A Robot's Moral Advice Is Not Appreciated Neither in Functional nor in Social Communication

Clarissa Sabrina Arlinghaus[1], Carolin Straßmann[2], and Annika Dix[3]

***Abstract*— This study (*N* = 317) investigated the influence of verbal communication (social vs. functional) on the acceptance of robot recommendations in non-moral, somewhat moral or very moral decision-making situations. The robot's communication style had no impact on the participants (1) being confident in their decision, (2) perceiving the robot's recommendation as helpful, and (3) making a decision dependent on the robot's recommendation. However, all three aspects were strongly influenced by the morality of the decision situation demonstrating higher algorithm aversion in moral contexts.**

I. INTRODUCTION

Robots could support human decision-making with recommendations based on algorithms but people prefer human advice which is known as algorithm aversion [1]. Algorithm aversion is especially pronounced in medical [2], [3] or moral decision-making [4].

Medical decision-making can quickly involve moral issues (e.g., allocating scarce lifesaving resources [25]) which combines both aspects promoting algorithm aversion. Therefore, we chose medical decisions with varying moral charge as relevant context to explore under which conditions recommendations from robots are more likely to be accepted. All robot's moral recommendations are characterized by a utilitarian mindset.

Besides non-moral and moral decision-making, our focus is on the robot's verbal communication style. Communication is a key aspect in designing human-robot interactions [5]. Many studies combine verbal communication with several non-verbal communication features like eye color, eye gaze, and gestures [6] so that the impact of verbal communication alone remains unclear. This is why we solely manipulated the robot's verbal communication style.

We examine whether social communication reduces algorithm aversion in non-moral or moral decision situations and thus leads to people accepting the robot's recommendations more compared to functional communication. We did a preregistered [7] German online experiment with videos of the robot Pepper (Softbank Robotics) giving recommendations for non-moral, somewhat moral, and very moral medical decisions in a functional or social verbal communication style. As far as we know, the effects of social communication in medical-moral human-robot interactions have not been researched yet. With our study, we want to contribute to a deeper understanding and more reflective use of communication styles and robot recommendations for non-moral or moral decision-making. This late-breaking report will only present our exploratory analyses. The full paper with the hypotheses-related analyses is currently under review [8].

[1] Bielefeld University, Bielefeld, North-Rhine Westphalia, Germany clarissa_sabrina.arlinghaus@uni-bielefeld.de

[2] University of Applied Sciences Ruhr West, Bottrop, North-Rhine Westphalia, Germany carolin.strassmann@hs-ruhrwest.de

[3] TU Dresden, Dresden, Saxony, Germany annika.dix@tu-dresden.de

II. RELATED WORK

People don't like machines making moral decisions [4]. However, prior studies show that different verbal communication styles do have an impact on accepting social robots. For example, polite language increases people's compliance to the robot's medical recommendations [9] and people are more likely to do the robot a favor when the robot is using indirect requests [10]. Therefore we explore, whether social communication can increase a robot's acceptance in moral decision-making compared to functional communication.

For creating social communication, we used the social concepts self-disclosure, content intimacy, and relational continuity constructional units [11] as well as we-phrases. The functional robot communication style instead consisted of rational arguments and the third-person passive voice. The communication styles can be found in the project link given in the beginning of research methods section.

Self-disclosure means sharing personal information about oneself [11]. Implementing self-disclosure is recommended for using robots in healthcare [12]. Robots' self-disclosure makes people attribute more mind to robots which is associated with moral agency [13]. *Content intimacy* is provided by detailed, intimate, and emotional elements of a conversation [11]. Content intimacy could counteract the criticized lack of emotions that may be responsible for aversion in moral context [14]. Emotional robots are more persuasive [15] and seen as more morally accountable [16]. Furthermore, people are more likely to cooperate with a robot in a moral dilemma when the robot is showing emotions [17]. *Relational continuity constructional units* describes referring to behavior before, during, or after an absence in a conversation [11]. This concept has not been researched much yet but is an aspect of social communication according to [11]. *We-phrases* are the use of the first person plural ("we"). It is a persuasion strategy known from politics [18] and also recommended for medical settings [19]. We-phrases could also contribute to a stronger ingroup perception which could mitigate algorithm aversion because [20] observed a preference for ingroup robots over outgroup humans. Moreover, robots using self-referential pronouns are perceived as more competent [21].

These broad findings suggest that social communication could increase the acceptance of robots in moral decision-

making situations compared to functional communication. Nevertheless, this has not yet been investigated. With this study, we address this research gap with the following pre-registered [7] exploratory research questions:

- *EQ1*: Does social communication (vs. functional communication) affect how confident the person is in their decision?
- *EQ2*: Does social communication (vs. functional communication) affect how helpful the person finds the robot's recommendation?
- *EQ3*: Does social communication (vs. functional communication) affect how much the person made their decision based on the robot's recommendation?

III. Research methods

This late-breaking report is part of a bigger pre-registered study [7], which is currently under review [8]. The materials and collected data are available under the following link: `https://osf.io/dpc3q/`

A. Sample

The sample (N = 317) consists of 179 women (56.5 %), 128 men (40.4 %), and three non-Binaries (1.0 %). Seven people (2.2 %) did not disclose their gender. The participants were aged between 18 to 70 years (M = 29.42, SD = 11.41). Most of them were attending university (59.62 %, n = 189) or in employment (34.38 %, n = 109). All participants gave informed consent. The study was approved by the university's data protection manager and ethics committee (SR-EK-284062022). Participation was voluntary, anonymous, and without any compensation.

B. Design

We used a 2 x 3 between-subjects design for our study. The combination of the factors communication (social vs. functional) and decision situation (non-moral vs. somewhat-moral vs very moral) resulted in six experimental conditions:

- *Condition 1*: functional communication - non-moral decision situation
- *Condition 2*: functional communication - somewhat decision situation
- *Condition 3*: functional communication - very moral decision situation
- *Condition 4*: social communication - non-moral decision situation
- *Condition 5*: social communication - somewhat decision situation
- *Condition 6*: social communication - very moral decision situation

C. Procedure

All participants were randomly assigned to one experimental condition. They were presented a non-moral, somewhat moral, or very moral decision situation in a hospital in text form depending on the assigned condition. The participants received a video with a recommendation given by the robot Pepper (Softbank Robotics) which was communicated in a functional or social communication style. After that, they were asked to rate the robot's recommendation. The whole experiment was conducted online in German language.

D. Decision situations

We developed three decision situations: non-moral, somewhat moral, very moral.

Non-moral: A new X-ray machine has to be purchased. Three machines are available with different periods of use, acquisition costs, depreciation, and annual profits. Participants should decide for one machine to buy.

Somewhat moral: New respiratory machines has to be purchased with a budget of 60,000 Euro. Three company offers are available with different numbers of machine and quality of the machines. Participants should decide for one offer to accept.

Very moral: An intensive care bed has to be allocated. Three patients with different ages and chances of survival are in need of intensive care.

The three decision situations were pre-tested in advance (N = 27, M = 26.15). In an online experiment, participants saw all the decision situations in randomized order and were asked to rate how morally challenging the situation was. The pretest demonstrated that the intended manipulation was successful. That means that the three decision situations vary in terms of morality (non-moral $<$ somewhat moral $<$ very moral) and thus form a three-stage increase in how morally challenging the situation is.

E. Recommendations

We developed six recommendations. For each decision situation, the same recommendation was developed in a functional and social communication style.

Non-moral: The robot recommends buying CT device 2 because of its shortest amortization period.

Somewhat moral: The robot recommends accepting company C's offer because of the highest number of people it would save.

Very moral: The robot recommends selecting patient C because of the biggest chance of survival.

The non-moral recommendation was inspired by [22]. The somewhat moral and very moral recommendations represent a utilitarian mindset. Making utilitarian choices is expected of robots in morally challenging situations [23] and also advised for medical moral decisions like vaccinations [24] or triage [25].

Self-disclosure, content intimacy, relational continuity constructional units [11], and "we-phrases" were used to create a social communication style. Meanwhile, the functional robot communication style consisted of rational arguments and the third-person passive voice. Consequently, social communication is based more on emotionality while functional communication is based more on rationality.

The six recommendations were pre-tested in advance (N = 23, M = 26.74). In an online experiment, participants saw all the recommendations in randomized order and were asked to rate how functional or social the communication was.

The pretest demonstrated that the intended manipulation was successful. That means that the social communication style is perceived as more social than the functional communication style.

F. Measures

Decision certainty was measured with one item (How confident are you in this decision?) on a 0 (not) to 100 (very) scale.

Decision help was measured with one item ("How helpful did you find the robot's recommendation?") on a 0 (not) to 100 (very) scale.

Decision dependency was measured with one item ("How much did your decision depend on the robot's recommendation?") on a 0 (not) to 100 (very) scale.

G. Analyses

To answer the research questions, we calculated a MANOVA with the factors communication (social vs. functional) and decision situation (not moral vs. somewhat moral vs. very moral) as the independent variables, and decision certainty, decision help, and decision dependency as the dependent variables. One multivariate outlier (Mahalanobis distance $>$ 16.266) and nine univariate outliers ($>$ 1.5*IQR) were deleted. Thus, we performed the two-factor MANOVA with data from 307 participants. We report Pillai's trace instead of Wilk's lambda because of non-normality, weak linearity, and heterogeneous covariance matrices. Because the assumption of homoscedasticity was not met, single-factor Welch-ANOVAs were used as a follow-up with the factor communication or decision situation as the independent variable and decision certainty, decision help, or decision dependency as the dependent variables. Welch-ANOVAS were followed up with Games-Howell post-hoc tests. Multiple testing was Bonferroni corrected.

Furthermore, we calculated Spearman's correlations between decision certainty, decision help, and decision dependency. We used Spearman's correlation instead of Pearson's correlation because of non-normality and weak linearity. Multiple testing was Bonferroni corrected.

IV. Results

With the two-factor MANOVA, no significant differences were found for the factor communication (V $<$ 0.01; F(3.00; 299.00) = 0.39; p = 0.761; ηp2 = 0.04), but for the factor decision situation (V = 0.39; F(6.00; 600.00) = 24.17; $p < 0.001$; ηp2 = 0.20). Furthermore, no interaction effect between both factors was observed (V = 0.04; F(6.00; 600.00) = 1.94; p = 0.073; ηp2 = 0.02).

The Welch-ANOVAS confirmed significant differences for decision certainty (F(2; 201.89) = 8.06; $p = 0.001$, ω2 = 0.04), decision help (F(2; 200.81) = 78.31; $p < 0.001$, ω2 = 0.31), and decision dependency (F(2; 170.77) = 77.12; $p < 0.001$, ω2 = 0.27) due to the decision situation.

Games-Howell post-hoc tests showed significant difference for decision certainty (ΔM = 13.04; 95% CI[4.70; 21.38]; $p = 0.008$), and decision help (ΔM = 19.29; 95% CI[10.93; 27.66]; $p < 0.001$) when comparing non-moral and somewhat moral, for decision help (ΔM = 39.45; 95% CI[20.93; 27.66]; $p < 0.001$), and decision dependency (ΔM = 42.85; 95% CI[33.65; 52.05]; $p < 0.001$) when comparing non-moral and very-moral decision situations, and for decision help (ΔM = 20.16; 95% CI[11.71; 28,61]; $p < 0.001$), and decision dependency (ΔM= 28.30; 95% CI[19.65; 36.94]; $p < 0.001$) when comparing somewhat moral and very-moral decision situations. These results are visualized in Figure 1.

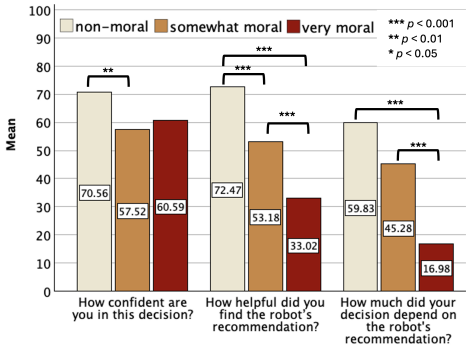

Fig. 1. Visualized results of the Games-Howell post hoc tests

How confident participants are in their decision did not correlate with how helpful the robot's recommendation was perceived (Spearman's p = 0.10, p = .224) or how much participants' relied on the robot's recommendation Spearman's p = -0.10, p = .225), but how helpful the robot's recommendation was perceived and how much participants' relied on the robot's recommendation correlated strongly (Spearman's p = 0.67, $p < .001$).

V. Discussion

In this study, we investigated how the verbal communication style of a robot affects its acceptance. We considered (1) how confident the person is in their decision, (2) how helpful the person finds the robot's recommendation, and (3) how much the person made their decision based on the robot's recommendation (cf. EQ1-3, p.2). The main result of our study is that the morality of the decision situation has a greater influence on the acceptance of the robot's recommendation than the verbal communication of the robot.

We found no differences between functional and social communication. In both communication styles, participants were similarly confident in their decision, found the robot's recommendation similarly helpful, and made their decision similarly dependent on the robot's recommendation. This could indicate that both affective and logical persuasion strategies are comparably effective in robots [26] and that, unlike [15], there is no superiority of affective robot communication. Perhaps differences would be found when comparing other communication styles. Maybe polite and non-polite [9] or direct and indirect [10] language differ more in their effectiveness than functional and social communication. It could also mean that the influence of verbal communication alone is too small to make a difference. Possibly the verbal

communication of robots must be combined with other non-verbal communication features in order to have an effect, as done in in previous studies [6], [17].

Although no differences could be observed between the two communication styles, the study highlights the impact of decision situation characteristics and demonstrates the importance of context consideration. We observed a decrease in acceptance as the decision situation becomes more morally challenging. This supports that people are averse to machines making moral decisions [4]. Our results show that this also applies to robo-advisors. The more moral the decision-making situation becomes, the stronger the aversion.

The lack of correlation between decision certainty and decision help or dependence is interesting. Contrary to [27]'s assumption that uncertainty influences conformity towards robots, we found no evidence that insecure participants found the robot's recommendation more helpful or depended more on it. However, the strong correlation between decision help and dependence is expected, as robots are more likely to be accepted if perceived as helpful.

To conclude, a robot's communication style did not affect decision certainty, perceived helpfulness, or reliance, but all three were strongly influenced by the morality of the situation, showing more algorithm aversion in moral contexts.

ACKNOWLEDGMENT

Clarissa Sabrina Arlinghaus attending the IEEE RO-MAN conference was supported by SAIL. SAIL is funded by the Ministry of Culture and Science of the State of North Rhine-Westphalia under the grant no NW21-05A.

REFERENCES

[1] B. J. Dietvorst, J. P. Simmons, and C. Massey, Algorithm aversion: People erroneously avoid algorithms after seeing them err, Journal of Experimental Psychology: General, vol. 144, no. 1, pp. 114–126, 2015. https://doi.org/10.1037/xge0000033

[2] C. Larkin, C. Drummmond Otten, and J. Árvai, Paging Dr. JARVIS! Will people accept advice from artificial intelligence for consequential risk management decisions?, Journal of Risk Research, vol. 25, no. 4, pp. 407-422, 2022. https://doi.org/10.1080/13669877.2021.1958047

[3] C. Longoni, A. Bonezzi, C. K. Morewedge, Resistance to medical artificial intelligence, Journal of Consumer Research, vol. 46, no. 4, pp. 629–650, 2019. https://doi.org/10.1093/jcr/ucz013

[4] Y. E. Bigman, and K. Gray, People are averse to machines making moral decisions, Cognition, vol. 181, pp. 21–34, 2018. https://doi.org/10.1016/j.cognition.2018.08.003

[5] L. Kunold, and L. Onnasch, A framework to study and design communication with social robots, Robotics, vol. 11, no. 6, 129, 2022. https://doi.org/10.3390/robotics11060129

[6] G. Maggi, E. Dell'Aquila, I. Cucciniello, and S. Rossi, S., "Don't get distracted!": The role of social robots' interaction style on users' cognitive performance, acceptance, and non-compliant behavior, International Journal of Social Robotics, vol. 13, no. 8, pp. 2057–2069, 2021. https://doi.org/10.1007/s12369-020-00702-4

[7] C. S. Arlinghaus, A. Dix, C. Straßmann, S. A. Pertuz, A. Podlubne, D. Göhringer, and S. Pannasch, Influence of robot's social communication on robot's evaluation and human decision making behavior in non-moral, somewhat moral, and very moral decision situations, 2022. https://doi.org/10.17605/OSF.IO/7G8K2

[8] C. S. Arlinghaus, C. Straßmann, and A. Dix, Increased morality through social communication or decision situation worsens the acceptance of robo-advisors. Preprint. https://doi.org/10.31219/osf.io/bufjh

[9] N. Lee, J. Kim, E. Kim, and O. Kwon, The influence of politeness behavior on user compliance with social robots in a healthcare service setting, International Journal of Social Robotics, vol. 9, pp. 727–743, 2017. https://doi.org/10.1007/s12369-017-0420-0

[10] S. Saunderson, and G. Nejat, Robots asking for favors: The effects of directness and familiarity on persuasive HRI, IEEE Robotics and Automation Letters, vol. 6, no. 2, pp. 1793-1800, 2021. https://doi.org/10.1109/LRA.2021.3060369

[11] A. C. Horstmann, N. Bock, E. Linhuber, J. M. Szczuka, C. Straßmann, and N. C. Krämer, Do a robot's social skills and its objection discourage interactants from switching the robot off?, PLoS ONE, vol. 13, no. 7, e0201581, 2018. https://doi.org/10.1371/journal.pone.0201581

[12] D. L. Johanson, H. S. Ahn, and E. Broadbent, Improving interactions with healthcare robots: A review of communication behaviours in social and healthcare contexts, International Journal of Social Robotics, vol. 13, no. 8, pp. 1835–1850, 2021. https://doi.org/10.1007/s12369-020-00719-9

[13] I. Saltik, D. Erdil, and B. A. Urgen, Mind perception and social robots: The role of agent appearance and action types, in Companion of the 2021 ACM/IEEE International Conference on Human-Robot Interaction (HRI 2021), 2021, pp. 210-214. https://doi.org/10.1145/3434074.3447161

[14] P. Formosa, and M. Ryan, Making moral machines: Why we need artificial moral agents, AI & Society, vol. 36, no. 3, pp. 839–851, 2021. https://doi.org/10.1007/s00146-020-01089-6

[15] S. Saunderson, and G. Nejat, Investigating strategies for robot persuasion in social human–robot interaction, IEEE Transactions on Cybernetics, vol. 52, no. 1, pp. 641–653, 2022. https://doi.org/10.1109/tcyb.2020.2987463

[16] P. H. Kahn, T. Kanda, H. Ishiguro, B. T. Gill, J. H. Ruckert, S. Shen, H. E. Gary, A. L. Reichert, N. G. Freier, and R. L. Severson, Do people hold a humanoid robot morally accountable for the harm it causes?, in Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction (HRI 2012), 2012, pp. 33-40. https://doi.org/10.1145/2157689.2157696

[17] Y. Takahashi, Y. Kayukawa, K. Terada, and H. Inoue, Emotional expressions of real humanoid robots and their influence on human decision-making in a finite iterated prisoner's dilemma game, International Journal of Social Robotics, vol. 13, pp. 1777–1786, 2021. https://doi.org/10.1007/s12369-021-00758-w

[18] S. Aboulenine, Persuasion in president Biden's inauguration speech, Traduction et Languages, vol. 20, no. 1, pp. 186-208, 2021.

[19] H. Pfaff, and J. Braithwaite, A Parsonian approach to patient safety: Transformational leadership and social capital as preconditions for clinical risk management — the GI factor, International Journal of Environmental Research and Public Health, vol. 17, no. 11, 3989, 2020. https://doi.org/10.3390/ijerph17113989

[20] M. R. Fraune, S. Sabanovic, and E. R. Smith, Some are more equal than others, Interaction Studies, vol. 21, no. 3, pp. 303–328, 2020. https://doi.org/10.1075/is.18043.fra

[21] P. S. C. Dautzenberg, G. M. I. Voß, S. Ladwid, and A. M. Rosenthal-von der Pütten, Investigation of different communication strategies for a delivery robot: The positive effects of humanlike communication styles, in Proceedings of the 30th IEEE International Conference on Robot & Human Interactive Communication (RO-MAN 2021), 2021, pp. 356-361. https://doi.org/10.1109/RO-MAN50785.2021.9515547

[22] G. Stephan, Amortisationsrechnung in der Radiologie, Radiologen WirtschaftsForum, vol. 9, no. 7-8, 2021.

[23] N. D. Starr, B. Malle, and T. Williams, T., I need your advice… Human perceptions of robot moral advising behaviors, ArXiv, 2021. https://doi.org/10.48550/arXiv.2104.06963

[24] J. Savulescu, Good reasons to vaccinate: mandatory or payment for risk?, Journal of Medical Ethics, vol. 47, no. 2, pp. 78–85, 2021. https://doi.org/10.1136/medethics-2020-106821

[25] W. Buckwalter, and A. C. Peterson, Public attitudes toward allocating scarce resources in the COVID-19 pandemic. PLoS ONE, vol. 15, no. 11, e0240651, 2020. https://doi.org/10.1371/journal.pone.0240651

[26] S. Saunderson, and G. Nejat, G., It would make me happy if you used my guess: Comparing robot persuasive strategies in social human–robot interaction, IEEE Robotics and Automation Letters, vol. 4, no. 2, pp. 1707–1714, 2019. https://doi.org/10.1109/lra.2019.2897143

[27] L. Masjutin, J. Laing, and G. W. Maier, Why do we follow robots? An experimental investigation of conformity with robot, human, and hybrid majorities, in Proceedings of the 2022 ACM/ IEEE International Conference on Human-Robot Interaction (HRI 2022), 2022, pp. 139-146. https://doi.org/10.1109/HRI53351.2022.9889675