Robot-stated limitations but not intentions promote user assistance

David Cameron[1] and **Ee Jing Loh**[2] and **Adriel Chua**[2] and **Emily Collins**[1] and
Jonathan M. Aitken[1] and **James Law**[1]

Abstract. Human-Robot-Interaction (HRI) research is typically built around the premise that the robot serves to assist a human in achieving a human-led goal or shared task. However, there are many circumstances during HRI in which a robot may need the assistance of a human in shared tasks or to achieve goals. We use the ROBO-GUIDE model as a case study, and insights from social psychology, to examine how a robots personality can impact on user cooperation. A study of 364 participants indicates that individuals may prefer to use likable social robots ahead of those designed to appear more capable; this outcome reflects known social decisions in human interpersonal relationships [7]. This work further demonstrates the value of social psychology in developing social robots and exploring HRI.

1 INTRODUCTION

The use of autonomous, mobile service robots in the workplace is expected to grow substantially in the coming years [16]. Robots are anticipated to work side-by-side with people, assisting or collaborating with employees on a variety of tasks. As a consequence, Human-Robot Interaction (HRI) research typically explores interactions based around a robot in a supportive or assistive role for the human user [14].

Current research in assistive robotics highlights the importance of ensuring users' trust that the robot can provide effective assistance. In scenarios of robot-assisted navigation, this may require: near-immediate user trust placed in a robot guide [28]; maintenance of appropriate levels of user trust, so neither over-reliance nor under-use occurs [19] [11]; and recovery of user trust after mistakes in automation [20, 15]. Across the literature, user trust is typically associated with a robot establishing its capacity in meeting goals as an autonomous agent.

However, there are instances in HRI where an assistive robot guide may *require* user support to achieve its aims or complete shared tasks. For example, a robot may hold incomplete information about its situation and require user input (e.g. asking for its location [29] or for directions [2]) or face a task requiring manual intervention by a user (e.g. autonomous mobile robots encountering physical barriers to progress [5, 29]). At this juncture, a robot cannot effectively demonstrate its capacity to operate as an autonomous agent and so typical channels for engendering user trust may be at risk. Robots requiring, rather than solely providing, user assistance in complex and social environments is an emerging topic [4] and so it becomes essential to study its impact on users' experience of HRI.

Research on the topic of robots requiring assistance identifies effective means for robots to determine their limits and when human intervention is required [29], or how to locate users that can offer assistance [30]. However, determining effective means of *how* socially adaptive robots request help to encourage user cooperation and assistance is still a challenge in HRI.

To approach this challenge, we draw from social psychological models exploring cooperation between agents. Social psychological insights can be beneficial in exploring HRI, given it is a novel and still-developing research topic [10]. Robots need to engage the user, particularly those that require user input (either directly through interacting with the robot or shaping the environment to meet the robot's needs). In particular, the emerging field of social robotics considers the optimization of a robot's morphology [16, 3] and even its 'personality' [12, 6] as important to user perceptions of experienced HRI.

In this paper, we compare the impact of two different robot personalities on individuals' willingness to use a robot requiring assistance. We construct a friendly robot personality that emphasises its *limitations* and a capable personality that emphasises its *intentions* in interactions. We further explore the impact that these personality features have on three factors identified in social psychology as having impact on human-human cooperation and assistance: *liking, trust, and ambiguity*. We test a model of robot personality; factors of liking, trust, and ambiguity; and individuals' willingness to help a robot requiring assistance in an applied HRI scenario.

1.1 Pathways to cooperation

The following paragraphs introduce the three factors antecedent to cooperation investigated in this experiment. We target specifically the under-explored point of interaction: in which a robot needs human assistance, within the broader context of it operating as an assistive robot in a workplace environment.

1.1.1 Liking

Future robotic systems are considered likely to operate more as teammates rather than tools [25]. Understanding the impact this change could have on HRI is therefore vital to adapt work-forces to accommodate future, robotic teammates. Occupational psychology models of teamwork and cooperation in the workplace may provide a good foundation for understanding HRI with these robotic teammates. Research from occupational psychology indicates that individuals prefer to populate their cooperative working networks with people they like, ahead of those more capable in the cooperative role (but not liked) [7, 33]. Liking of individuals is considered to play a substantial role in motivations for cooperation and, particularly as the relationship develops, displace antecedent cognitive motivations for cooperation [24].

[1] Sheffield Robotics, University of Sheffield, UK, email: {d.s.cameron, e.c.collins, jonathan.aitken, j.law}@sheffield.ac.uk
[2] Department of Psychology, University of Sheffield, UK email: {ejloh2, dxachua1}@sheffield.ac.uk

1.1.2 Trust

A prominent model from occupational psychology of interpersonal cooperation in teammates identifies trust as its foundation [22]. McAllister argues that to successfully achieve in a task requiring two agents working together, both need to trust each other. Specifically, cooperation is thought to benefit from both *affective trust* (built by personable interactions from the partner) and *cognitive trust* (built from evidence that one's partner carries out responsibilities competently) [22]. Analogues for both forms are seen in HRI [14]. A user's perception of a robot's performance (analogous to cognitive trust) and a user's perception of a robot's attributes, such as personality (analogous to affective trust), are found to positively contribute towards user trust in robots [14].

1.1.3 Ambiguity

Classic social psychology research indicates that ambiguity in assistive scenarios results in substantial detriment to individuals' proactive assistive behavior [9]. For many, HRI in cooperative environments may currently be entirely novel. As a result, individuals could face ambiguity in HRI situations and, without an indication of a robot's limitations, uncertainty regarding *whether* the robot requires user cooperation. Alternatively, lack of a clear plan or intention communicated by a robot may create further ambiguity in *how* the individual may best cooperate with the robot, limiting the action taken.

2 Scenario

To explore how a robot's personality may influence the above factors and user cooperation and assistance towards a robot, it is essential to consider an interactive social robotics scenario in which these circumstances arise. The ROBOtic GUidance and Interaction DEvelopment (ROBO-GUIDE) project [18] offers an ideal scenario to explore these factors. ROBO-GUIDE is implemented on the Pioneer LX mobile platform to autonomously navigate a multistory building, leading users (visitors to the building) to their chosen destination.

Critically, there are elements in the tour which require user intervention to remove barriers to progress, such as using an elevator to navigate between floors. While the ROBO-GUIDE platform can identify which floor it is on [23], it is currently unable to call for an elevator itself: it can neither manually operate elevator buttons nor remotely command elevator operation. As a result, the robot relies on user cooperation to press buttons to call the elevator and select the required floor to progress. To direct the user, statements of the robot's limitations or intentions are communicated using the onboard speech synthesizer.

Our focus for this study is the point at which the platform changes between floors as it navigates the multistory building. Use of ROBO-GUIDE requires users to place trust in an autonomous way-finding robot, whilst also assisting the robot in overcoming obstacles or barriers to progress. For both the user and the robot, this scenario is identified as a simple and low-risk circumstance in which a human user can act in order to meet a robot's needs [5].

2.1 Promoting user cooperation in HRI

We identify how scenario-specific robot-stated limitations or intentions can influence factors for user cooperation and assistance. The following sub-sections show proposed control-condition statements or requests in quotes and normal font, whereas supplemental, experimental statements are in the same quotes and italics.

2.1.1 Liking

ROBO-GUIDE's primary purpose is to lead new visitors to the robotics laboratory. It is anticipated that user liking of the robot would be greater with supplemental phrases regarding its offering 'face-to-face' direct assistance to the human user. To promote *liking*, we include friendly and relatable references by the robot that is assisting the user as a tour guide: "Please follow me; *I am here as your tour guide*".

2.1.2 Trust

It is anticipated that developing *affective trust* overlaps substantially with developing user liking [22] and can be built by demonstrating trust in others [21, 32]. To promote affective trust, we supplement requests from the robot for user cooperation at the elevator with identification of its limitations: "Please press the down button; *I cant quite reach the buttons*".

In contrast, *cognitive trust* is developed through demonstration of an agent's competency in meeting its intended and/or required responsibilities [22]. It is anticipated that users' cognitive trust of the robot's way-finding capabilities would be greater with the inclusion of additional phrases that directly address its intended aims: "Please follow me *to get to the robot labs*".

2.1.3 Ambiguity

It is anticipated that many visitors would be unfamiliar with HRI or social robots and could find the experience unusual. Interacting with a novel robot, especially a robot that needs help, is anticipated to create ambiguous situations for users. To reduce ambiguity, we use friendly and relatable references by the robot about its limitations, which a user could help with in the HRI scenario: "Please press ground floor; *good thing you're here to do that for me*".

It is further anticipated that ROBO-GUIDE's statements of intentions when requesting help will reduce ambiguity and promote cooperation. Without declaring why the tasks are to be completed, requests for help could be ambiguous in their purpose. Intentions should demonstrate the robot has a clear goal it is trying to achieve but that it is now facing an obstacle and so asks for help: "Please press the down button *to call the lift*" and "Please press ground floor *for the Robot Labs*".

2.1.4 A model of user cooperation

We predict that a robot stating its limitations and intentions when requesting assistance will promote individuals' willingness to cooperate with assisting the robot. Pathways by which this is anticipated to occur, factors of liking, trust, and ambiguity, are over-viewed in Figure 1.

The robot stating its limitations is predicted to: promote user liking, promote users' affective trust towards the robot, and limit ambiguity in the situation. The robot stating its intentions is predicted to: promote users' cognitive trust towards the robot and limit ambiguity in the situation. These outcomes are in turn anticipated to positively impact on users' perceptions of assisting the robot

3 METHOD

3.1 Design

A 2x2 independent measures design was implemented. The four conditions comprised of the presence or absence of key statements re-
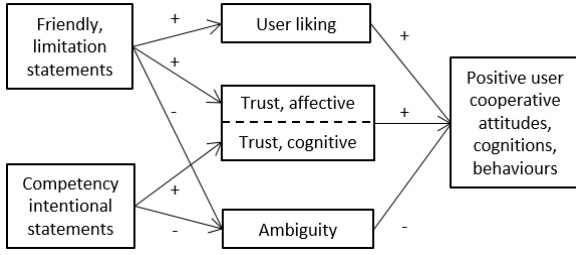
Figure 1. Pathways by which robot-stated limitations and intentions impact on user-cooperation

garding i) the robot's limitations, and ii) intentions. Participants were randomly allocated to a single condition and the Qualtrics survey engine prohibited repeat participation.

3.2 Materials

3.2.1 Videos

Four HRI videos were prepared; one for each condition. The videos represented a typical use of the ROBO-GUIDE: an individual is greeted by the robot at the building entrance and instructed to follow it. The robot and user travel along a corridor to an elevator, at which point the robot instructs the user to press the call button (Example of HRI in Figure 2). On entering the elevator, the robot again instructs the user to press the relevant button for the floor. The robot and user then leave the elevator at the target floor and travel along another corridor to the robot-lab - the final destination.

All four films have minimal visual differences as non-critical scenes were used across conditions. Critical scenes differ in audio (i.e. the condition-specific words spoken by the robot) and subtitles of the robot's speech. Videos lasted 90 seconds.
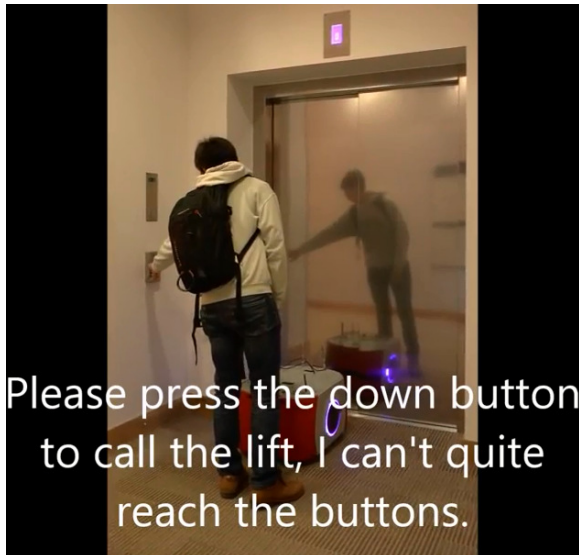

Figure 2. Example of user assisting robot in HRI (Intentions and Limitations condition)

3.2.2 Questionnaires

Perspectives on the observed HRI were assessed using the Godspeed Questionnaire [1]. Measurements of liking were taken using the relevant Godspeed sub-scale. In addition, 3 ad-hoc measures were developed to assess participants' perspectives of the trustworthiness of ROBO-GUIDE, ambiguity of the interaction, and their likelihood of using ROBO-GUIDE.

Trustworthiness was assessed using an ad-hoc 16-item scale derived from words previously identified as strongly associated with trust or mistrust in automated systems [17]. Each of the following items was scored on a 7-point scale from *not at all* to *extremely*, headed '*To what extent would you describe ROBO-GUIDE as...*' *capable, competent, confident, deceptive, false, honest, honorable, incapable, incompetent, loyal, misleading, reliable, trustworthy, unreliable, unsure, untrustworthy*.

Ambiguity was assessed using a 4-item measure (alpha = 0.93) concerning the clarity of the robot's requests regarding, *why*, *whether*, *when* and *how* a user may need to assist the robot. Each item was scored on a 7-point scale from *strongly agree* to *strongly disagree*.

Likelihood of use was assessed using a 4-item measure (alpha = 0.78), with each item scored on a 7-point scale from *strongly agree* to *strongly disagree*. Items included: perceiving the robot as being fun to use, and intentions to use the robot in an unfamiliar building.

3.3 Participants

Participants were recruited through staff and student university volunteer mailing lists; 442 participants signed up for the study. A survey timer indicated that 78 participants did not fully watch the HRI video stimulus so were excluded from the study. Of the remaining 364 participants: 196 were female, 152 were male, and 16 declined to identify gender; 240 participants were British, 114 were from overseas, 10 declined to identify nationality; aged M = 27.67, SD = 8.99. Participants were given the opportunity to win one of two 50 gift vouchers as recompense for their time.

3.4 Procedure

The survey was distributed through online mailing lists to reach a broad audience and delivered through the online survey engine Qualtrics (Version 06.2015; Copyright 2015; Qualtrics, Utah). Participants were first presented with a study information and consent page. On agreeing to participate, individuals were randomly allocated to one of the four conditions described above and presented with an HRI video. Following the video, participants completed the Godspeed questionnaire, measures rating trustworthiness of ROBO-GUIDE, interaction ambiguity, and willingness to use the ROBO-GUIDE. At the close of the questionnaire, participants could share demographic details and register for a chance to win offered gift vouchers. The study took participants approximately 10 minutes to complete.

4 RESULTS

4.1 Randomisation Check

Participant numbers were evenly distributed across conditions $x^2(1, N = 364) = 0.01$, $p = 0.92$. For each condition, there were even distributions of participants in terms of gender and nationality (max $x^2(1, N = 354) = 3.31$, $p = 0.07$).

4.2 Primary results

4.2.1 Liking

There was a significant main effect for the robot stating its limitations $F(1,356) = 53.407$, $p < 0.01$. Participants who saw the interactions in which the robot stated its limitations reported liking the robot to a substantially greater extent (M = 3.77, S.E. = 0.05) than those who saw control statements (M = 3.24, S.E. = 0.05). This is a large effect observed (d = 0.77). There was also a significant main effect for the robot stating its intentions $F(1,356) = 7.25$, $p < 0.01$. Participants who saw the interactions in which the robot stated its intentions reported liking the robot to a substantially greater extent (M = 3.60, S.E. = 0.05) than those who saw control statements (M = 3.40, S.E. = 0.05). This is a small effect observed (d = 0.28). There was no significant interaction effect between the robot stating both its limitations and intentions on participants' liking $F(1,356) = 0.87$, $p = 0.79$.

4.2.2 Trust

The 16 items for trust were subjected to a principal axis factor analysis for the full sample of the 364 participants; missing values were treated pairwise. Three factors with Eigenvalues greater than 1.00 [8] were extracted from the matrix, explaining 59% of the variance. Inspection of the pattern matrix showed that all items loaded above 0.40 on one of the three factors. Items on the first factor are mainly related to falsity: *deceptive, false, misleading, unreliable, unsure,* and *untrustworthy.* Items on the second factor are mainly concerned with affective trust: *honest, honorable, loyal,* and *trustworthy.* Finally, items on the third factor are mainly concerned with cognitive trust *capable, competent, confident, reliable,* and two negatively scored items of *incapable* and *incompetent.*

However, there were no significant main effects for the robot stating its limitations nor intentions for either factors of affective trust (Limitations $F(1,351) = .11$, $p = .75$; Intentions $F(1,351) = .21$, $p = .65$) or cognitive trust (Limitations $F(1,351) = 2.14$, $p = 0.14$; Intentions $F(1,351) = .01$, $p = .98$). Furthermore, there were no significant interaction effects.

4.2.3 Ambiguity

There was a significant main effect for the robot stating its limitations $F(1,352) = 134.95$, $p < 0.01$. Participants who saw the interactions in which the robot stated its limitations reported substantially less ambiguity in the situation (M = 1.87, S.E. = 0.10) than those who saw control statements (M = 3.52, S.E. = 0.10). This is a large effect observed (d = 1.22). There was no main effect for the robot stating its intentions on participants' perceptions of ambiguity in the interaction $F(1,352) = 0.43$, $p = 0.52$.

There was a significant interaction effect between the robot stating both its limitations and its intentions $F(1,352) = 4.72$, $p = 0.03$. The robot solely stating its intentions promoted greater perceptions of ambiguity in comparison to control statements (M = 3.72, SE = 0.14 versus M = 3.32, SE = 0.14). However, in combination with the robot stating its limitations, participants' ratings of ambiguity were lower in comparison to the robot solely stating limitations only (M = 1.76 SE = 0.14 versus M = 1.98, SE = 0.14). Simple main effects analysis showed that the robot-stated intentions resulted in more ambiguity than control statements, when presented without robot-stated

limitations (p = .04), but there were no differences when accompanied with robot-stated limitations (p = .28). These results are presented in Figure 3.
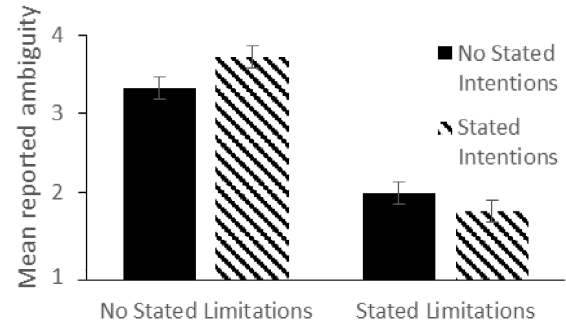

Figure 3. Interaction effects for intentions and limitations for user perceived ambiguity

4.2.4 Willingness to use

There was a significant main effect for the robot stating its limitations $f(1,352) = 8.74$, $p < 0.05$. Participants who saw the interactions in which the robot stated its limitations reported greater willingness to use the robot (M = 4.57, SE = 0.11) than those who saw control statements (M = 4.26, SE = 0.11). There was no significant main effect for the robot stating its intentions on participants' willingness to use the robot $F(1,352) = 1.07$, $p = 0.30$. There was no significant interaction effect between the robot stating both its limitations and intentions on participants' willingness to use the robot $F(1,352) = 0.09$, $p = 0.86$.

Mediation analysis was run to determine if the main effect of the robot-stated limitations promoting individuals' willingness to use ROBO-GUIDE could be explained by the effects seen above for robot-stated limitations on user liking and ambiguity[3]. Indirect effects were computed following 1,000 bootstrapped samples at 95% confidence intervals. The positive relationship between robot-stated limitations and individuals' willingness to use ROBO-GUIDE was mediated by user liking but not ambiguity (see Table 1). Figure 4 identifies the pathways examined in the mediation analysis and coefficients between constructs.

[htbp]

Table 1. Mediators for robot-stated performance limitations on individuals' willingness to use ROBO-GUIDE

Dependent variable	Mediators	Point estimate effects	95% confidence intervals	
			Lower	Upper
Willingness to Use	Liking	0.56	0.39	0.74
	Ambiguity	0.12	-0.04	0.031
	Total	0.31	0.01	0.62
$R^2 = 0.27$, $F(3,352) = 42.54$, p < 0.01				

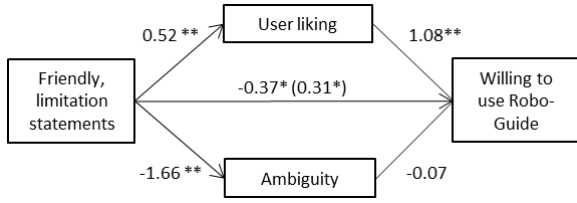[3] User trust is not included as there were no main effects observed.

Figure 4. Regression coefficients for the relationship between robot-stated limitations and individuals' willingness to use ROBO-GUIDE as mediated by user liking and ambiguity. The regression coefficient between robot stated limitations and willingness to use ROBO-GUIDE, controlling for user liking, is in parenthesis (* p < 0.5; ** p < 0.01)

4.3 Demographics

There were significant main effects of nationality for both participants' reports of liking the robot $F(1,343) = 6.03$, $p = 0.02$ and their perceptions of the robot as being trustworthy $F(1,342) = 21.68$, $p < 0.01$. Individuals from overseas reported liking the robot more (M = 3.66, SE = 0.07) than those from the UK did (M = 3.45, SE = 0.05) and showed more affective trust towards the robot (M = 4.50, SE = 0.12), than those from the UK did (M = 3.81, SE = 0.09). These are small and medium effects (d = 0.28; d = 0.53) respectively. There were no significant main effects for gender (Max $F(1, 354) = 2.58$, $p = 0.11$).

There were significant interaction effects between nationality and gender for both perceptions of robot competency $F(1, 343) = 5.83$, $p = 0.02$ and willingness to use such a robot $F(1,343) = 7.80$, p 0.01. In both cases, female participants from the UK scored higher (M = 5.31, SE = 0.08; M = 3.96, SE = 0.13) than female participants from overseas (M = 4.94, SE= 0.13; M = 3.60, SE = 0.21), respectively; whereas, male participants from overseas scored higher (M = 5.26, SE = 0.12; M = 4.12, SE = 0.20) than male participants from the UK (M = 5.11, SE = 0.10; M = 3.48, SE = 0.16), respectively.

5 Discussion

The results demonstrate a substantial difference in individuals' responses to observing a friendly versus capable social robot. The contrast between a friendly, limitations-focused, and capable, intentions-focused, robot is seen in individuals' willingness to use the robot in the future; only the former significantly differs from the control condition. Moreover, the observed effect can be explained by the mediating influence of individuals liking a limitations-focused robot. As observed in occupational psychology [7], individuals prefer to further interact with those they like, rather than those presented as being more capable. This study provides new evidence that strategies used by individuals in interpersonal relationship development can be extended to apply to social robotics HRI.

One key strength of the current work is the available sample size, which would require an extensive effort to produce in a field robotics study. The large sample size enabled a sufficiently powered 2x2 design study and a full factor analysis of the ad-hoc trust scale. Again, the factor analysis indicates that it is extremely worthwhile exploring HRI in terms of social psychological models of interpersonal relationships. Factors extracted from the user responses closely corresponded to the constructs of affective and cognitive trust [22]. Participants may be applying their own understanding of human-human social relationships to novel contexts (HRI) containing social agents.

There are several outcomes apparent due to individual differences, such as an interaction effect of gender and nationality on willingness to use ROBO-GUIDE. This emphasises the importance of human-focused study design and robotic development [4]. The demographic differences in individuals' responses indicate that developing a socially adaptive robot responsive to users, rather than attempting a 'one-size-fits-all' approach could greatly benefit user perceptions of HRI. In terms of application to the ROBO-GUIDE scenario, this could comprise of adjusting synthetic personalities following a user's feedback in preparation for their next use of ROBO-GUIDE.

5.1 Further Directions

The promising findings so far offer many directions to further examine user assistance in HRI.

The scenario chosen demonstrates a user assisting a robot at minimal cost to themselves. While the explored factor of ambiguity is important in determining proactive assistive behaviour [9], relative cost or risk to the individual is also a key factor [26]. The present work indicates the two robot personalities both reduce ambiguity in interaction. Alternative scenarios (such as in manufacturing), in which a user must expend greater effort to assist a robot, may further see changes in the impact of ambiguity on a user's willingness to assist.

A complex social interaction between user and robot, such as the scenario presented, can be further explored across a variety of social dimensions and factors not-yet considered. Empathy towards robots in need or at risk [31] may contribute to user willingness to assist. This may impact on the current pathways identified, particularly towards a friendly robot or for individuals more likely to anthropomorphise the robot [27]. Alternatively, the interaction could be further considered in a broader, social context, rather than the applied, occupational psychology background for this workplace-inspired interaction scenario and draw from related social cognition literature [13].

Further revisions to the scenario could include altering the robot's behavioural competency. It is possible that the robot performing effectively throughout the task (elevator operation not-withstanding) places a ceiling on individuals' trust towards the robot and limits the impact of the experimental conditions. As identified in the literature [14], errors from the robot impinge on user trust. Introductions of errors (such as departing the elevator on the wrong floor) may lower baseline user trust, enabling any potential influence from limitation or intentional statements. This could further develop work examining the impact of robots addressing their mistakes on improving user HRI experiences [34], again suggesting the importance of user liking ahead of competency in social robots.

This study forms a solid foundation to conduct a field experiment exploring user behaviour in the HRI scenario. Direct measures of user behaviour, such as human-robot 'interpersonal' distance [35] and user speed of response to robot requests, would build on the established results of willingness to engage, liking of, and ambiguity felt towards the social robot.

Our findings have implications for the development of social robots and study of social HRI. The study demonstrates that individuals may apply known human-human interpersonal social decisions to social robots. The careful development of robot personalities to account for this may be instrumental in fostering positive user experiences of social robots and effective HRI.

ACKNOWLEDGEMENTS

REFERENCES

[1] Christoph Bartneck, Dana Kulić, Elizabeth Croft, and Susana Zoghbi, 'Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots', *International journal of social robotics*, **1**(1), 71–81, (2009).

[2] Andrea Bauer, Klaas Klasing, Georgios Lidoris, Quirin Mühlbauer, Florian Rohrmüller, Stefan Sosnowski, Tingting Xu, Kolja Kühnlenz, Dirk Wollherr, and Martin Buss, 'The autonomous city explorer: Towards natural human-robot interaction in urban environments', *International Journal of Social Robotics*, **1**(2), 127–140, (2009).

[3] Cynthia Breazeal and Brian Scassellati, 'How to build robots that make friends and influence people', in *Intelligent Robots and Systems, 1999. IROS'99. Proceedings. 1999 IEEE/RSJ International Conference on*, volume 2, pp. 858–863. IEEE, (1999).

[4] David Cameron, Jonathan M Aitken, Emily C Collins, Luke Boorman, Adriel Chua, Samuel Fernando, Owen McAree, Uriel Martinez-Hernandez, and James Law, 'Framing factors: The importance of context and the individual in understanding trust in human-robot interaction', in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Workshop on Designing and Evaluating Social Robots for Public Settings*, (2015).

[5] David Cameron, Emily C. Collins, Adriel Chua, Samuel Fernando, Owen McAree, Uriel Martinez-Hernandez, Jonathan M. Aitken, Luke Boorman, and James Law, 'Help! I cant reach the buttons: Facilitating helping behaviors towards robots', in *Biomimetic and Biohybrid Systems, Living Machines 2015*, volume 9222 of *Lecture Notes in Computer Science*, 354–358, (2015).

[6] David Cameron, Samuel Fernando, Emily Collins, Abigail Millings, Roger Moore, Amanda Sharkey, Vanessa Evers, and Tony Prescott, 'Presence of life-like robot expressions influences childrens enjoyment of human-robot interactions in the field', in *Fourth International Symposium on New Frontiers in Human-Robot Interaction*, 36–42, (2015).

[7] Tiziana Casciaro and Miguel Sousa Lobo, 'Competent jerks, lovable fools, and the formation of social networks', *Harvard Business Review*, **83**(6), 92–99, (2005).

[8] Raymond B Cattell, 'The scree test for the number of factors', *Multivariate behavioral research*, **1**(2), 245–276, (1966).

[9] Russell D Clark and Larry E Word, 'Why don't bystanders help? because of ambiguity?', *Journal of Personality and Social Psychology*, **24**(3), 392, (1972).

[10] Emily C Collins, Abigail Millings, and Tony J Prescott, 'Attachment to assistive technology: A new conceptualisation', in *Proceedings of the 12th European AAATE Conference (Association for the Advancement of Assistive Technology in Europe)*, (2013).

[11] Munjal Desai, Kristen Stubbs, Aaron Steinfeld, and Holly Yanco, 'Creating trustworthy robots: Lessons and inspirations from automated systems', (2009).

[12] Samuel Fernando, Emily C Collins, Armin Duff, Roger K Moore, Paul FMJ Verschure, and Tony J Prescott, 'Optimising robot personalities for symbiotic interaction', in *Biomimetic and Biohybrid Systems*, 392–395, Springer, (2014).

[13] Susan T Fiske, Amy JC Cuddy, and Peter Glick, 'Universal dimensions of social cognition: Warmth and competence', *Trends in cognitive sciences*, **11**(2), 77–83, (2007).

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

[15] Guido Herrmann, Martin Pearson, Alexander Lenz, Paul Bremner, Adam Spiers, and Ute Leonards, *Social Robotics: 5th International Conference, ICSR 2013, Bristol, UK, October 27-29, 2013, Proceedings*, volume 8239, Springer, 2013.

[16] Pamela J Hinds, Teresa L Roberts, and Hank Jones, 'Whose job is it anyway? a study of human-robot interaction in a collaborative task', *Human-Computer Interaction*, **19**(1), 151–181, (2004).

[17] Jiun-Yin Jian, Ann M Bisantz, and Colin G Drury, 'Foundations for an empirically determined scale of trust in automated systems', *International Journal of Cognitive Ergonomics*, **4**(1), 53–71, (2000).

[18] James Law, Jonathan M. Aitken, Luke Boorman, David Cameron, Adriel Chua, Emily C. Collins, Samuel Fernando, Uriel Martinez-Hernandez, and Owen McAree, 'Robo-guide: Towards safe, reliable, trustworthy, and natural behaviours in robotic assistants', in *Towards Autonomous Robotic Systems (TAROS) 2015*, volume 9287 of *Lecture Notes in Computer Science*, 149–154, (2015).

[19] John D Lee and Katrina A See, 'Trust in automation: Designing for appropriate reliance', *Human Factors: The Journal of the Human Factors and Ergonomics Society*, **46**(1), 50–80, (2004).

[20] Erika Mason, Anusha Nagabandi, Aaron Steinfeld, and Christian Bruggeman, 'Trust during robot-assisted navigation', in *2013 AAAI Spring Symposium Series*, (2013).

[21] Roger C Mayer, James H Davis, and F David Schoorman, 'An integrative model of organizational trust', *Academy of Management Review*, **20**(3), 709–734, (1995).

[22] Daniel J McAllister, 'Affect-and cognition-based trust as foundations for interpersonal cooperation in organizations', *Academy of Management Journal*, **38**(1), 24–59, (1995).

[23] Owen McAree, Jonathan M Aitken, Luke Boorman, David Cameron, Adriel Chua, Emily C Collins, Samuel Fernando, James Law, and Uriel Martinez-Hernandez, 'Floor determination in the operation of a lift by a mobile guide robot', in *Proceedings of the European Conference on Mobile Robotics*, pp. 1–6, (2015).

[24] Carolyn Y Nicholson, Larry D Compeau, and Rajesh Sethi, 'The role of interpersonal liking in building trust in long-term channel relationships', *Journal of the Academy of Marketing Science*, **29**(1), 3–15, (2001).

[25] Scott Ososky, David Schuster, Elizabeth Phillips, and Florian G Jentsch, 'Building appropriate trust in human-robot teams', in *2013 AAAI Spring Symposium Series*, (2013).

[26] Irving M Piliavin, Jane A Piliavin, and Judith Rodin, 'Costs, diffusion, and the stigmatized victim.', (1975).

[27] Laurel D Riek, Tal-Chen Rabinowitch, Bhismadev Chakrabarti, and Peter Robinson, 'How anthropomorphism affects empathy toward robots', in *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*, pp. 245–246. ACM, (2009).

[28] Paul Robinette, Alan R Wagner, and Ayanna M Howard, 'Building and maintaining trust between humans and guidance robots in an emergency', (2013).

[29] Stephanie Rosenthal, Joydeep Biswas, and Manuela Veloso, 'An effective personal mobile robot agent through symbiotic human-robot interaction', in *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems: volume 1-Volume 1*, pp. 915–922. International Foundation for Autonomous Agents and Multiagent Systems, (2010).

[30] Stephanie Rosenthal and Manuela M Veloso, 'Mobile robot planning to seek help with spatially-situated tasks.', in *AAAI*, volume 4, p. 1, (2012).

[31] Astrid Marieke Rosenthal-von der Pütten, Frank P Schulte, Sabrina C Eimler, Laura Hoffmann, Sabrina Sobieraj, Stefan Maderwald, Nicole C Krämer, and Matthias Brand, 'Neural correlates of empathy towards robots', in *Proceedings of the 8th ACM/IEEE international conference on Human-robot interaction*, pp. 215–216. IEEE Press, (2013).

[32] Mark A Serva, Mark A Fuller, and Roger C Mayer, 'The reciprocal nature of trust: A longitudinal study of interacting teams', *Journal of Organizational Behavior*, **26**(6), 625–648, (2005).

[33] Ramadhar Singh and Xue Ling Tor, 'The relative effects of competence and likability on interpersonal attraction', *The Journal of social psychology*, **148**(2), 253–256, (2008).

[34] DHJ Snijders, 'Robot's recovery from invading personal space', in *23rd Twente Student Conference on IT*, volume 23. University of Twente, (2015).

[35] Michael L Walters, Kerstin Dautenhahn, René Te Boekhorst, Kheng Lee Koay, Dag Sverre Syrdal, and Chrystopher L Nehaniv, 'An empirical framework for human-robot proxemics', *Procs of New Frontiers in Human-Robot Interaction*, (2009).