Agency Effects on Robot Trust in Different Age Groups

Patrick Holthaus
p.holthaus@herts.ac.uk
University of Hertfordshire
School of Physics, Engineering and Computer Science
Hatfield, United Kingdom

Ali Fallahi
a.fallahi@herts.ac.uk
University of Hertfordshire
School of Physics, Engineering and Computer Science
Hatfield, United Kingdom

Frank Förster
f.foerster@herts.ac.uk
University of Hertfordshire
School of Physics, Engineering and Computer Science
Hatfield, United Kingdom

Catherine Menon
c.menon@herts.ac.uk
University of Hertfordshire
School of Physics, Engineering and Computer Science
Hatfield, United Kingdom

Luke Wood
l.wood@herts.ac.uk
University of Hertfordshire
School of Physics, Engineering and Computer Science
Hatfield, United Kingdom

Gabriella Lakatos
g.lakatos@herts.ac.uk
University of Hertfordshire
School of Physics, Engineering and Computer Science
Hatfield, United Kingdom

ABSTRACT

Trust plays a major role when introducing interactive robots into people's personal spaces, which, in large part, depends on how they perceive the robot. This paper presents the initial results of an investigation into the perception of robot agency as a potential factor influencing trust. We manipulated a robot's agency to see how trust would change as a result. Our preliminary results indicate age as a confounding factor while we did not find differences when priming robot autonomy.

ACM Reference Format:
Patrick Holthaus, Ali Fallahi, Frank Förster, Catherine Menon, Luke Wood, and Gabriella Lakatos. 2024. Agency Effects on Robot Trust in Different Age Groups. In *International Conference on Human-Agent Interaction (HAI '24), November 24–27, 2024, Swansea, United Kingdom.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3687272.3690903

1 INTRODUCTION

Autonomous systems, such as self-driving cars or socially assistive robots, appear to humans as autonomous agents. However, it is not uncommon for people's expectations of the degree of agency possessed by such systems to be incorrect [7, 10, 17]. At the same time, people also have a tendency to over-trust these robots [3, 26], hinting at a potential relation between the two concepts of trust and expectation of agency. We believe that understanding the relation between the robot's perceived *agency* and trust at different ages helps roboticists to better calibrate the behaviour of the robot and evoke an optimal level of trust in it.

Trust plays a significant role in determining the adoption of automation [18, 27]. Operators are unlikely to use an automated system that they perceive as unreliable or untrustworthy [19]. Most roboticists agree that a lack of trust in a robot can lead to disengagement [22] while over-trust in a robot can cause over-reliance, leading to risks arising [3], for example, in emergency situations [21]. In human-robot interaction, robot performance is one of the major influences on people's trust in the robot [15]. In addition to performance, a robot's attributes (e.g. personality or anthropomorphism) can also influence trust in that robot [8]. Since trust in robots also relies on external factors such as the manufacturer [28], it is important to disentangle how the concept of trust relates to these individual components, i.e. how much users trust the robot itself, its operator or manufacturer.

Robot autonomy and behaviour of robots can affect the trust as a psychological state of human users [11, 12]. There has been an increasing focus in human-robot interaction research [24] on people's perception of robot agency. Robot autonomy is one of the main driving factors of how people will perceive robot agency. According to Luck et al. [14] an object qualifies as an agent when it achieves a beneficial function for either another object or itself, thereby achieving autonomy. Autonomy pertains to the ability of an agent to operate within its environment independently [1]. Consequently, this paper investigates the potential connection between autonomy and agency on the one hand and trust on the other hand. For that, we present the initial results of an online experiment in which we manipulate a robot's perceived agency to determine how these changes affect trust to address the question "How does perceived agency influence people's trust towards an interactive robot?".

2 METHODS

We designed and conducted an online study via Microsoft Forms to examine people's perception of agency across Generation X, Y, and Z. Our study examined the relationship between expectations of a robot's agency, people's age and gender and levels of trust in that robot. Participants were invited via social media, personal contacts, and mailing lists. The study was approved by the University of Hertfordshire (UH) Health, Science, Engineering & Technology Ethics Committee with Delegated Authority (ECDA) under protocol number LMS/SF/UH/05127. We manipulated participants' perceptions

HAI '24, November 24–27, 2024, Swansea, United Kingdom

ACM ISBN 979-8-4007-1178-7/24/11
https://doi.org/10.1145/3687272.3690903

of the robot's agency via manipulation of perceived autonomy before the interaction. The study used text- and image-based primers to describe the robot as autonomous or remote-controlled in two groups, keeping its behaviour identical. We hypothesise that higher levels of perceived agency lead to elevated levels of trust in the robot and that trust also depends on people's age and gender.

In total, 76 participants consented and participated in the study. Of those, 38 identified as female, 37 as male and one as non-binary. Moreover, 20 participants were in Generation Z (younger than 26), 46 were in Generation Y (between 27 and 42 years) and 10 were in Generation X+ (older than 43 years). Participants were randomly assigned to one of two experimental conditions with 39 participants in condition A (autonomous) and 37 in condition R (remote-controlled). In line with other researchers, we rely on a mix of objective and subjective methods to determine how trustworthy a robot is [2, 9, 23] in our experiment. The experiment consisted of four phases, including a pre-test and introduction phase, a manipulation phase for priming the conditions, a virtual interaction phase, and a post-test and debriefing phase. After consenting to participate in the experiment, we first collected participant demographics, experience with robots, the Ten-Item Personality Inventory (TIPI) [6] and Negative Attitude Towards Robots (NARS) [16] scale to identify potential effects on the dependent variables.

In a between-subjects experimental design, we then primed participant in two condition groups to perceive the robot as either autonomous (A) or remote-controlled (R). Participants were then led through an interactive questionnaire. Priming was achieved by a text snippet in both conditions and an additional photo of a control room showing computer monitors observing the experimental area in condition R. Participants were then instructed to virtually interact with the robot where they watched videos showing the robot from a first-person perspective. In between the presentation of videos, participants had the option to accept or deny robot suggestions, which we used as objective measures of trust placed in the robot. Depending on a participant's answer, the robot would then execute the requested function, for example, take the glass from the table or leave without the glass if asked to do so. After the virtual interaction, participants were asked to evaluate the robot using the Robot Social Attribute Scale (RoSAS) [5]. We used the Multi-Dimensional Trust Measuring Tool (MDMT) [25] and some individual questions to determine whether participants' opinion of the robot and trust placed in it correlated with the primed condition. Due to the nature of the data (Shapiro-Wilk test), we used non-parametric tests, i.e. Wilcoxon rank-sum test for independent samples to analyse the effects of condition, age, and gender.

3 INITIAL RESULTS

When presented with a choice during the virtual interaction with the robot, participants did not significantly differ in their response based on the experimental condition, participant age, or gender.

We found a significant difference between the two conditions when analysing TIPI, one of the pre-test questionnaires, potentially influencing our findings. The factor *Agreeableness* was higher in condition R than in A ($W = 938$, $p < 0.05$). No effects have been found in the interaction choices or post-interaction questionnaires.

The experiment yielded several differences between the different age groups, both in the pre-test and the post-interaction questionnaires. Generation X+ had a more negative attitude towards robots in *Situational Interactions* (NARS S1) than Generation Y ($W = 661.5$, $p < 0.01$) and Generation Z ($W = 160.5$, $p < 0.01$). Generation Y had a significantly higher negative attitude towards *Emotions* with robots (S3) than Generation Z ($W = 612$, p-value = 0.05)· Generation X+ scored significantly higher than Generation Y on the *Agreeableness* factor in TIPI ($W = 323$, $p < 0.05$), while Generation Z scored higher than Generation Y in terms of *Conscientiousness* ($W = 309$, $p < 0.05$). No other effects between the generations were observed in the pre-test. In the post-test, Generation X+ rated the robot lower than Generation Z in terms of *Warmth* on RoSAS ($W = 51$, $p < 0.05$) and Generation Z expressed higher ratings than Generation Y on the *Sincerity* subscale of MDMT. No further effects between the generations were found.

Statistical analysis has been conducted solely between males and females only a single participant disclosed a different gender (non-binary). The following statistically significant effects have been found in the pre-test: Females expressed a significantly higher negative attitude towards robots on NARS, both in *Situational Interactions* (S1) ($W = 938$, $p < 0.05$) and in *Emotions* (S3) ($W = 476.5$, $p < 0.05$). Moreover, females also scored higher on TIPI, i.e. on the *Agreeableness* ($W = 438.5$, $p < 0.01$) and *Conscientiousness* subscales ($W = 495$, $p < 0.05$). No effects between the disclosed gender groups have been found in the post-test or interaction choice.

4 DISCUSSION AND CONCLUSIONS

Our initial results do not show any significant differences between the two conditions or participant gender considering people's trust in the robot. However, participant age had several effects. We can therefore not conclusively answer our research question of how trust in a robot is influenced by perceived robot agency.

While we still suspect a relation, the experiment showed that our way of priming autonomy was insufficient to elicit the desired effects. Perhaps a stronger priming method, considering people's preferences for personal data handling [20], using additional video footage of an operator, or repeated priming for the duration of the experiment, would be necessary. Participants' personality might have a confounding factor that could be targeted in a future experiment. In line with other literature [13], older adults in our experiment were generally more likely to display a negative attitude towards robots than younger adults and, in some parts, showed lower levels of trust in the robot. The clear effects and trends suggest that age is an important factor in the perception of robots and older people's attitude toward them, implying that different age groups have varying predispositions to trusting them. Our findings contrast earlier work [4] which did not find a difference in negative attitudes towards robots between the generations.

In this paper, we presented an initial investigation into the potential relationship between people's perception of a robot's agency and their trust towards that robot. We primed different levels of a robot's autonomy to alter people's perception of its agency and see how trust would change. In future work, we aim to further explore the connection between perceived robot agency and trust.

REFERENCES

[1] Hussein A Abbass, Jason Scholz, and Darryn J Reid. 2018. *Foundations of trusted autonomy*. Springer Nature.

[2] Farshid Amirabdollahian, Kerstin Dautenhahn, Clare Dixon, Kerstin Eder, Michael Fisher, Kheng L Koay, Evgeni Magid, Tony Pipe, Maha Salem, Joe Saunders, et al. 2013. Can you trust your robotic assistant? *Social Robotics* (2013).

[3] Alexander M. Aroyo, Jan de Bruyne, Orian Dheu, Eduard Fosch-Villaronga, Aleksei Gudkov, Holly Hoch, Steve Jones, Christoph Lutz, Henrik Sætra, Mads Solberg, and Aurelia Tamò-Larrieux. 2021. Overtrusting robots: Setting a research agenda to mitigate overtrust in automation. *Paladyn, Journal of Behavioral Robotics* 12, 1 (2021), 423–436. https://doi.org/doi:10.1515/pjbr-2021-0029

[4] Uba Backonja, Amanda K Hall, Ian Painter, Laura Kneale, Amanda Lazar, Maya Cakmak, Hilaire J Thompson, and George Demiris. 2018. Comfort and attitudes towards robots among young, middle-aged, and older adults: a cross-sectional study. *Journal of Nursing Scholarship* 50, 6 (2018), 623–633.

[5] Colleen M Carpinella, Alisa B Wyman, Michael A Perez, and Steven J Stroessner. 2017. The robotic social attributes scale (rosas) development and validation. In *Proceedings of the 2017 ACM/IEEE International Conference on human-robot interaction*. 254–262.

[6] Samuel D Gosling, Peter J Rentfrow, and William B Swann Jr. 2003. A very brief measure of the Big-Five personality domains. *Journal of Research in personality* 37, 6 (2003), 504–528.

[7] Victoria Groom and Clifford Nass. 2007. Can robots be teammates?: Benchmarks in human–robot teams. *Interaction studies* 8, 3 (2007), 483–500.

[8] Peter A Hancock, Deborah R Billings, Kristin E Schaefer, Jessie YC Chen, Ewart J De Visser, and Raja Parasuraman. 2011. A meta-analysis of factors affecting trust in human-robot interaction. *Human factors* 53, 5 (2011), 517–527.

[9] Patrick Holthaus and Alessandra Rossi. 2023. Common (good) practices measuring trust in HRI. In *RO-MAN 2023 Workshop on Trust, Acceptance and Social Cues in Human-Robot Interaction - SCRITA*. https://arxiv.org/abs/2311.12182

[10] Lynn M Hulse, Hui Xie, and Edwin R Galea. 2018. Perceptions of autonomous vehicles: Relationships with road users, risk, gender and age. *Safety science* 102 (2018), 1–13.

[11] Bing Cai Kok and Harold Soh. 2020. Trust in robots: Challenges and opportunities. *Current Robotics Reports* 1, 4 (2020), 297–309.

[12] Allison Langer, Ronit Feingold-Polak, Oliver Mueller, Philipp Kellmeyer, and Shelly Levy-Tzedek. 2019. Trust in socially assistive robots: Considerations for use in rehabilitation. *Neuroscience & Biobehavioral Reviews* 104 (2019), 231–239.

[13] Stephanie Lehmann, Esther Ruf, and Sabina Misoch. 2020. Emotions and attitudes of older adults toward robots of different appearances and in different situations. In *International Conference on Information and Communication Technologies for Ageing Well and E-Health*. Springer, 21–43.

[14] Michael Luck, Mark d'Inverno, et al. 1995. A Formal Framework for Agency and Autonomy.. In *Icmas*, Vol. 95. 254–260.

[15] Peter E. McKenna, Muneeb I. Ahmad, Tafadzwa Maisva, Birthe Nesset, Katrin Lohan, and Helen Hastie. 2024. A Meta-analysis of Vulnerability and Trust in Human-Robot Interaction. *J. Hum.-Robot Interact.* (apr 2024). https://doi.org/10.1145/3658897 Just Accepted.

[16] Tatsuya Nomura, Tomohiro Suzuki, Takayuki Kanda, and Kensuke Kato. 2006. Measurement of negative attitudes toward robots. *Interaction Studies. Social Behaviour and Communication in Biological and Artificial Systems* 7, 3 (2006), 437–454.

[17] Sven Nyholm. 2020. *Humans and robots: Ethics, agency, and anthropomorphism*. Rowman & Littlefield Publishers.

[18] Thomas Nys. 2015. Autonomy, trust, and respect. In *The Journal of Medicine and Philosophy: A Forum for Bioethics and Philosophy of Medicine*, Vol. 41. Journal of Medicine and Philosophy Inc., 10–24.

[19] Raja Parasuraman and Victor Riley. 1997. Humans and automation: Use, misuse, disuse, abuse. *Human factors* 39, 2 (1997), 230–253.

[20] Lewis Riches, Kheng Lee Koay, and Patrick Holthaus. under review. Privacy in Robotics: Investigating the Acceptability of Personal Data Knowledge and Perceived Risk of Disclosure. *International Journal of Social Robotics* (under review).

[21] Paul Robinette, Wenchen Li, Robert Allen, Ayanna M Howard, and Alan R Wagner. 2016. Overtrust of robots in emergency evacuation scenarios. In *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 101–108.

[22] Maha Salem, Gabriella Lakatos, Farshid Amirabdollahian, and Kerstin Dautenhahn. 2015. Towards safe and trustworthy social robots: ethical challenges and practical issues. In *Social Robotics: 7th International Conference, ICSR 2015, Paris, France, October 26-30, 2015, Proceedings 7*. Springer, 584–593.

[23] Maha Salem, Gabriella Lakatos, Farshid Amirabdollahian, and Kerstin Dautenhahn. 2015. Would you trust a (faulty) robot? Effects of error, task type and personality on human-robot cooperation and trust. In *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction*. 141–148.

[24] J. Gregory Trafton, J. Malcolm McCurry, Kevin Zish, and Chelsea R. Frazier. 2024. The Perception of Agency. *J. Hum.-Robot Interact.* (jan 2024). https://doi.org/10.1145/3640011

[25] Daniel Ullman and Bertram F Malle. 2019. Measuring gains and losses in human-robot trust: Evidence for differentiable components of trust. In *2019 14th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*. IEEE, 618–619.

[26] Alan R. Wagner, Jason Borenstein, and Ayanna Howard. 2018. Overtrust in the robotic age. *Commun. ACM* 61, 9 (aug 2018), 22–24. https://doi.org/10.1145/3241365

[27] Michael L Walters, Mohammedreza A Oskoei, Dag Sverre Syrdal, and Kerstin Dautenhahn. 2011. A long-term human-robot proxemic study. In *2011 RO-MAN*. IEEE, 137–142.

[28] Marie Waung, Pam McAuslan, and Sridhar Lakshmanan. 2021. Trust and intention to use autonomous vehicles: Manufacturer focus and passenger control. *Transportation research part F: traffic psychology and behaviour* 80 (2021), 328–340.