The impact of human-like and device-like signals on people's mental models of robots

Helena Anna Frijns
helena.frijns@tuwien.ac.at
Vienna University of Technology
Vienna, Austria

Oliver Schürer
oliver.schuerer@tuwien.ac.at
Vienna University of Technology
Vienna, Austria

ABSTRACT

People use mental models to guide their expectations of how interaction with a robot will unfold. We note that in the literature on Human-Robot Interaction, interaction with robots is described with terms that treat robots as devices, but in other work robots are anthropomorphized. Communicative actions by robots are also designed in more human-like or more device-like ways. We need to account for both types of signals when designing interactions with robots. We propose directions for future research into how different types of signals as well as combinations of signals influence people's mental models of robots.

CCS CONCEPTS

• **Human-centered computing → Interaction design theory, concepts and paradigms**; • **Computer systems organization** → *Robotics*;

KEYWORDS

mental models, human-robot interaction, design

1 INTRODUCTION

Upon encountering a robot, it can initially be difficult to understand exactly what this particular robot is capable of. The concept of mental models in Human-Robot Interaction (HRI) refers to expectations people have of (for instance) a robot's behaviour or capabilities. People use their existing experiences of interactions with robots, tools, devices and even other people to inform expectations of what interacting with a robot will be like. The field of cognitive HRI studies mental models people have of the robot and the task [17]. Findings by Kwon et al. [15] suggest that both the appearance and behaviour of a robot influence people's perception of its capabilities. Kiesler [14] argues that people hold implicit mental models of the robots they interact with that include task-specific expectations regarding the way a system will function.

Design can also evoke particular expectations. The design of the embodiment of a robot, the design of its behaviour, and the way tasks are shaped all contribute to the expectations people have of interactions with that robot. Knowing about the relationships between the way a robot is designed and the mental model that people form or call upon when they engage in interactions with a robot will help HRI researchers and roboticists design for these interactions. Relevant is also the concept of the *design metaphor* [5], which refers to the associations and expectations particular design choices call upon on the user's side. We note that forming expectations is not restricted to robots only, and that it does not necessarily have to do with construing a robot as an agent. Instead, people hold and form all kinds of beliefs regarding the way things and people in their environment function, and the same thing happens in the case of robots.

In order to understand how robot embodiments, behaviours and interactions can be designed in such a way that people are supported in forming a correct mental model of the robot, we need to develop an understanding of how different types of cues, signals, and actions influence people's mental models of robots. In this extended abstract we will discuss existing research on communication and interaction between humans and robots. This will yield a distinction between different ways of conceptualizing human-robot interaction and communication. We will also propose directions for future research.

2 CONCEPTUALIZING ROBOTS AS DEVICE-LIKE OR ANTHROPOMORPHIC

In the HRI literature we can distinguish different ways of talking about collaboration and interaction with (social) robots: interacting with a device versus communicating with an agent-like entity. Researchers have developed prototypes that investigate robot-specific communicative cues such as LED lights [7], vibration and sound [20]. We can apply interaction design principles such as those proposed by Norman [18] to the design of robots, thereby considering them as interactive devices. However, robots are sometimes considered as more agentic depending on their autonomy level. Many capabilities are developed that take inspiration from human behaviours. In the field of computational HRI, capabilities have been developed for generating and interpreting verbal communication, non-verbal behaviour, perceiving humans, and affective expressions, as well as interpreting and communicating attention and intentional behaviour [21]. Some researchers discuss interaction with robots in terms of communication with another agent. This would require human-like capabilities, and anthropomorphizes robots to some extent. Effective communication would require the ability to estimate mental states of others to determine which information needs to be communicated to an interaction partner using communicative actions [12]. To summarize, both human-like and artificial signals and behaviours have been developed for robots.

The term anthropomorphism frequently reappears in discussions on the design of social robots, and is often included as a category in design taxonomies [9][1]. We see such a distinction between artificial and human-like signalling in Hegel et al.'s typology of

HRI'20, Workshop on Mental Models of Robots, March 23, 2020, Cambridge, UK

signals and cues for social robotics [11]. However, we argue that human-likeness and device-likeness are not strictly separable when it comes to the design of (social) robot signals. We will give some examples: emotion expression, which is important in communication between humans, can be mimicked in an artificial way by means of LED lights. [20]. Saunderson and Nejat note that non-human-like robot behaviours such as exaggerated motion can actually lead people to rate a robot as more likeable and human-like [19]. Signals such as facial expressions, when expressed by a robot, cannot just be classified as either device-like or anthropomorphic, but lie on a continuum. For instance, a facial expression on a screen will be perceived as more artificial, while a facial expression on an android may be perceived as more anthropomorphic. Simple verbal responses to a command will be perceived as more device-like than full capability for mixed-initiative dialogue.

We pose that no strict distinction can be made between human-like signals and more device-like signals. We propose that both device-like signals and human-like signals can be useful for enhancing the understandability of robot actions, depending on the application. Potential advantages of using artificial signals are decreased cost and complexity compared to more human-like ways of communicating [20]. Functional speech may help set expectations at a more appropriate level than human-like speech if the robot lacks other human-like behaviours [2].

3 DIRECTIONS FOR FUTURE RESEARCH

1. Affordances and external devices

It is important that people can understand the actions that are possible in the current state of the device and get feedback on their actions, so that the interaction is more efficient, pleasant and effective. In HRI, a robot's affordances should be communicated to the human interaction partner, more specifically: the human needs to be able to find out how the robot can be interacted with. Fischer [8] notes that a robot's affordances can be signalled both explicitly (e.g. by means of a visible touch screen) and implicitly (e.g. by indirectly addressing affordances through speech). In line with Fischer, we propose that more research needs to be done on the way robots signal affordances and the way this influences people's mental models of robots.

Clark noted that in games such as chess, the external representation that is the chessboard keeps track of the activity and its current state. Such external representations are physical, their elements or *markers* refer to aspects of the joint activity, and these markers have a spatial location, can be manipulated, and are accessible to the participants in the joint activity [3, p. 45]. We can conceptualize of interfaces that are part of an interaction between a human and a robot as external representations. We propose that more research be done on the role of additional, external or integrated devices on the way a robot is perceived and how these devices can support an interaction. Consider for instance a cobot that is programmed by means of a touch screen. Information displayed on the screen can help decrease a person's cognitive load, but the presence of the screen may also result in people perceiving the robot as more device-like. Where do people think the locus of control is? Do they see the robot arm as a tool, or would that depend on the robot's actions or possibilities for interacting with the arm directly?

2. Composite signals

Having access to both human-like and device-like signals can help set adequate expectations. In addition, viewing the design of a robot's embodiment and behaviour as an *assemblage* of different types of signals and strategies, both human-like and device-like, will both support analysis of the effects of design choices and open up opportunities that may be missed if the main strategy is to make the robot more (or less) anthropomorphic. The relationship between combinations of different types of signals and the interpretation of these combinations is complex. Preliminary findings by Lee et al. [16] suggest that users will describe a robot's features as human-like, rather than perceiving a robot as anthropomorphic in its entirety. At the same time, signals and cues can obtain new meaning when used in combination with other signals and cues. For the design of Flobi, a humanoid robotic head, a modular design was chosen to investigate the effect of different hairstyles and clothes on perceived gender and social roles [10]. We also refer to Clark's notion of composite signals. As opposed to calling a signal either linguistic or non-linguistic, he calls the method of signalling linguistic or non-linguistic. Signals can be made up of combinations of signalling methods, forming *composite signals* [4]. One signal can be composed out of different signs using multiple methods of signalling. This also applies to robot signals, and robots can combine signals that are more device-like or more human-like. Timing is a relevant factor here: communication is not just multimodal (see [6][19]); the way signals from different modalities are organized has relevance to the meaning of the formed composite signals.

We propose that more research be done on combinations of signals and that researchers empirically establish a link between the type of signal or combination of signals used and the way people conceive of the robot's capabilities. Researchers have compared people's perception of human-like versus machine-like robots [13]. The interaction between combinations of robot behaviours and human mental models, however, remains a research challenge (also refer to [19, p.595]). It has been proposed that closely following existing social roles can help set user expectations appropriately for specific tasks [14]. We would argue against such an approach, as this could unnecessarily promote stereotypes in society. Rather, we pose that the effects of sets of behaviours and capabilities (such as voice or gaze) on user expectations should be studied. It would also be useful to know if displaying particular capabilities will prompt people to generalize to other capabilities and at what point this leads to inaccurate mental models.

There are some limitations to carrying out studies into the design of social robots. Carrying out studies on design features can become costly and complex if the aim is to try out many different variations, findings may not generalize to other systems, and appropriate measurement tools can still be lacking [5]. Cha et al. [2] developed a survey with the aim of measuring perceived robot capability, but it is desirable to develop more ways of measuring people's perception of a robot's capabilities, as well as the next actions they expect a robot to take. Despite the challenges, we propose that viewing a robot's behaviour and embodiment as a vehicle for a mix of machine-like and human-like signals, as well as studying the effects of combining different types of signals, will be useful in the effort to design understandable robot behaviour.

ACKNOWLEDGMENTS

This work was supported by TU Wien through the Doctoral College TrustRobots.

REFERENCES

[1] C. Bartneck and J. Forlizzi. 2004. A design-centred framework for social human-robot interaction. In *RO-MAN 2004. 13th IEEE International Workshop on Robot and Human Interactive Communication (IEEE Catalog No.04TH8759)*. IEEE, pp. 591–594. https://doi.org/10.1109/ROMAN.2004.1374827

[2] Elizabeth Cha, Anca D. Dragan, and Siddhartha S. Srinivasa. 2015. Perceived robot capability. In *2015 24th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*. IEEE, 541–548. https://doi.org/10.1109/ROMAN.2015.7333656

[3] Herbert H. Clark. 1996. Joint actions. In *Using Language*. Cambridge University Press.

[4] Herbert H. Clark. 1996. Signaling. In *Using Language*. Cambridge University Press.

[5] Eric Deng, Bilge Mutlu, and Maja J Mataric. 2019. Embodiment in Socially Interactive Robots. 7, 4 (2019), pp. 251–356. https://doi.org/10.1561/2300000056

[6] Bruno Dumas, Denis Lalanne, and Sharon Oviatt. 2009. Multimodal Interfaces: A Survey of Principles, Models and Frameworks. In *Human Machine Interaction*, Denis Lalanne and Jürg Kohlas (Eds.). Vol. 5440. Springer Berlin Heidelberg, pp. 3–26. https://doi.org/10.1007/978-3-642-00437-7_1

[7] S. Embgen, M. Luber, C. Becker-Asano, M. Ragni, V. Evers, and K. O. Arras. 2012. Robot-specific social cues in emotional body language. In *2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*. pp. 1019–1025. https://doi.org/10.1109/ROMAN.2012.6343883

[8] Kerstin Fischer. 2018. When Transparent does not Mean Explainable. In *Proceedings of 'Explainable Robotic Systems', Workshop in conjunction with the HRI 2018 Conference*. Chicago.

[9] Terrence Fong, Illah Nourbakhsh, and Kerstin Dautenhahn. 2003. A survey of socially interactive robots. 42, 3 (March 2003), pp. 143–166. https://doi.org/10.1016/S0921-8890(02)00372-X

[10] Frank Hegel. 2013. A Modular Interface Design to Indicate a Robot's Social Capabilities.

[11] Frank Hegel, Sebastian Gieselmann, Annika Peters, Patrick Holthaus, and Britta Wrede. 2011. Towards a typology of meaningful signals and cues in social robotics. In *2011 RO-MAN*. IEEE, pp. 72–78. https://doi.org/10.1109/ROMAN.2011.6005246

[12] Thomas Hellström and Suna Bensch. 2018. Understandable robots - What, Why, and How. *Paladyn, Journal of Behavioral Robotics* 9, 1 (July 2018), pp. 110–123. https://doi.org/10.1515/pjbr-2018-0009

[13] Pamela J Hinds, Teresa L Roberts, and Hank Jones. 2004. Whose Job Is It Anyway? A Study of Human-Robot Interaction in a Collaborative Task. 19, 1-2 (2004), pp. 151–181.

[14] S. Kiesler. 2005. Fostering common ground in human-robot interaction. In *RO-MAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005*. pp. 729–734. https://doi.org/10.1109/ROMAN.2005.1513866 ISSN: 1944-9437.

[15] M. Kwon, M. F. Jung, and R. A. Knepper. 2016. Human expectations of social robots. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. pp. 463–464. https://doi.org/10.1109/HRI.2016.7451807

[16] Hee Rin Lee, Selma Sabanovic, and Erik Stolterman. [n. d.]. How Humanlike Should a Social Robot Be: A User-Centered Exploration. ([n. d.]), pp. 135–141.

[17] Bilge Mutlu, Nicholas Roy, and Selma Sabanovic. 2016. Cognitive Human-Robot Interaction. In *Springer Handbook of Robotics* (2nd ed.). Springer, pp. 1907–1933.

[18] Donald A. Norman. 2013. *The design of everyday things* (revised and expanded ed.). Basic Books.

[19] Shane Saunderson and Goldie Nejat. 2019. How Robots Influence Humans: A Survey of Nonverbal Communication in Social Human–Robot Interaction. *International Journal of Social Robotics* 11, 4 (Aug. 2019), pp. 575–608. https://doi.org/10.1007/s12369-019-00523-0

[20] Sichao Song and Seiji Yamada. 2017. Expressing Emotions through Color, Sound, and Vibration with an Appearance-Constrained Social Robot. In *Proceedings of the 2017 ACM/IEEE International Conference on Human-Robot Interaction - HRI '17*. ACM Press, pp. 2–11. https://doi.org/10.1145/2909824.3020239

[21] Andrea Thomaz, Guy Hoffman, and Maya Cakmak. 2016. Computational Human-Robot Interaction. *Foundations and Trends in Robotics* 4, 2-3 (2016), pp. 104–223. https://doi.org/10.1561/2300000049