Image 3.3), a greeting gesture close to waving.

Ben visually monitors the robot's conduct carefully enough to, firstly, initiate his own response early on, and, secondly, revise his projected action as soon as he gets a different understanding of the robot's emerging conduct. He mirrors and adapts to the robot's emerging conduct, he also times and tailors his actions for this robot as it is behaving there and then. Thus, robot-recipient design is embedded in a local, emergent sequence of actions. It seeks to facilitate the robot's work in pursuing its course of action, and to encourage it to proceed. In this case, it exhibits relatively

low expectations regarding the robot's capacities. Ben orients to the robot as a not fully competent co-participant, for the least not an ordinary one.

So far, we have outlined a set of practices characteristic of robot-recipient design where each seem to aim to enhance the recognisability of the gesture. First, participants can bring the body part in question closer to the robot, within what they assume to be its proximal space, in an attempt to facilitate the robot's detection and recognition of the gesture (Mead and Mataric 2016). Secondly, participants can make the particular body part more visible, larger, simply by spreading fingers for the robot to better detect the hand. Thirdly, they can shape their gesture so that it resembles, if not mirrors, the robot's prior or ongoing action, potentially to 'fit' into the range of actions the robot expects next. Finally, participants can emphasise the definitional features of those gestures, for example by expanding the movement of waving.

Whilst we have identified these practices by focusing the analysis on isolated sequences, or paired actions, the encounters with the robots span over longer stretches. They are composed of series of meaningful sequences, and the lived experience of the interaction as 'successful' depends on this continuity and progressivity. In our data, participants frequently encountered problems following the greetings, and yet they rarely just turned their back to the robot. We

looked into the methods and resources they rely on then to try to move the interaction forward anyway.

## 6 Remedial actions in the face of robots' lack of response

There are many cases when, for the participants, the robot does not respond as expected or at all. In these cases, participants often pursue a relevant response by re-producing their action, by transforming it, or even by exploring the robot's 'body'. They do so in ways that display their evolving understanding of, and attempt to continuously adapt to, the robot's competences.

### 6.1 Pursuing a response with a recipient-design hand proffer

As an example, the following fragment occurs after a first successful sequence between the robot and the participants. In their path down the hallway, Edward and Franck approached the robot. The latter responded to their greeting in a way that they both treated as adequate and timely. This initial sequence, thus, provisionally enacted the robot as a potential interactional partner. Fragment 4 starts as Edward walks away from the robot and Franck moves closer.

#### Fragment 4

```

01      × (1.1)
        × Im4.1
02 F   tu serres la main après?×
        you shake hands afterwards?
        × Im4.2
03      (1.0)
04 F   °(XXX XXXX)° × (.) bonjour:r#
        °(XXX XXXX)° (.) hello:
        × Im4.3
        r                                     #head and gaze down
05      (1.0)#×(1.2)#(0.3)×
        r   #-----#head follows waving hand
        × Im4.4      × Im4.5

```

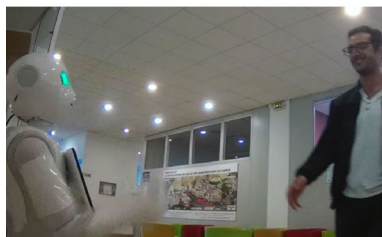


Image 4.1

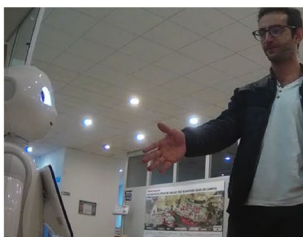


Image 4.2



Image 4.3



Image 4.4

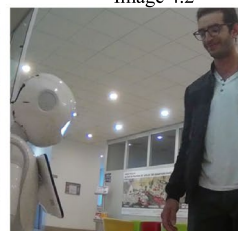


Image 4.5



Franck starts extending his arm (Image 4.1) and produces a yes/no question “you shake hands afterwards?” (line 2). In the course of this question, his arm stops in a ‘hand proffer’ position (Image 4.2). The utterance on line 2, whilst designed as if it were addressing the robot, is intelligible as a formulation made available to Edward of what he is currently doing, that is, exploring further the interactional ability of the robot. It depicts a step forward in this exploration both in the temporal sense, as it comes after the exchange of greetings; and in a normative sense, with the word “after” marking that a handshake would be an expected next action in such a greeting sequence. He thus orients to the robot being potentially endowed with some interactional abilities. This is made possible by what the robot did just before, that is, returning a verbal greeting. At the same time, with the utterance on line 2 formatted as a question, Franck literally questions the robot’s interactional competence and thereby does not orient to the robot as a fully competent co-participant.

The hand proffer displays a particular and interesting form of recipient design: the hand is extended and held vertical and flat, with the fingers separate. While reminiscent of the shape of Phil’s hand in the waving gesture in Fragment 3, also with fingers apart which aimed to make the hand more visible, this hand shape does something more in the case of an offer to shake hands. Let us consider for a moment how a handshake is ordinarily done. A handshake is a collaborative accomplishment, in which the hands of both participants approach and mutually and continuously adjust their shape, in anticipation of the upcoming shake, and in the actual shake itself. A handshake is an instance of what Merleau-Ponty called intercorporeality (Merleau-Ponty 1964). First, when social action is achieved in this way, agencies are blurred, in the sense that one cannot distinguish a hand that shakes and a hand that is shaken: “*The reason why I have evidence of the other man’s being there when I shake his hand is that his hand is substituted for my left hand, and my body annexes the body of another person in that “sort of reflection” it is paradoxically the seat*

*of. My two hands “coexist” or are “compresent” because they are one single body’s hands. The other person appears through an extension of that compresence; he and I are like organs of a single intercorporeality”* (Merleau-Ponty 1964). Intercorporeality provides for an embodied world, known in common (Meyer et al. 2017). Second, as a temporal accomplishment the handshake is not sequential, in the sense that it would be made of recognizable discrete units of action and projected discrete responses: ‘responsive’ adjustments are continuous and mutual.

Franck’s hand proffer here is very different. In part because the robot does not display coordinated responsive behaviour, it is designed as a single, discrete unit of behaviour, to be recognised as such, and as projecting a next action. Indeed, facing a lack of response, Franck appears to pursue a response by other means, such as the wave in line 5, Image 4.4. And in that respect, the way the gesture is done is significant. On the one hand, the flat hand highlights the lack of continuously coordinated embodied response from the robot. Had there been any, the hand would have been approached in a continuously evolving grasping shape. Additionally, it provides a schematised, stylised configuration of the hand, maximising the recognisability of the gesture as a discrete unit of embodied behaviour sequentially projecting a responsive shake. This is what we call here the ‘hand proffer’. As a recipient-designed piece of embodied conduct, the hand proffer provides an opportunity for the robot to display interactional competence—recognising it for what it is and providing the projected next action, shaking—on one level, while on another level highlighting the robot’s inability to produce continuous embodied adjustments, and thus placing a boundary on its interactional competence.

## 6.2 Re-designing a handshake in the course of its production

Let us consider Fragment 5 where there is another, different attempt at shaking hands, this time in a closing sequence.

Fragment 5

```

01          #*(0.2)
           #Im5.1
02 s      *extends arm-->
03 S      est-ce que je peux te serrer# la main#
           may I shake your hand

```



```

05          (2.2)* (0.8)#(1.0)#(0.6)* (0.5)#(0.5)*#
06 s          *rotates hand
07          #Im5.4
08          #Im5.5
09          *waggles robot's fingers
10          #Im5.6
11          #Im5.7

```



After a pause in the interaction, Sam starts to extend his arm, and then requests permission to shake hands with the robot: “May I shake your hand” (line 3). His hand is positioned into a hand proffer co-extensively with the verbal turn. This combination of talk and gesture to initiate pre-closings displays an uncertainty regarding the robot’s ability to produce an adequate response, i.e., a responsive handshake. First, asking for permission orients to potential reciprocity issues. By explicitly formulating the relevant activity as a handshake, it suggests a concern with the robot’s capacity to visually recognise the gesture for what it means. As to the gesture itself (Image 5.2), it displays a similar exaggerated flat hand as in Fragment 4, highlighting recognisability in one of the ways identified in the previous section. In other words, rather than involving continuous adjustment, the offer to shake hands is done as a sequentially

implicative, self-contained move, the hand proffer. This suggests a difficulty for Sam to anticipate the robot’s response, and to “take the attitude of the other”, as Mead’s social psychology would have it (Mead 1967 [1934]). Still, with this hand proffer, Sam offers the robot the opportunity to shake hands. Should the robot follow suit, its competence to produce a proper response (which involves responding to the offer, understanding the handshake, initiating it, etc.) could be ‘discovered’ in and through interaction.

As the verbal utterance unfolds, the robot raises its left hand (Image 5.3). Though it is fortuitous, and though this is the wrong hand for shaking, this can be understood as a relevant response to an offer to shake hands, and Sam seems to understand it in this way. First, he holds the hand proffer for 2.5 s, thus providing an extensive slot for the robot to respond to it. Then, lacking any kind of response, he rotates

his hand and brings it close to the robot's left hand, thus projecting to grasp it (Image 5.4).

Sam then holds his rotated right hand for a second (line 6), giving the robot a relatively brief time to grasp it in return. This suggests mounting doubts regarding the robot's capacity to adjust intercorporeally to the offer. Furthermore, Sam pre-empts the robot's response by initiating a kind of handshake by himself. He grasps the robot's fingers (Image 5.5) to form a kind of partial handshake grasp, and then slightly raises his hand, as if trying to shake. Then he releases his initial grip and waggles his fingers on the robot's (Image 5.6), and eventually removes his hand (Image 5.7).

The subtle embodied interactional work which emerges in this short sequence exhibits a dynamically evolving stance towards the robot. In a handshake, both parties have to continuously adjust their shake to what they mutually experience through their hands.<sup>5</sup> Such an adjustment involves haptic responses so smoothly attuned as to blur individual agencies and achieve a tactile intercorporeality. When Sam grasps the robot's fingers and moves their joint hands up, he provides an opportunity for the robot to show a kind of finely tuned, responsive haptic coordination. His release and reconfiguration of the grasp may then be understood as a form of action pivot (Lerner and Raymond 2017) involving another form of coordination. In changing his grasp and waggling the fingers, Sam has recipient-designed a move which no longer projects a handshake, trying instead for any kind of haptic reaction to the waggling at the level of the robot's fingers. We may note that this does not involve intercorporeality anymore, but a stimulus–response haptic organisation where agencies are sharply differentiated: the participant is pressing, and testing for some reaction on the part of the 'pressed' robot's hand, casting the robot as a more passive participant. Sam's eventual disengagement becomes accountable as an orientation towards trouble: his first attempt to get the robot to shake his hand failed, his second attempt self-repairing his gesture also failed to elicit any kind of haptic response. Sam's disengagement projects a change of interactional project.

To sum up, what we observe here is a particular way for the participant to address interactional troubles in the human–robot encounter. Trouble here is the lack of what

could be deemed a proper response from the robot, as Sam successively re-produces and reconfigures his initial actions. Instead of making the robot accountable, Sam revises his moves and recipient-designs them so as to make various forms of responses possible on the part of the robot. Each of the successive attempts solicits different levels of competence. The series of reconfigurations and targeted responses displays a kind of hierarchical orientation, from more elaborate to increasingly simple ones, thus enacting the robot as an agent with diminishing interactional competence: initiating a handshake (enacting the robot as a potential, intercorporeal, shaker), and testing for haptic sensitivity (enacting the robot as a machine with or without tactile capacities). This way of making interactional troubles perceptible and attempt to resolve them displays both an uncertainty with respect to the robot's competences, and an orientation to them as 'discoverables'.

## 7 Implications for other gestures in human–robot interaction

Our analyses mainly focused on handwaves and hand profers inviting handshakes, two standard gestures recurrent in the openings and closings of interactions. The issues we address with respect to robot–recipient design—that those gestures can be emphasised, slightly transformed, mirrored, and overall that they are key resources to go through the first steps of an interaction—are relevant to many other embodied actions. Head nods and pointing gestures, for instance, are extensively used in robotic experiments because of the same features. Let us consider Fragment 6 as an example. As mentioned earlier, our quasi-naturalistic experiments with a museum guide robot involves pointing gestures. Following the robot's question: "Out of these designs, which would you prefer to have in your home?", participants can be expected to point towards one of the objects of the exhibition. In the following fragment, even though the robot's head is not turned towards Sandra (the participant in the white jacket), she self-selects to answer: she looks at and points to one of the objects (Image 6.1) and initiates a verbal response: "I: like this one." (line 1).

<sup>5</sup> See the review of handshakes in human–robot interaction by Prasad et al. 2020. Distinguishing different phases (reaching phase, grasping phase, shaking phase) can be very useful but obviously does not apply in this case where the robot fails to respond to the invitation to shake hands.

### Fragment 6

01 S \*I: like this\* one.  
 ×Im.6.1  
 02 S ×it can hold sandwiches.\* hhhe×  
 ×Im.6.2 ×Im.6.3 ×Im.6.4



Image 6.1



Image 6.2



Image 6.3



Image 6.4

Sandra makes a first attempt to obtain the recipient's gaze (Goodwin 1980) by elongating the vowel on "I" at the beginning of her answer. Once her answer is complete, she turns her head to the robot (Image 6.2); and she maintains her head in this direction as she expands her answer ("It can hold sandwiches.", line 2), as if seeking further to obtain the robot's attention. Indeed, sustaining her head towards the robot displays the expectation that the robot turn towards the object she is pointing at and thereby understands which object "this one" refers to. The robot does not move. Sandra briefly looks at the object again while retrieving her pointing (Image 6.3), and, still expecting and pursuing some form of response from the robot, she turns towards it again while producing an exhaled laughter ("hhhe", line 2, Image 6.4).

Like in the above extracts, she demonstrably expects a certain level of interactional competence on the part of the robot: that it be able to recognise this pointing gesture as such, by turning to the object. By repeatedly turning her head towards the robot, while pointing to the object and after for lack of a response, she also gives the robot several opportunities to respond, showing that she would take even a late response as an appropriate one.

With this fragment, we also want to highlight a simple and yet perhaps overlooked phenomenon: participants expect the robot to be able to understand an action which its prior action (especially a first pair part like the question in Fragment 6) makes relevant. That is, an answer to the robot's question "Out of these designs, which would you prefer to have in your home?" is very likely to involve pointing towards one of the objects and, therefore, the robot should be able to follow the pointing gesture and turn towards the object. We come back to this point in the following discussion.

## 8 Discussion

In this paper, we have outlined a set of practices characteristic of robot-recipient design, whereby participants seem to aim to enhance the recognisability of their gesture: bringing

the gesturing body part closer to the robot; making the body part more visible, larger; shaping the gesture so that it resembles, if not mirrors, the robot's prior or ongoing action; and emphasising the definitional features of a particular gesture, such as extending the breadth of a waving movement. If one action fails, that is, is not taken up, participants readily reproduce it in a different form, re-design it in case the robot can recognise the second. Or, they can initiate a new action and change interactional trajectory, again in the eventuality that the robot may then follow suit.

These practices are part of what we refer to as 'robot-recipient design', not only the recipient design production of isolated actions, but also the ways in which the production and reproduction of actions exhibit participants' continuous, moment-by-moment assessment of robots' interactional competencies. We believe it has a number of consequences for both how we understand human-robot interaction and how we can design the robot's contribution.

Firstly, our approach differs from most HRI research relying on questionnaires, post hoc evaluations of interactions based on subjective experience, or quantitative analyses of video-recorded interactions. Using quasi-naturalistic experiments and focusing on a limited number of single instances of interaction, we focus on the quality of actions and their sequential organisation, from the participants' perspective. The ethnomethodological, conversation-analytic approach lets us take into consideration what is accomplished in interaction on a moment-by-moment basis. It sheds light, for instance, on participants' practical reasoning as they shift back and forth in their attribution of competence to the robot, depending on the action it produces and when in the course of the interaction. It also reveals aspects of interaction that cannot emerge through interviews, such as participants' sensitivity to what happened just before in the interaction.

Interactional competence can thus be defined as the capacity to produce a timely first or responsive action, and thus be considered at a merely technical level. By 'technical', we do not mean 'mechanistic', we refer rather to



an approach that focuses on what is exhibited and publicly available. Interactional competence is a concrete and explorable characteristic of a robot, through robot–recipient design and probably other phenomena. Analysing a selection of cases from our video data and providing a transcript for each, we hope to have shown that participants’ evolving perceptions of the robot’s competencies are observable in their actions. We showed in our analyses that close attention to both the quality of actions and their sequential organisation was crucial. In Fragment 2, Phil indeed “waves” to the robot, but his attribution of (lesser) competence to the robot is apparent in the particular shaping and positioning of his hand, and the delay in the gesture compared to usual paired waving gestures. The quality of embodied actions and their sequential organisation are two dimensions; participants can play with in their endeavour to make the interaction progress. Ultimately, it is indeed the progression of interaction that shows that robot and participants are jointly engaged in an interaction in which the other is actively participating, and therefore competent to do so.

With this approach, we also argue for a distinction between notions that are instantiated in concrete phenomena (such as recipient design or interactional competence the way we approached them in this paper) and more abstract or composite notions commonly studied in HRI, such as robots’ likeability, their ability to provide a sense of familiarity, or trust. The latter tend to require interpreting a mix of observable and/or measurable facts and participants’ subjective experiences. For example, a number of studies take “anthropomorphism” or “anthropomorphising behaviour” as a starting point to study how human participants consider a robot, through how they interact with it (e.g., Eyssel et al. 2011; Fink 2012; Fischer 2021; Lemaignan et al. 2014; Salem et al. 2013). In fact, when they set about interacting with a robot, participants can but use the same resources and methods they would with a fellow human. Does this mean that they anthropomorphise the robot? More precisely, participants can, and they probably do, dissociate the ability to engage in basic interactional sequences, from the capacity to understand complex social actions (with several layers of meaning, such as humour, irony, offence, etc.) and their implications for interpersonal relationships, or even more to feel emotions or be endowed with a personality and moral rights. Therefore, participants can produce similar actions as they would with fellow humans and yet expect a different sort of uptake or response from a robot than they would from a fellow human, and different long-term consequences.

Thus, this approach bridges the divide between ‘subjective assessment’ based on questionnaires or the like, and ‘objective assessments’ based on quantitative measures (e.g., Salem et al. 2013): no tool or measure is imposed by the researcher, and it relies on objective facts, whilst taking participants’ perspective. It allows to unpack both meaningful

and objective social phenomena, because they are available out there and not located in the mind, even though they cannot be counted and computed.

Throughout our analyses, we have shown that human participants constantly monitor robots’ conduct, to assess and potentially adapt to their actions. The data presented here suggest that participants can have two different—overlapping and on a continuum rather than mutually exclusive—a priori attitudes towards the robot, which largely influence the trajectory of the encounter: exploring and adapting (fragments 1, 2, 3 and 5 typically), or testing (Fragment 4). In exploring and adapting, participants appear curious about, and therefore explore, what the robot is capable of, and they appear more tolerant and persistent when the robot fails to act or to respond appropriately and in a timely fashion. They would probably judge a human fellow similarly failing to act or respond in this way as rude, inattentive or unavailable. Here participants would wait, and/or reproduce and reshape their previous action. Participants also accept a broad range of action types as relevant or conform, and make sense of the robot’s actions according to their sequential environment. In testing, participants appear less curious about the robot, they test its ability to recognise an action and produce an appropriate response; and if it does not respond appropriately and immediately, they turn away from the robot. Although our study does not allow to make any general assumptions and even less draw conclusions in this regard, we can hypothesise that different types of audiences may tend toward one type of attitude more than the other. In any case, these a priori attitudes largely influence the trajectory of the encounter in how much effort participants will put in making the interaction work, and they are exhibited in robot–recipient design.

We acknowledge that our study has a few limitations due to biases in our data and analytic choices. First, our participants were probably particularly cooperative and benevolent towards the robot. The participants in our main dataset, people walking in a university hallway, may form a particular audience which cannot be conflated with ‘the general public’. Besides, we extracted and analysed only the interactions with passers-by who chose to stop by the robot and engage, while many did not. Lastly, we can hypothesise that once the participants who chose to approach the robot had read the signs explaining that the encounter was video recorded for research purposes, they felt more committed to persevering to make the interaction with the robot work. In our second dataset, the participants may have felt somehow obliged to act as good research participants because of the quasi-experimental situation. On the other hand, in all, our participants also largely varied in terms of age and socio-economic background, so that it is hard to speculate on other biases. In any case, even though these characteristics of our data and approach can be considered as limitations



in conventional HRI research, they do not distort qualitative analyses of single, short instances of interactions, and therefore do not undermine the reliability and robustness of the results. Our analyses reveal examples of problems participants can encounter with robots, their assumptions and the methods they can use to manage interactions with robots.

## 9 Conclusion

Recipient design is integral to our everyday interactions. With fellow humans, people take for granted that any prospective co-participant is a competent user of language, be it verbal or bodily. On encountering a robot for the first time, people do not take these competencies for granted, and therefore, they face particular problems to produce and shape their own actions for a recipient.

We hope to have shown that the study of recipient design in human–robot interaction through sequential analysis of embodied action is novel, and interesting in several respects. First, such an analysis explores the initial assumptions of participants, their *a priori* expectations regarding the competencies of the robot as a co-participant. The more uncertain the user, the more recipient design and repair are connected: recipient design can be understood as an attempt to pre-empt potential trouble, and repair exhibits the actual occurrence of trouble. Second, this focus shows that recipient design is tentative and exploratory, as participants can be seen to revise their initial assumptions and to generate new understandings of the robot’s competencies at any moment. Third, recipient design appears as local, contingent, and revisable, grounded in the immediate interactional environment; and it does not follow an upward or downward trajectory, towards a higher or lower level of competence, as one might expect.<sup>6</sup> These aspects of recipient design are particularly revealing in actions that seek to elicit a response from the robot, as well as when an initial action fails to obtain an appropriate response from the robot and is then re-produced, reshaped, or transformed. Thus, the robot’s status as an interactional partner can be seen to change from moment-to-moment in the course of an interaction (on a continuum from competent to incompetent), in brief sequences of actions, depending on how smooth the interaction turns out to be.

A core component of robot–recipient design is human participants’ sustained and careful attention to the robot’s just previous, emerging, and expected conduct, at every moment of the interaction. In the study, participants seemed

to produce actions that reflected low expectation of the robot’s competence. We identified some of the ways participants shape their embodied actions to maximise their recognisability for the robot, and reproduce earlier, failing actions in a new shape. Robot–recipient design is not a set of general rules that participants apply when they seek to interact with a robot, it is a way of evaluating the robot’s actions in the light of the sequential progression of the interaction, and shapes how they adapt their own conduct accordingly. Participants pre-suppose that their actions are organised sequentially, and so they expect the robot to also organise its conduct likewise.

There are implications to this. First, the examples we have shown of gestures being reshaped and transformed, and how these transformations are produced, may be useful in helping designers distinguish first from ‘repairing’ actions, or to emphasise that repeating actions are done in different ways and may need to be taken into account in seeking to understand an action in context. Second, our findings point to concrete solutions to the principle of transparency according to which robots should display in their design—their appearance and embodied conduct—what they are capable of, and no more (Baillie et al. 2019; Malle et al. 2021; Złotowski et al. 2020). Whilst interactional capacities may be difficult to display through static aspects of appearance, participants can infer those capacities from robots’ actions and what they project or make relevant. First, participants expect a robot to be able to perceive and understand the type of actions that it has produced itself. Second, they expect the robot to be able to understand actions which a prior action on its part has made relevant—be it the answer to a question, a returned greeting, etc. As we showed in particular with Fragment 6, a robot should not produce actions that make relevant next actions it will not be able to respond to. For example, if the robot cannot recognise a pointing gesture, it should not ask a question that makes relevant an answer including a pointing gesture to refer to an object. Thus, our basic design recommendations are that a robot should (1) produce actions which it is itself capable of understanding; and (2) for which it is capable of processing the range of possible next actions. Whilst constraining, these ‘rules’ are more specific and applicable than what HRI research has proposed so far. In sum, we are suggesting an interactional perspective on transparency. It is not just a property of a robot or a robot’s actions but one that is tied to prior conduct and also to what is expected to follow.

The openings of interactions not only provide users with their first impressions of what the robot can do, impressions that shape their future conduct and attitudes. They also provide an insight about how people understand the robot, the assumptions they have about it, whether this is from its appearance or its initial movements. We suggest that taking these first moments seriously and scripting them

<sup>6</sup> As Fischer notes: “we have found that participants respond to the robot’s actions with anthropomorphizing behaviours, such as reciprocating a greeting, offering help, asking what it wants etc., in one moment, but that in the next moment, they may also treat it like a machine” (2021: 4:23), and that is why it is important to get to grips with how it comes to be the case.

in detail can not only reveal the complexities of seemingly simple mundane actions like waving and hand shaking, but also suggest ways in which we can start to analyse the ways humans engage with robots not just through talk, but through the embodied actions both produce in the local environment.

**Acknowledgements** This research was supported by King's Business School (King's College, London) as part of an Incubator project.

**Data availability statement** The datasets analysed during the current study are not publicly available because of the agreement made with the video-recorded participants, but are available from the corresponding author on reasonable request.

## Declarations

**Conflict of interest** The authors have neither relevant financial or non-financial interests to disclose, nor competing interests to declare that are relevant to the content of this article. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no financial or proprietary interests in any material discussed in this article.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Arminen I, Licoppe C, Spagnolli A (2016) Respecifying mediated interaction. *Res Lang Soc Interact* 49(4):290–309
- Baddoura R, Venture G (2013) Social vs. useful HRI: experiencing the familiar, perceiving the robot as a sociable partner and responding to its actions. *Int J Soc Robot* 5(4):529–547
- Baddoura R, Venture G (2015) This robot is sociable: Close-up on the gestures and motion of a human responding to a proactive robot. *Int J Soc Robot* 7:489–496
- Baillie L, Breazeal C, Denman P, Foster ME, Fischer K, Cauchard JE (2019) The challenges of working on social robots that collaborate with people. In: CHI conference on human factors in computing systems extended abstracts (CHI'19 extended abstracts)
- Ben-Youssef A, Varni G, Essid S, Clavel C (2019) On-the-fly detection of user engagement decrease in spontaneous human-robot interaction using recurrent and deep neural networks. *Int J Soc Robot* 11:815–828
- Bergmann K, Eyssel F, Kopp S (2012) A second chance to make a first impression? How appearance and nonverbal behavior affect perceived warmth and competence of virtual agents over time. In: Nakano Y, Neff M, Paiva A, Walker M (eds) *Intelligent virtual agents*, vol 7502. Lecture notes in computer science. Springer, Berlin Heidelberg, pp 126–138
- Button G, Sharrock W (1995) On simulacrum of conversation: toward a clarification of the relevance of conversation analysis for human-computer interaction. In: Thomas PJ (ed) *The social and interactional dimensions of human-computer interfaces*. Cambridge series on human-computer interaction. Cambridge University Press, New York, pp 107–125
- Cafaro A, Vilhjálmsdóttir HH, Bickmore T (2016) First impressions in human-agent virtual encounters. *ACM Trans Comput Hum Interact* 23(4):1–40 (article 24)
- de Graaf M, Ben Allouch S, van Dijk J (2017) Why do they refuse to use my robot?: reasons for non-use derived from a long-term home study. In: *Proceedings of the ACM/IEEE international conference on human-robot interaction*, p 224–233
- Deppermann A (2015) When recipient design fails: egocentric turn-design of instructions in driving school lessons leading to breakdowns of intersubjectivity. *Gesprächsforschung* 16:63–101
- Dourish P (2006) Implications for design. In *Proceedings of the SIGCHI conference on human factors in computing systems*, p 541–550
- Drew P (1997) 'Open' class repair initiators in response to sequential sources of troubles in conversation. *J Pragmat* 28(1):69–101
- Eyssel F, Kuchenbrandt D, Bobinger S (2011) Effects of anticipated human-robot interaction and predictability of robot behavior on perceptions of anthropomorphism. In: *Proceedings of the ACM/IEEE international conference on human-robot interaction*, p 61–67
- Fink J (2012) Anthropomorphism and human likeness in the design of robots and human-robot interaction. *Soc Robot* 7621:199–208
- Fischer K (2010) Why it is interesting to investigate how people talk to computers and robots: introduction to the special issue. *J Pragmat* 42(9):2349–2354
- Fischer K (2021) Tracking anthropomorphizing behaviour in human-robot interaction. *ACM Trans Hum Robot Interact* 11(1):1–28 (Article 4)
- Fischer K, Lohan K, Foth K (2012) Levels of embodiment: linguistic analyses of factors influencing HRI. In: *Proceedings of the ACM/IEEE international conference on human-robot interaction*, p 463–472
- Garfinkel H (1967) *Studies in ethnomethodology*. Prentice-Hall, Englewood Cliffs
- Gehle R, Pitsch K, Dankert T, Wrede S (2017) How to open an interaction between robot and museum visitor? Strategies to establish a focused encounter. In: *Proceedings of the ACM/IEEE international conference on human-robot interaction*, p 187–195.
- Heath C, Luff P (2018) The naturalistic experiment: video and organizational interaction. *Organ Res Methods* 21(2):466–488
- Hindmarsh J, Heath C (2000) Sharing the tools of the trade. *J Contemp Ethnogr* 29(5):523–562
- Hyuk Park C, Ros R, Kwak SS, Huang CM, Lemaignan S (2020) Editorial: Towards real world impacts: design, development, and deployment of social robots in the wild. *Front Robot AI* 7:600830
- Kendon A, Ferber A (1973) A description of some human greetings. In: Kendon A (ed) *Conducting interaction. Patterns of behavior in focused encounters*. Cambridge University Press, Cambridge, pp 153–207
- Kontogiorgos D, van Waveren S, Wallberg O, Pereira A, Leite I, Gustafsson J (2020) Embodiment effects in interactions with failing robots. In: *Proceedings of the CHI conference on human factors in computing systems*, p 1–14
- Koshik I (2005) Alternative questions used in conversational repair. *Discourse Stud* 7(2):193–211
- Kuzuoka H, Suzuki Y, Yamashita J, Yamazaki K (2010) Reconfiguring spatial formation arrangement by robot body orientation. In: *Proceedings of the 5th ACM/IEEE international conference on human-robot interaction*, p 285–292

- Lemaignan S, Fink J, Dillenbourg P (2014) The dynamics of anthropomorphism in robotics. In: Proceedings of HRI'14, March 3–6, Bielefeld, Germany, p 226–227
- Lerner GH, Raymond G (2017) On the practical re-intentionalization of body behavior: action pivots in the progressive realization of embodied conduct. In: Raymond G, Lerner GH, Heritage J (eds) *Enabling Human Conduct: Studies of talk-in-interaction in honor of Emanuel A. Schegloff*. John Benjamins Publishing, Amsterdam/Philadelphia, pp 299–313
- Licoppe C (2017) Skype appearances, multiple greetings and 'coucou': the sequential organization of video-mediated openings. *Pragmatics* 27(3):351–386
- Luff P, Gilbert Frohlich GND (eds) (1990) *Computers and conversation*. Academic Press, London
- Malle BF, Fischer K, Young JE, Moon A, Collins E (2021) Trust and the discrepancy between expectations and actual capabilities of social robots. In: Zhang D, Wei B (eds) *Human–robot interaction: Control, analysis, and design*. Cambridge Scholars Publishing, New York
- Mead GH (1967 [1934]) *Mind, self and society*. University of Chicago Press, Chicago
- Mead R, Mataric M (2016) Robots have needs too: How and why people adapt their proxemic behavior to improve robot social signal understanding. *J Hum Robot Interact* 5(2):48–68
- Merleau-Ponty M (1964) The primacy of perception. In: Edie J (ed) *Northwestern University Press*, Evanston
- Meyer C, Streeck J, Jordan JS (eds) (2017) *Intercorporeality: emerging socialities in interaction*. Oxford University Press, Oxford
- Mori M, MacDorman KF, Kageki N (2012) The uncanny valley. *IEEE Robot Autom Mag* 19(2):98–100
- Neville M (2015) The embodied turn in research on language and social interaction. *Res Lang Soc Interact* 48(2):121–151
- Oloff F (2018) “Sorry?”/“Como?”/“Was?": open class and embodied repair initiators in international workplace interactions. *J Pragmat* 126:29–51
- Paetzel M, Perugia G, Castellano G (2020) The persistence of first impressions: the effect of repeated interactions on the perception of a social robot. In: Proceedings of the ACM/IEEE international conference on human–robot interaction, p 73–82
- Parviainen J, Coeckelbergh M (2021) The political choreography of the Sophia robot: beyond robot rights and citizenship to political performances for the social robotics market. *AI Soc* 36:715–724
- Pelikan H, Broth M (2016) Why that Nao? How humans adapt to a conventional humanoid robot in taking turns-at-talk. In: Proceedings of the CHI conference on human factors in computing systems, p 4921–4932
- Porcheron M, Fischer JE, Reeves S, Sharples S (2018) Voice interfaces in everyday life. In: Proceedings of the CHI conference on human factors in computing systems, Paper 640
- Prasad V, Stock-Homburg R, Peters J (2020) Human-robot handshaking: a review. *Int J Soc Robot* 14(1):277–293
- Rosenthal-von der Pütten A, Krämer NC, Becker-Asano C, Ogawa K, Nishio S, Ishiguro H (2014) The uncanny in the wild. analysis of unscripted human-android interaction in the field. *Int J Soc Robot* 6:67–83
- Ruijten PAM, Haans A, Ham J, Midden CJH (2019) Perceived human-likeness of social robots: testing the Rasch model as a method for measuring anthropomorphism. *Int J Soc Robot* 11(3):477–494
- Sacks H, Schegloff EA, Jefferson G (1974) A simplest systematics for the organization of turn-taking for conversation. *Language* 50(4):696–735
- Salem M, Eyssel F, Rohlfing KJ, Kopp S, Joubin F (2013) To err is human(-like): effects of robot gesture on perceived anthropomorphism and likability. *Int J Soc Robot* 5(3):313–323
- Schegloff EA (1992) Repair after next turn: the last structurally provided defense of intersubjectivity in conversation. *Am J Sociol* 97(5):1295–1345
- Sidnell J (2013) Basic conversation analytic methods. In: Sidnell J, Stivers T (eds) *The handbook of conversation analysis*. Chichester, Wiley Blackwell, pp 77–99
- Sidnell J, Stivers T (eds) (2012) *The handbook of conversation analysis*. Wiley-Blackwell, Chichester
- Tolmie P (2011) Uncovering the unremarkable. In: Szymanski M, Whalen J (eds) *Making work visible*. Cambridge University Press, Cambridge, pp 53–73
- Weiss A, Spiel K (2021) Robots beyond science fiction: mutual learning in human–robot interaction on the way to participatory approaches. *AI Soc*. <https://doi.org/10.1007/s00146-021-01209-w>
- Yamazaki K, Yamazaki A, Okada M, Kuno Y, Kobayashi Y, Hoshi Y, Pitsch K, Luff P, von Lehm D, Heath C (2009) Revealing Gauvain: engaging visitors in robot guide's explanation in an art museum. In: Proceedings of CHI'09, April 4–9, Boston, MA, USA
- Yamazaki A, Yamazaki K, Ohyama T, Kobayashi Y, Kuno Y (2012) A techno-sociological solution for designing a museum guide robot: regarding choosing an appropriate visitor. In: Proceedings of HRI'12, March 5–8, Boston, MA, USA
- Złotowski J, Khalil A, Abdallah S (2020) One robot doesn't fit all: aligning social robot appearance and job suitability from a Middle Eastern perspective. *AI Soc* 35:485–500
- Goodwin C (1980) Restarts, pauses, and the achievement of a state of mutual gaze at turn-beginning. *Sociological Inquiry* 50(3–4):272–302
- Yamasaki A, Ogina H, Yamasaki K, Kuzuoka H (2014) Interaction analysis of visitors-robot (Talk Torque-2) interactions in science museum – How body torque of the robot is efficient in human-robot interaction? In: *IEICE Transactions on Information and Systems (Japanese Edition)* Vol. 97(1):28–38.

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.