Human-Robot Interactions -

Effects of Agency, Autonomy and Embodiment on User Trust

We welcome applications of PhD students in the domain of human-robot interaction (HRI) with a focus on trust and its relation to perceived agency and autonomy, and to robot embodiment.

PhD Topic Summary

Assistive robots play an increasing role in society, particularly in older people's care, where trust in the system is crucial for the success of the technology. However, people often have mis-calibrated levels of trust in such systems, with older people's trust levels frequently being lower than in younger people.
This project aims to contribute to the overall understanding of people's trust in autonomous systems as mechanisms for trust calibration are not as well researched across the population. The project will explore 'agency' as a potential influencing factor for trust in assistive technology. Robot agency relates to whether it is perceived to be autonomous or remote controlled and how it looks, e.g., like a human or a machine. The project will systematically explore these two parameters in Human-Robot Interaction experiments to be conducted at the UH Robot House. Findings will facilitate better alignment of autonomous systems with the needs of people of different ages. The planned work will improve the lives of a potentially vulnerable population, particularly by increasing understanding of older people's trust towards unfamiliar technologies.

Background

There is a general tendency today to over-trust robots [1], for example during emergencies [2]. This is especially true for older people where age-related limitations can lead to over-trust in automation and consequently to over-reliance on some assistive technologies [3]. However, there is also contrary evidence that older people lack trust in technology, like navigation systems [4] or self-driving vehicles [5]. Hence, there is an urgent need to understand the exact factors that influence trust across population.
It is known that the perception of a robot affects a human's trust [6, 7], including its perceived autonomy [8], appearance [9], and behaviours [10], which all individually contribute to the user's concept of a robot's agency [11].
Interactive robots often set false expectations on their degree of autonomy [12]. Human-like robot appearances also make it difficult to assess the capabilities of a robotic agent [13]. We strongly suspect such discrepancies to cause wrong concepts of robot agency and consequently lead to improperly calibrated trust, especially in older people.
A focused analysis of how people develop an understanding of robot agency is thus required to understand how they place their trust in such a device.

Aims of this PhD are:

1. To establish a relation between perceived agency, autonomy, embodiment, and trust towards an assistive robot.
2. To conduct HRI experiments, in which to manipulate the perceived agency of different robot embodiments.
3. To find derive guidance on how designers can make more trustworthy assistive systems that meet the needs of users of different age groups who struggle to find correctly calibrated levels of trust.

Person Specifications

Applicants should optimally have a background in computer science or a similar field. However, given the interdisciplinary nature of the field, applicants with a background in ethology or psychology may apply too. Experience in planning, designing, and conducting experiments with human participants (preferably in the field of HRI) is desirable but not essential. Applicants will need a background in at least one programming language.

Supervisory Team

Dr. Gabriella Lakatos

Dr. Diego Faria

Dr. Patrick Holthaus

Dr. Farshid Amirabdollahian

Contact

For informal enquiries about this PhD, please contact Dr. Gabriella Lakatos at g.lakatos@herts.ac.uk

References

[1] Wagner A., Borenstein, J. & Howard A. (2018). Overtrust in the robotic age. Communications of the ACM, 61(9): 22-24.

[2] Robinette, P., Howard, A. & Wagner, A. (2017), Conceptualizing overtrust in robots: Why do people trust a robot that previously failed? In 'Autonomy and Artificial Intelligence: A Threat or Savior?', Springer, Cham, pp. 129–155.

[3] Pak, R., et al. (2017). Does the domain of technology impact user trust? Theoretical issues in ergonomics science, 18(3): 199-220.

[4] Fuchsberger, V., et al. (2012). Supporting older adults' indoor and outdoor wayfinding: Getting to know the user in terms of traveling and navigation. In Proceedings of the AAL Forum 2012.

[5] Hulse, L., Xie, H., & Galea, E. (2018). Perceptions of autonomous vehicles: Relationships with road users, risk, gender and age. Safety Science, 102: 1-13.

[6] Hancock P., et al. (2011). A meta-analysis of factors affecting trust in human-robot interaction. Human factors, 53(5): 517-527.

[7] Schaefer K. The perception and measurement of human-robot trust. Dissertation. University of Central Florida Orlando, Florida, 2013.

[8] Beer, J., Fisk, A., & Rogers, W. (2014). Toward a framework for levels of robot autonomy in human-robot interaction. Journal of human-robot interaction, 3(2), 74–99.

[9] van Maris, A., et al. (2017). The influence of a robot's embodiment on trust: A longitudinal study, in 'Proceedings of the International Conference on Human-Robot Interaction (ACM/IEEE HRI 2017)', 313–314.

[10] Salem, M., Lakatos, G., Amirabdollahian F., et al. (2015). Would you trust a (faulty) robot? Effects of error task type and personality on human-robot cooperation and trust. In 'Proceedings of the International Conference on Human-Robot Interaction (ACM/IEEE HRI 2015)', 141-148.

[11] van der Woerdt, S., & Haselager, P. (2019). "When robots appear to have a mind: The human perception of machine agency and responsibility." New Ideas in Psychology 54: 93-100.

[12] Groom V., & Nass C. (2007). Can robots be teammates?: Benchmarks in human-robot teams. Interaction Studies, 8(3): 483-500.

[13] Hegel, F., Gieselmann, S., Peters, A., Holthaus, P., & Wrede, B. (2011). Towards a Typology of Meaningful Signals and Cues in Social Robotics. In 'Proceedings of International Symposium on Robot and Human Interactive Communication (IEEE RO-MAN 2011)', 72–78.