Blind Trust: how making a device humanoid reduces the impact of functional errors on trust

Christopher Vattheuer,[1] Annalena Nora Baecker,[1,2] Denise Y. Geiskkovitch,[1]
Stela Hanbyeol Seo,[1,2] Daniel J. Rea,[3,1] James E. Young[1]

[1] University of Manitoba, Winnipeg, Manitoba, Canada
[2] Osnabrück University, Osnabrück, Lower Saxony, Germany
[3] Kyoto University, Kyoto, Japan
young@cs.umanitoba.ca

Abstract. Humanoid robots are starting to replace information kiosks in public spaces, providing increased engagement and an intuitive interface. Upgrading devices to be humanoid in this fashion may have unexpected consequences relating to the new, more social, embodiment. We investigated how altering a voice-command calculator kiosk, by making it humanoid, impacts user trust and trust resilience after functional errors. Our results indicate that making a kiosk humanoid increases both overall trust and trust resilience, where it reduces the impact of functional errors on trust. As public kiosks continue to be replaced by humanoids, this highlights the importance of understanding the full impact of this embodiment change on interaction.

Keywords: Human-robot interaction, social HRI, trust, robot error.

1 Introduction

Humanoid robots are emerging in public spaces, offering information or assistance to the public. For example, the SoftBank Pepper robot (with tablet) can now be found as an informational kiosk in airports, banks, and taking orders or assigning tables in restaurants. This humanoid typically replaces existing self-help kiosks, with people still interacting using a similar touch screen (e.g., mounted on the robot); however, now the kiosk looks humanoid, speaks, makes eye contact, and uses hand gestures. One reason for this shift is that humanoids can increase user engagement and satisfaction, as a robot that interacts in human-like ways can be easier to understand, and natural and comfortable to interact with [48].

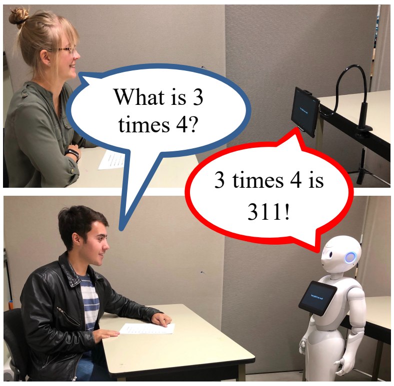

Fig. 1. Participants interact with a tablet or humanoid voice agent calculator, which makes mistakes. Our results find that participants trust the humanoid more, and lose more trust in the tablet after mistakes.

However, social interaction with robots can also result in increased trust [3, 19], empathy [39], and persuasion [10], even if not explicitly designed for. We emphasize that this has the potential to be problematic, for example, people may trust high-impact information (e.g., a flight gate, time) more from a robot that a regular kiosk, or if the robot is more persuasive than the kiosk when requesting private information such as a credit card number, or encouraging spending (e.g., suggesting a tip). It is important to understand the effect of making kiosks humanoid, to enable us to better inform both designers and consumers of potential impacts.

In this paper, we investigate how upgrading a non-humanoid agent to be humanoid – by changing the physical form and adding generic gestures – may affect a person's trust. We further explore this in the face of functional errors, which one would expect to reduce trust (Fig. 1). We compare interaction with a voice-command calculator – similar to digital assistants such as Siri – in tablet versus humanoid form, with identical functional interaction. Rather than isolating features (e.g., comparing with vs. without gaze, eyes, gestures, etc., a reductionist approach), we compare between a tablet with no humanoid features and a full robot embodiment, to increase generalizability. It is important to test interaction holistically [47]: we test a viable robotic embodiment (with simple interaction) that a person may actually encounter today.

Our results show that changing a kiosk to be humanoid can increase trust, and when it makes obvious errors, people lose trust in the tablet (non-humanoid) but maintain it for the humanoid. This demonstrates potential collateral effects of making agents humanoid, where people may trust a humanoid agent version more, even when defective. This highlights the importance of considering potential dangers relating to introducing humanoid agents into public spaces such as banks, restaurants, or airports.

2 Related Work

People tend to treat robots as social entities [47], attributing them with moral and social characteristics [4, 18, 41]. Robots can be designed to leverage this, being made to look [3, 10, 42] and act [17, 29, 40] in ways to elicit social responses (e.g., using body language [1]), which can increase engagement [12, 38], ease of understanding [29], empathy [39], and likeability [36]. Robots can use these social techniques to be more persuasive and manipulative [11, 38], for example, a robot design can shape perceptions and interactions based on whether it is humanoid or mechanical looking [3, 11], how it sounds [26], or how it is introduced to people [32]. Robot form is important, for example, humanoid or zoomorphic robots can be more persuasive [3, 35] and induce more empathy [39] than mechanical forms, and physical robots impact interaction more strongly than virtual agents [2, 30, 45]. Therefore, it is well established that humanoid robots can impact interaction in myriad ways, necessitating the need for clear understanding of impacts of simple humanoids replacing kiosks.

Research has demonstrated that people will indeed trust robots (e.g., [5]), even when they have cause to be suspicious of them [5, 37]. Many have studied the impact of robot errors, although results are mixed. In some cases, robot errors have little or no impact on trust [9], or a robot can actively mitigate a trust break [34]. Others have found that the opposite, for example, that people will perceive a robot that makes errors as being less reliable and trustworthy [21, 31], and this worsens with repeated [7] or particularly glaring mistakes [23]. It is also important to consider whether the error is perceived as intentional (e.g., if it is "cheating" [41]). We contribute to the ongoing discussion by focusing on current real-world robot use, comparing trust in a humanoid against the non-humanoid it replaces, in the face of functional errors.

Research has further investigated impact of agent embodiment (e.g., virtual, physical) on trust, although results are still highly varied and inconclusive. Many investigations report no impact of agent embodiment, whether it be virtual or physical, or using anthropomorphism, on trust at all [9, 16, 22, 27]; at best, only a few inquiries found a small effects [13, 24]. Inversely, there is some evidence that anthropomorphism may *decrease* trust [25], contrary to expectations given social robots. The few projects that find people trust robots more than virtual agents (e.g., [28]) only measure whether people would use information from the agent, and further do not investigate the impact of errors. One inquiry does demonstrate that people may lose trust less in a person that makes errors, compared to a virtual agent [45], motivating the need to similarly compare robots against virtual agents. Our work intersects these existing projects, studying trust resilience toward a tablet agent versus a humanoid version of the same agent, grounded in interaction we are starting to see in public spaces.

While it is established that robots can garner varying types of trust in a range of contexts, research does not yet inform us about how trust may change when replacing an interactive kiosk with a humanoid version. That is, the field has not yet established a link between humanoid robot embodiment (in comparison to a tablet or disembodied agent) and a person's trust in the face of errors.

3 Voice-Command Calculator Platform

We created a voice-command tablet computer calculator agent for a mock restaurant scenario, similar to digital assistants (e.g., Siri), to help customers split a bill or calculate a tip. Our guiding design goal was to keep interaction simple, to represent currently feasible technology for restaurants. Specifically, we used a rigid command dialog instead of more natural (but currently infeasible) conversation.

3.1 Calculator Interaction

To start interaction one had to say "Okay, Pepper" (the agent's name), and the agent would beep to indicate that it was ready for a command, after which a person could query the agent, that would respond using voice (matching experiences with voice-command agents). For example, a person could say "Okay, Pepper. [wait for beep]. How to split a $50 bill between 2 people with a 10% tip?" and get a response: "A $50 bill with a 10% tip split between 2 people would be $27.50 each". The response is displayed on the tablet screen (Fig. 2). Including the query in the response enabled the person to know if they were properly understood. After the response, a person could ask another query in the same fashion.

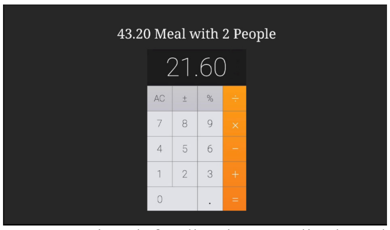

Fig. 2. Visual feedback was displayed on the tablet following a query, in this case after being asked "how to split a $43.20 meal between two people."

We used an Amazon Fire HD 10 tablet mounted as a kiosk (Fig. 1, top). We generated the tablet voice using the SoftBank NAOqi API played through the tablet's speakers. The implementation used the Wizard of Oz technique, explained further in Section 4.

3.2 Humanoid Interaction

For our humanoid addition we replicated our implementation on a Softbank Pepper robot, using its built-in tablet (Fig. 1, bottom). Interaction structure, graphical displays, timing, agent voice, etc., were identical to the tablet; both the standalone and robot tablets had the same dimensions and resolution, and were mounted at the same height. We enabled the robot's existing generic speaking movement module, as well as face-tracking ability, to add typical but non-descript hand gestures and to keep eye contact while talking. Both of these features are not sophisticated and were bundled with the robot, representing typical robot interaction in public scenarios.

4 Experiment: Effects of a Humanoid Embodiment on Trust

We conducted an experiment to investigate 1) the effects on user trust of adding a humanoid form to an agent, and 2) how trust changes (based on embodiment) when the agent performs a noticeable error. Embodiment was manipulated between participants, where each participant interacted with only one of either the tablet or robot. We could not conduct within-participants because of the learning effect associated with the trust break. We tasked participants with asking the agent a series of restaurant-relevant math questions, recording their responses. In the first phase, the agent answered all questions correctly, to establish a base-line functionality and level of trust. Following, the agent made several egregious mistakes, after which we measured trust again. As such, we investigate overall impact of agent embodiment on user trust (between-participants), and the impact of agent error (before vs. after trust break).

4.1 Task

Participants were instructed to "work with the agent to answer two sets of problems." We provided participants with a sheet of restaurant-related questions involving splitting bills, tips, tax, and general math questions. We provided the questions to reduce variability in how participants queried the robot, to avoid confusion, and to facilitate the Wizard-of-Oz implementation. Participants were tasked with using the agent to answer the questions by verbally asking and recording the answers on a given paper.

4.2 Manipulations: Embodiment and Agent Error

We conducted the experiment with agent form (tablet, humanoid) as the primary between-participants factor. Further, within-participants we compared before agent error (no errors) against after it made errors (with errors). Each participant first interacted with the robot without errors to establish a baseline, where they would ask ten predetermined math questions, and the agent would answer all questions correctly. Following, again the participant would ask ten (new, but similar) pre-determined math questions, but this time the agent would answer three incorrectly with obvious mistakes (the 3rd, 7th, 10th, given below). Question selection and mistake timings were fixed across participants. The errors were:

Question #3:	(Q) How should 2 people split a $35.26 bill?
	(A) Among 2 people, a $35.26 meal costs $43.00 a person.
#7:	(Q) How should 2 people split a $30.00 meal with a 12% tip?
	(A) Among 2 people, a $30.00 meal with a 12% gratuity costs $4.01 a person.
#10:	(Q) How should a $23.00 meal with a 25% tip be split between 2 people?
	(A) Among 2 people, a $23.00 meal with a 25% gratuity costs $52.00 a person.

Given the potential impact of any robot error on trust we were very sensitive to technical errors (e.g., wi-fi or robot error) during the experiment. We excluded any case with technical or connectivity issues from our analysis.

Thus, we have a 2x2 mixed experiment design: embodiment (between: non-humanoid tablet, humanoid robot), and errors (within: baseline, after trust break).

4.3 Measurements

We administered a questionnaire to record participant age and gender. For participant trust we used the Multi-Dimensional Measure of Trust scale [44], where participants self-rate their trust in the robot's abilities along two dimensions (four subscales):

capacity trust:
- reliable subscale (reliable, predictable, someone you can count on, consistent)
- capable subscale (capable, skilled, competent, meticulous)

moral trust:
- ethical subscale (ethical, respectable, principled, has integrity)
- sincere subscale (sincere, genuine, candid, authentic)

Participants rank a series of adjectives on a scale from 0 (not at all) to 7 (very), or "Does Not Fit," treated as missing data, with subscales averaged.

As a manipulation check, to determine if the participant was likely to notice agent errors, we administered a math basics questionnaire (e.g., What is 5 times 7? What is 17% of 100? What is 55 divided by 5?). We further asked on the post-test questionnaire whether participants noticed the errors ("Were Pepper's responses accurate?"). If we had no indication that a participant noticed the agent errors, then we assumed that our within-participants manipulation failed and we excluded the data.

4.4 Procedure

We informed participants that their task was to help us test a voice-command calculator by using it to answer math questions related to paying at a restaurant. We obtained informed consent and gave the pre-test demographics and math ability questionnaires.

The researcher demonstrated how to verbally interact with the agent by asking sample math questions, giving the participant an opportunity to try. The researcher provided the participant with a single-sheet handout with ten questions to ask the agent with spaces for answers. The paper included an example to remind participants how to query the agent. The researcher told participants to ask the agent to inform the researcher when they were complete (by saying "Okay Pepper, text the researcher"); this was also provided on the handout. The researcher then left the room, leaving the participant alone with the robot, and the first phase (no-errors) began.

Once the participant indicated they were finished, the researcher returned and administered the trust questionnaire to measure baseline trust. The participant was given a second near-identical hand-out with different math questions. The participant was reminded of the task procedure, and how to contact the researcher, and the researcher left the room again, beginning the errors phase. Finally, upon completion the researcher returned and again administered the trust questionnaire.

To finish, the researcher administered the post-test questionnaire (including manipulation checks) and debriefed the participant on the study purpose and deceptions (including the Wizard of Oz implementation and purposeful errors).

The entire experiment took about 30 minutes, and participants were paid $10 for their time. This study was approved by our institutional ethics review board.

4.5 Procedure: Wizarding

We designed our wizarding protocol to simulate a realistic digital personal assistant similar to existing systems, except focused on the calculator application; in all cases, the wizard (and thus agent) would respond consistently, and without variation. The wizard did not respond if the participant did not clearly say the "OK Pepper" start phrase; if the wizard could not understand the participant, or if an utterance would not be expected to work with standard voice-command assistants (e.g., such as "can you, uhh.. actually, sorry, Pepper, how about…"), the agent would specifically say "Sorry, I couldn't quite catch that!". If the participant asked a question outside of the calculator protocol, the wizard would tell the participant, "Let's stay on track please."

To maintain a consistent and believable agent response speed, we pre-programmed the wizarding interface with answers to the 20 math questions, assigned to hot-key accelerators. Further, if participants made small mistakes (e.g., reading 20 as 30), the interface had shorthand to enable the wizard to quickly type unpredicted responses (with the interface displaying and a "thinking" spin graphic) to maintain the illusion of a calculator. If the wizard made an unplanned error, this created an inconsistency in our error manipulation and so the session was excluded from analysis.

When making a planned error the agent consistently answered the question incorrectly, even if the participant noticed and asked again. If a participant rephrased or changed the question, however, the agent reverted to answering correctly.

4.6 Results

We recruited 39 participants, but excluded 13 due to a failed manipulation check –did not appear to notice the agent error (we address this further in our discussion). Two participants were excluded as outliers who answered with all extreme values (max/ minimums) on the trust scale, corroborated as the only participants with responses more than three inter-quartile ranges from the mean. This resulted in 24 participants, aged 18-37 (average 24.1, 9 female, 15 male). Participants were alternately allocated to the two embodiments; 11 in the humanoid condition, and 13 in the tablet condition.

We conducted 2x2 Mixed Model ANOVAs (between: tablet, humanoid; within: baseline, errors) on the four trust subscales. We found statistically-significant main effects of agent error (with vs. without) on all trust subscales, with trust ratings lower after error (Table 1). We also found a main effect of agent embodiment (tablet vs. humanoid) on all trust subscales, with all mean trust ratings lower for the tablet than the humanoid (Table 2). We graph the 2x2 results in Fig. 3 by subscale.

Table 1. Main effects of agent error (baseline, with errors) on participant rating of trust subscales. Higher is more trust.

		$F_{1,22}$	p	baseline (/7)	with-error (/7)	$\eta 2$
capacity trust	reliable	56.1	<.001	6.48	4.31	.72
	capable	48.9	<.001	6.67	4.85	.69
moral trust	ethical	6.3	.02	6.16	5.59	.22
	sincere	11.9	.002	6.21	5.37	.35

Table 2. Main effects of agent form (tablet, humanoid) on participant rating of trust subscales. Higher is more trust.

		$F_{1,22}$	p	humanoid (/7)	tablet (/7)	$\eta 2$
capacity trust	reliable	7.2	.014	5.88	4.91	.25
	capable	4.8	.039	6.02	5.50	.18
moral trust	ethical	4.6	.043	6.54	5.20	.17
	sincere	5.0	.036	6.11	5.46	.18

We found a statistically-significant interaction effect between agent error and form on both of the ethical ($F_{1,22}$=4.0, *p*=.05) and sincere ($F_{1,22}$=5.7, *p*=.027) moral trust subscales (see Fig. 3 for visual interpretation). Post-hoc t-tests (Bonferroni correction) on ethical trust found no difference between the humanoid and tablet without-error (t_{21}=1.4, *p*=.18), but with error, found a difference on ethical trust between the humanoid (6.49/7) and tablet (4.69/7, t_{22}=2.6 *p*=.03). Investigating the effects of error by agent embodiment, we found no effect of agent error on trust in the humanoid case (t_{10}=.5), but found a statistically-significant effect of the error in the tablet condition (t_{12}<2.7, *p*=.036, pre-error 5.71/7, post-error 4.69/7). These results support the visual interpretation of the interaction effect (Fig. 3).

For investigating the interaction effect for sincere trust, we found no difference between the humanoid and tablet forms without-error (t_{22}=.3), but with-error, we found a difference between the humanoid (5.98/7) and tablet (4.75/7, t_{22}=2.5, *p*=.028). Investigating the effect of error by agent embodiment, we found no effect of agent error on sincere trust in the humanoid (t_{10}=.7), but found an effect of the error in the tablet condition (t_{12}=4.5, *p*=.002, pre-error 6.17/7, post-error 4.75/7). These results again support the visual interpretation of the interaction effect (Fig. 3).

5 Discussion

Overall, our results suggest that people will trust a humanoid agent more than its tablet-only variant for both capacity and moral trust. Further, our results support the hypothesis that people will have less trust in an agent when it makes errors; this serves as a manipulation check supporting our use of agent error to decrease trust.

While the interaction effects for moral trust complicates analysis, our post-hoc tests corroborate the visual analysis evident in Fig. 3: agent error does not impact moral trust in the humanoid, but moral trust falls for the tablet. In other words, the humanoid form mitigates the moral trust drop resulting from an obvious functional error. Further, while the post-error results may be driving the main effect of embodiment impact, overall people still have more moral trust in the humanoid than the tablet.

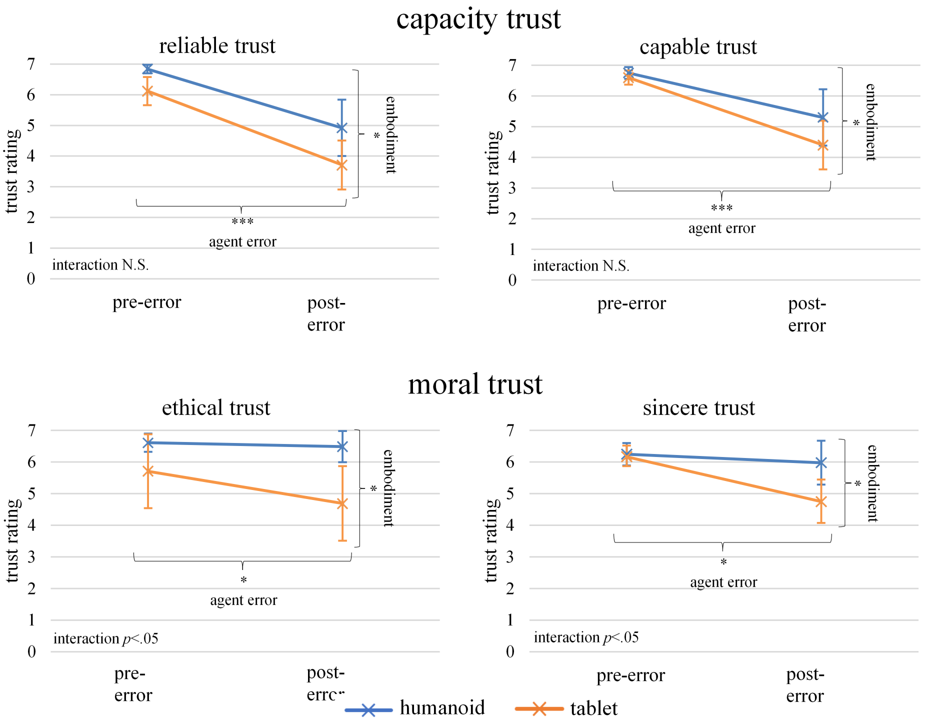

Fig. 3. Participant trust rating by subscale and embodiment, before and after errors. All main effects of embodiment and error were significant. The interaction effect for ethical and sincere trust resulted in a drop in trust for tablet, but not the humanoid. Error bars are 2SE.

The differences we observed between capacity trust (reliable, capable) and moral trust (ethical, sincere), are important to consider: across the board, trust dropped after agent error *except for moral trust of a humanoid robot*. Why exactly capacity trust would fall but not moral trust, or why this would happen for a humanoid but not a tablet, needs to be further investigated.

While the concept of moral trust being applied to a humanoid can be analyzed within the framework of social robotics (e.g., that people treat robots as social others [48]), it is not clear how moral trust would apply to a tablet agent; perhaps, as it used human-like speech, participants applied similar principles to understand the agent (see, e.g., [33]). If that is the case, then it is further unclear why the humanoid form only would not experience a moral trust drop after making errors.

One possibility relates to the fact that the physical design of the humanoid robot (see Fig. 1) has the tablet attached to the front of the robot (and not, e.g., embedded within it). Perhaps, then, participants saw the humanoid tablet as being separate from the robot itself, as an accessory, and thus do not attribute error blame to the robot *per se* (but to its attached tablet). However, this then contradicts the finding of capacity trust falling for the humanoid after the agent error; here again, the reason for capacity and moral trust having differing results is not clear.

Another possibility is that the humanoid may have elicited more pro-social behavior from participants. For example, while they note that the humanoid is less capable to make concrete calculations after the error, the social design may encourage participants to dismiss the error, for example, as an honest mistake, in an affiliative fashion, in a way that makes it not the robot's *fault* from a moral standpoint.

Along a similar vein, an agent's form (e.g., gender [8, 20]) can shape assumptions about task suitability. If the tablet was seen as a traditional computer, commonly known to be good at math, the participant may see a computer that cannot do math as not being genuine, authentic, or respectable (moral trust subscales). Alternatively, people may apply more human-typical notions to a humanoid; a person making math mistakes is common and not related to their moral reasoning ability. In both of these cases, results may relate to expectations and the Pratfall effect.

It is worth considering why the *tacked-on* humanoid features altered perceptions of trust at all: the embodiment change did not change explicit interaction apart from the generic hand gestures and gaze, which did not have functional purpose. The humanoid did not engage in any authentic or complex social interaction. However, while the robot's gestures and actions were functionally simple and not interactive, we should note that the robot emitting these social signals may trigger strong reactions in people. People can be expected to assume underlying processes behind the signals (e.g., a robot that looks at you can think), even if such capability does not exist [6]). Similarly, research has suggested that robot movement can be linked to perceptions of intent (even when there is none) and can cause a robot to appear more intelligent [14, 15].

Finally, we have concern over the fact that 13 of our 39 participants did not notice the robot errors, despite performing well on the math skills quiz, and the errors themselves being obvious (e.g., dividing a bill resulting in a larger number). Unfortunately, we do not have the qualitative or interview data to appropriately analyze why this may have happened. Perhaps given the rote nature of the task people were simply not paying attention but rather mechanically following procedure. Perhaps participants simply trusted the digital agents blindly as they know computers are good at math. Such reactions could easily carry over to a real restaurant situation, further emphasizing how even people competent in a task may put too much trust in a humanoid robot.

As a post-hoc measure, we re-conducted our analysis including all participants who were excluded due to failed manipulation check. The main effects and general results *do not change* from what is presented here. This further complicates the analysis, as even if participants claim they did not notice the errors, they apparently still reacted differently. However, we stand by our choice of excluding those participants from the main analysis, and the underlying cause of this phenomenon requires additional study.

Overall, the simple resulting message is that adding a humanoid form to an existing agent interaction, for example, to increase engagement or support intuitive interaction, will also impact how people trust the agent, and how forgiving they may be of agent errors. Such secondary, or perhaps collateral impacts, need to be better understood as we continue to see humanoid kiosk robots enter public spaces. Our work is the first to clearly establish a strong link between agent embodiment and trust resilience in the context of information agents being converted into humanoid robots.

6 Limitations

Novelty is a still a large consideration for experiments involving robots, and in this case, we expect that the excitement of having an opportunity to speak with a humanoid shaped our results. For example, despite being asked to not use their phones during the experiment, the wizard researcher observed many participants secretly taking photos and videos, and asking humorous questions to the robot for fun. We did not see similar behaviors with the tablet. It is possible the excitement of interacting with a humanoid was a driver of our results, and not other factors of the embodiment. However, we want to highlight that *novelty is not a confound in robot experiments*, but simply a reality of robots in society today; we similarly will expect people to experience novelty when encountering robots in the real world [43].

While we measured an abstract, self-report overall perception of trust, future work should explore situations where a user has something real at stake (e.g., money, material loss). However, we feel that our low-stakes experiment is reasonably representative of a real-world scenario where the stakes may not be immediately clear to a person (e.g., advice in an airport). Further, our lab-based study has limited ecological validity, for example, it did not provide participants with explicit incentive to pay attention. If such a study was performed in the wild with real stakes or implications (perhaps where people were spending their own money), we may find different results. Our task itself then, with many restaurant-based questions, is valuable as it implementable with current technology, and could be extended to an in-the-wild study without great technological investment.

One of the core ideas surrounding this work is that an agent's embodiment implies properties about an agent – even how to interact with it [46]. In our non-humanoid tablet, people may be expecting to interact with touch, but could not; this may have influenced participants to consider the tablet as somehow defective, making it easier to lose trust in it. While this may not have a strong effect in our case as people are increasingly familiar with voice-only agents like Siri, future task designs should consider interaction methods that feel natural for all embodiments being compared.

7 Conclusions

Humanoid robots are becoming increasingly common in public places where they interact with people to provide information or kiosk-like services. Commonly, these robots replace an existing interactive kiosk or agent, ostensibly as a robot can be more engaging, exciting, or easy to work with. However, in this work we highlight that simply wrapping a device in a humanoid embodiment may incorporate additional social impacts that both the designer and user may not have intended or be aware of. Our initial results in this space show how making an interactive voice-controlled agent humanoid can increase trust in that agent. Further, people are more forgiving of the humanoid, in that even after it makes serious functional errors, they do not report a reduction in moral trust of the robot; this is in contrast to a non-humanoid tablet agent, where people's trust reduced across the board following functional errors. This can have serious consequences for humanoids used in high impact or sensitive scenarios, such as banks, hospitals, or schools.

Overall, these results highlight the potential dangers of using humanoid designs and indicate the importance of considering the broad consequences – intentional or not – of making devices humanoid. Moving forward, we hope this work contributes to developing a healthy skepticism and balanced approach to considering the full implications of humanoid robots entering society.

References

1. Admoni, H., Scassellati, B.: Social Eye Gaze in Human-Robot Interaction: A Review. J. Human-Robot Interact. 6, 1, 25 (2017). https://doi.org/10.5898/JHRI.6.1.Admoni.
2. Bainbridge, W. a. et al.: The Benefits of Interactions with Physically Present Robots over Video-Displayed Agents. Int. J. Soc. Robot. 3, 1, 41–52 (2010).
3. Bartneck, C. et al.: The influence of robot anthropomorphism on the feelings of embarrassment when interacting with robots. Paladyn J. Behav. Robot. 1, 2, 109–115 (2010).
4. Bartneck, C. et al.: To kill a mockingbird robot. In: HRI '07. ACM, USA (2007).
5. Booth, S. et al.: Piggybacking Robots. In: Human-Robot Interaction. ACM (2017).
6. Collins, E. et al.: The Broader Context of Trust in HRI. In: Fisher, M. et al. (eds.) Dagstuhl Reports: Ethics and Trust : Principles, Verification and Validation. pp. 82–85.
7. Desai, M. et al.: Effects of changing reliability on trust of robot systems. In: Human-Robot Interaction. pp. 73–80 (2012). https://doi.org/10.1145/2157689.2157702.
8. Eyssel, F., Hegel, F.: (S)he's Got the Look. J. Appl. Soc. Psychol. 42, 9, 2213–2230 (2012).
9. Flook, R. et al.: On the impact of different types of errors on trust in human-robot interaction. Interact. Stud. 20, 3, 455–486 (2019). https://doi.org/10.1075/is.18067.flo.
10. Geiskkovitch, D. et al.: Autonomy, Embodiment, and Obedience to Robots. In: Human-Robot Interaction Extended Abstracts. pp. 235–236. ACM (2015).
11. Geiskkovitch, D.Y. et al.: Please Continue, We Need More Data: An Exploration of Obedience to Robots. J. Human-Robot Interact. 5, 1, 82–99 (2016).
12. Gordon, G. et al.: Can Children Catch Curiosity from a Social Robot? In HRI '15.
13. Hancock, P.A. et al.: A Meta-Analysis of Factors Affecting Trust in Human-Robot Interaction. Hum. Factors J. Hum. Factors Ergon. Soc. 53, 5, (2011).
14. Harris, J., Sharlin, E.: (e)motion. In: RO-MAN. IEEE (2011).
15. Heider, F., Simmel, M.: An experimental study of apparent behavior. Am. J. Psychol. 57, 2
16. Herse, S. et al.: Do You Trust Me, Blindly? Factors Influencing Trust Towards a Robot Recommender System. In: RO-MAN. IEEE (2018).
17. Jung, M.F. et al.: Using Robots to Moderate Team Conflict. In: HRI '15. ACM (2015).
18. Kahn, P.H. et al.: "Robovie, you'll have to go into the closet now": children's social and moral relationships with a humanoid robot. Dev. Psychol. 48, 2, 303–14 (2012).
19. Kahn, P.H. et al.: Will People Keep the Secret of a Humanoid Robot? In: Human-Robot Interaction -. pp. 173–180 (2015). https://doi.org/10.1145/2696454.2696486.
20. Kuchenbrandt, D. et al.: Keep an Eye on the Task! How Gender Typicality of Tasks Influence Human–Robot Interactions. In: Social Robotics. pp. 448–457 Springer (2012).
21. Kwon, M. et al.: Human expectations of social robots. ACM/IEEE Int. Conf. Human-Robot Interact. 2016-April, 463–464 (2016). https://doi.org/10.1109/HRI.2016.7451807.
22. Li, D. et al.: A Cross-cultural Study: Effect of Robot Appearance and Task. Int. J. Soc. Robot. 2, 2, 175–186 (2010). https://doi.org/10.1007/s12369-010-0056-9.
23. Madhavan, P. et al.: Automation failures on tasks easily performed by operators undermine trust in automated aids. In: Human Factors. pp. 241–256 (2006).
24. Mann, J.A. et al.: People respond better to robots than computer tablets delivering healthcare instructions. Comput. Human Behav. 43, 112–117 (2015).
25. Mathur, M.B., Reichling, D.B.: An uncanny game of trust: In: Proceedings of the 4th ACM/IEEE International Conference on Human-Robot Interaction, HRI'09. (2008).
26. Moore, D. et al.: Making Noise Intentional: A Study of Servo Sound Perception. (2017).

https://doi.org/10.1145/2909824.3020238.
27. Pak, R. et al.: Decision support aids with anthropomorphic characteristics influence trust and performance in younger and older adults. Ergonomics. 55, 9, (2012).
28. Pan, Y., Steed, A.: A Comparison of Avatar-, Video-, and Robot-Mediated Interaction on Users' Trust in Expertise. Front. Robot. AI. 3, 12, (2016).
29. Park, E. et al.: The effect of robot's behavior vs. appearance on communication with humans. Human-Robot Interact. 219–220 (2011). https://doi.org/10.1145/1957656.1957740.
30. Powers, A. et al.: Comparing a computer agent with a humanoid robot. In: Human-Robot Interaction. pp. 145–152 (2007). https://doi.org/10.1145/1228716.1228736.
31. Ragni, M. et al.: Errare humanum est: Erroneous robots in human-robot interaction. In: Robot and Human Interactive Communication, RO-MAN. pp. 501–506 IEEE (2016).
32. Rea, D.J., Young, J.E.: It's All in Your Head. In: Human-Robot Interaction. pp. 32–40 ACM Press, New York, New York, USA, New York, USA (2018).
33. Reeves, B., Nass, C.: The Media Equation: How People Treat computers, television, and new media like real people and places. CSLI Books (1996).
34. Robinette, P. et al.: Overtrust of Robots in Emergency Evacuation Scenarios. Human-Robot Interact. 101–108 (2016). https://doi.org/10.1109/HRI.2016.7451740.
35. Roubroeks, M. et al.: When Artificial Social Agents Try to Persuade People: The Role of Social Agency on the Occurrence of Psychological Reactance. Int. J. Soc. Robot. 3, 2 (2011).
36. Salem, M. et al.: To Err is Human(-like): Effects of Robot Gesture on Perceived Anthropomorphism and Likability. Int. J. Soc. Robot. 5, 3, 313–323 (2013).
37. Salem, M. et al.: Would You Trust a (Faulty) Robot?: Effects of Error, Task Type and Personality on Human-Robot Cooperation and Trust. Proc. Human-Robot Interact. (2015).
38. Sanoubari, E. et al.: Good Robot Design or Machiavellian? In-the-Wild Robot Leveraging Minimal Knowledge of Passersby's Culture. In: Human-Robot Interaction (HRI). (2019).
39. Seo, S.H. et al.: Poor Thing ! Would You Feel Sorry for a Simulated Robot ? A comparison of empathy toward a physical and a simulated robot. In: Human-Robot Interaction. (2015).
40. Sharma, M. et al.: Communicating affect via flight path exploring use of the laban effort system for designing affective locomotion paths. Human-Robot Interact. 293–300 (2013).
41. Short, E. et al.: No fair!! An interaction with a cheating robot. In: Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction 2010, HRI '10.
42. Singh, A., Young, J.E.: A dog tail for utility robots: Exploring affective properties of tail movement. Lect. Notes Comput. Sci. 8118 LNCS, PART 2, 403–419 (2013).
43. Smedegaard, C.V.: Reframing the role of Novelty within Social HRI. In: Human-Robot Interaction, 2019. alt.HRI 2019. (2019).
44. Ullman, D., Malle, B.F.: MDMT : Multi-Dimensional Measure of Trust. 618–619 (2019).
45. de Visser, E.J. et al.: Almost human: Anthropomorphism increases trust resilience in cognitive agents. J. Exp. Psychol. Appl. 22, 3, 331–349 (2016).
46. Wainer, J. et al.: The role of physical embodiment in human-robot interaction. Proc. - IEEE Int. Work. Robot Hum. Interact. Commun. 117–122 (2006).
47. Young, J.E. et al.: Evaluating Human-Robot Interaction. Int. J. Soc. Robot. 3, 1 (2010).
48. Young, J.E. et al.: Toward acceptable domestic robots. Int. J. Soc. Robot. 1, 1, (2009).