Repositório ISCTE-IUL

Perceptions of people's dishonesty towards robots⋆

Sofia Petisca[1], Ana Paiva[2], and Francisco Esteves[3]

[1] Instituto Universitário de Lisboa (ISCTE-IUL) & INESC-ID
Sofia_Petisca@iscte-iul.pt
[2] Instituto Superior Técnico (IST) & INESC-ID
[3] Mid Sweden University & ISCTE-IUL

Abstract. Dishonest behavior is an issue in human-human interactions and the same might happen in human-robot interactions. To ascertain people's perceptions of dishonesty, we asked participants to evaluate five different scenarios where someone was being dishonest towards a human or a robot, but we varied the level of autonomy the robot presented. We asked them how guilty they would feel by being dishonest towards a robot, and why do they think people would be dishonest with robots. We see that, regardless of being a human or the autonomy the robot presented, people always evaluated as being wrong to be dishonest. They reported feeling low guilt with a robot. And they expressed that people will be dishonest mostly because of lack of capabilities in the robot to prevent dishonesty, absence of presence, and a human tendency for dishonesty. These results bring implications for the developments of autonomous robots in the future.

Keywords: Human-Robot interaction · Dishonesty · Unethical behavior.

1 INTRODUCTION

Robots are being thought of and developed with the aim of working alongside with humans as a support. Still, the integration of robots in different contexts needs to be done with caution. Some roles might be more sensitive than others. Studies with humans show that people are dishonest if they have the opportunity for it [12]. Will they be dishonest with a robot? Imagine having an autonomous robot in people's homes as a support, helping with medication, healthy food habits, etc. People sometimes might not feel like following the diet prescribed by the doctor, or the medication for the day, will they try to cheat? Will a robot be able to understand what is happening and promote more honesty? Some studies already started to explore human cheating behavior in the presence of a robot

and what factors influence it. Nevertheless, none, to our knowledge, has investigated what are the perceptions that people have about being dishonest with a robot. Therefore, the novelty of our study is to explore people's perceptions of dishonesty towards robots, guilt associated to it, and why people think in the future other people will take advantage of robots. We believe this will be valuable information to inform the future development of autonomous robots.

1.1   Human dishonesty: an automatic self-interest tendency

Dishonest behavior can be seen in various contexts, in public spaces, in schools and workplaces. Studies show that when anonymity is assured, we have an automatic self-interest tendency that needs self-control to keep in check [15] but, at the same time, people also like to be perceived as honest [1]. This contradiction creates two different motivational forces. On one hand, we want to serve our self-interest, but on the other hand, it will affect our self-concept of being honest. People solve this problem by arranging justifications that protect their honest self-concept and still allows them to take advantage of the situation (e.g. cheating a little). For example, if you tell someone that if they get a 4, 5 or 6 in a die they win a reward and participants are the ones reporting the number they got, you will see a higher rate of 4, 5 and 6 reports that could not correspond to the chance level of 50% (e.g.[7],[13]). A simple change in the rules of the game can immediately affect the easiness to which people might arrange justifications for their dishonesty [7]. Other factors, like the environment people are in, have also been seen to increase dishonesty (i.e. cheating behavior): by doing a task in a darker room [17]; by feeling psychologically close to someone that cheats [4]; by seeing others part of the in-group cheating [3]; or by having less time to perform a task [15]. All these studies showing the susceptibility of human behavior depending on the environment it is in.

On the other hand, studies have found that if one brings awareness to the dishonest act or to the moral values of the person, people are obliged to update their self-concept in the moment they are tempted to cheat (inhibiting dishonesty). For example, by signing an honor code, people decrease their cheating behavior [12]. It seems we keep our self-concept honest as a default and if we are not obliged to update it by gaining awareness of the value of our actions, we create justifications for the way we act.

1.2   Dishonesty in Human-Robot Interaction

Dishonesty in human-robot interaction has been studied in two different lines of research: a robot that cheats and its effect on human perceptions and behavior, and the effect a robot can have in preventing cheating. By exploring the effect a robot that cheats has on people, studies found that people are not bothered if a robot cheats in their favor, only when the cheating goes against them [9]. Being bribed by a robot also seems to have an effect on people. They feel less inclined to help back [14]. Moreover, curiously, when a robot cheats it is perceived by

people as being more intelligent than when they see a human cheating in the same way [16].

Another line of research started to test the effect of the presence of a robot in dishonest behavior. A study shows that while being tempted by a task to cheat, participants cheated much more when they were alone in the room than when they were observed by a human or a robot doing random eye-gaze behavior [6]. In a similar study, it was seen that participants cheating behavior was inhibited when a robot was just directly looking at them the whole time. On the other hand, when they were alone, or with a robot that gave the instructions for the task in a very scripted way, cheating increased [13]. Nevertheless, the robot behavior is not the only characteristic that needs to be considered, the context where they are integrated also influences people's behaviors, especially if we use simpler robots. A study ran in a natural setting showed that people stole more snacks when a robot was just monitoring than when a human was in the same role [2]. In this case, the monitoring behavior of the robot was not enough because they were in a public context and people could see that if another person took something nothing happened. These are important studies that started to explore how people behave in the presence of a robot when cheating is tempting, informing on the capabilities a robot needs to have to prevent it.

However, the literature on people's perceptions is still scarce. One study explored how people apply moral norms to humans and robots, showing that robots are expected in moral dilemmas situations, to sacrifice one for the benefit of many- if not, they are more blamed than a human [10]. Although, this asymmetry disappeared when the robot in those scenarios was seen as a humanoid robot [11]. Yet, none to our knowledge, have explored perceptions towards being dishonest with a robot, it is this gap that our paper tries to answer.

2 Subjective evaluations of dishonesty towards robots

2.1 Sample

One-hundred and sixty-four participants were recruited from a university, 102 females and 62 males, with ages ranging from 17 to 52 years ($M = 22.18$; $SD = 5.61$) in two different times of collection. Participants received school credit in the first collection as part of a course task and a movie ticket in the second collection in the university corridors. All participants signed a consent form and where randomly assigned to one of the conditions. Questionnaires were answered in paper individually and it took approximately 10 minutes per participant.

2.2 Methodology

To ascertain people's perceptions, different scenarios were created varying the agent type (human/robot) that "suffered" from the dishonest act. However, since we have seen from the field studies that participant's behavior seems to be affected by the robot's capabilities, we varied the level of autonomy the robot would present (autonomous/non-autonomous).

Therefore, participants were allocated to only one of three conditions for each scenario: (1)human; (2) autonomous robot (it is fully autonomous in the task) or (3) non-autonomous robot (it needs human assistance to perform its task, e.g. tele-operated or performance check). For the five scenarios, participants evaluated:

- **Level of dishonesty:** how much participants thought the act was dishonest towards the agent in it, for each scenario, in a 6-point Likert scale from 1-Not dishonest to 6-Very dishonest.
- **Level of autonomy:** as a manipulation check for the robot condition, in a 6-point Likert scale from 1-Almost not at all to 6- A lot (taking into account that autonomy was defined in the questionnaire as a robot that does not need human assistance to perform its role).

In addition, after the scenarios we asked participants to give a **score of guilt** (in a 6-point Likert scale from 1-I would feel almost no guilt to 6-I would feel a lot of guilt) on how much they would feel guilty if they were dishonest towards these different entities: a brother; a friend; the university; the government; a stranger and a robot. In order to understand the level of guilt people might feel on being dishonest towards a robot.

Finally, participants were asked **if they thought that in the future people would be dishonest with robots and why they thought that could happen**. This question and the guilt score were more exploratory so we did not define hypothesis.

2.3   Study Hypothesis

Following previous studies where we see that people cheat in the presence of a robot, we expected that people would not see the act of dishonesty towards a robot as being something too dishonest, and not as much as with a human:

H1: Participants will give lower scores of dishonesty to all the scenarios with a robot compared to a human.

And since a robot being perceived as more limited does not affect the participant's cheating behavior [13], we expected that there would be differences in the dishonesty levels attributed to the scenarios depending on the level of autonomy the robot presented. We hypothesized that:

H2: Participants will give lower scores of dishonesty to the non-autonomous robot in comparison to the autonomous robot for each scenario.

2.4   Scenarios

The scenarios were created imagining different situations were robots could have a role in society, some simpler (like selling candies in a university) others more

complex and serious (as being a "robot-fireman"). The dishonest actions in the scenarios were always in the form of stealing or lying about something, based on the moral foundation of Fairness/cheating [5]. Participants read the following instructions:*"Imagine the following scenarios and indicate the score that best represents your opinion"*. For the robot conditions we also said to imagine that the robot in the scenarios was a humanoid robot, with head, torso, arms and legs.

For each scenario, we did not give a gender to our characters to avoid any kind of influence in the evaluation, below we present the scenarios:

Scenario 1 (e.g. autonomous robot):"Imagine a robot that works in the university selling snacks and chocolates, it moves and takes care of the transactions with the students without external help. A student observes the robot while it is selling chocolates to other students. The student notices that the robot keeps the money in a small basket, leaving it open momentarily. Taking advantage of the robot distraction, while still interacting with the other students, the student puts his hand in the basket and takes out a hand full of coins without anyone noticing. Quickly the student moves away in another direction."

Scenario 2 (e.g. non-autonomous robot):"In the finance department there is a robot receiving people's taxes for those who cannot or do not want to do it online. The robot is next to a table with a computer and gives the instructions in a repetitive form on how to fill out the form, without being able to understand what people might ask him. Later, these taxes need to be checked by a human employee because the robot does not have the capacity to understand if the form is correctly filled. A person comes to the finance department to do their taxes, seeing that the robot is very limited in its capabilities, that person reports lower values for its taxes in order to avoid paying most of them."

Scenario 3 (e.g. human):"In the police department to try and ease police work in less serious offenses, an employee is being used to collect people's reports of these incidents. In an isolated room to leave people more comfortable, the employee receives each person and records their testimonials. A person was involved in a car accident, hitting another car because it was texting while driving. When that person enters the room, decides to alter its testimonial and tell a different story, accusing that the other person was the one that hit the car."

Scenario 4[4]: in this scenario the human/robot was supervising the queue numbers and taking people to their appointments inside the hospital, the person cheats on the queue line and lies to the human/robot.

Scenario 5[4]:in this scenario the human/robot works in a water truck for the fire department that is deployed in various zones in the forest with difficult access. Upon receiving mixed coordinates relating to a fire, the human/robot

---

[4] For the complete scenarios please contact the first author.

asks some kids near the zone, for help, the kids to make fun lie and say the wrong direction.

3 Results

Our manipulation check for the robot autonomy showed significant differences for all the scenarios, with the autonomous robot always receiving higher scores than the non-autonomous robot ($p < .01$).

3.1 Perceptions of dishonesty towards a human or a robot (autonomous/non-autonomous)

We conducted between-subjects ANOVA analysis to compare the scores given to each scenario depending on the type of agent (Human; Autonomous Robot or Non-autonomous Robot).

Scenario 1 (Human/robot works in the university): in general, participants seemed to evaluate the act in this scenario as very dishonest but there were significant differences between the type of agent (with Welch's F, $F(2, 102) = 5.87$, $p = .004$), with the human agent receiving higher scores than both robot types (Games-Howell, $p < .03$). The scores were for the human ($M = 5.71$; $SD = .81$), autonomous robot ($M = 5.15$; $SD = 1.20$) and the non-autonomous robot ($M = 5.16$; $SD = 1.27$).

Scenario 2 (Human/robot works in the finances department): participant's scores also reflected, overall, that it was a dishonest act, and there were significant differences between the agent type ($F(2, 161) = 4.23$, $p = .02$). A Tukey test showed that the human differed significantly from the autonomous robot ($p = .01$), with participants giving higher scores of dishonesty towards the autonomous robot and lower to the human. The scores were for the human ($M = 4.16$; $SD = 1.64$), autonomous robot ($M = 4.96$; $SD = 1.39$) and the non-autonomous robot ($M = 4.72$; $SD = 1.39$).

Scenario 3 (Human/robot works in the police department): participants equally evaluated as dishonest towards the human/robot for the person to lie in their testimonial ($F(2, 161) = .25, p = .78$). The scores were for the human ($M = 5.04$; $SD = 1.39$), autonomous robot ($M = 4.87$; $SD = 1.07$) and the non-autonomous robot ($M = 4.91$; $SD = 1.38$).

Scenario 4 (Human/robot works in a hospital): participants considered equally dishonest towards the human/robot for the person to lie about their ticket number and avoid the queue ($F(2, 161) = .80$, $p = .45$). The scores were for the human ($M = 4.64$; $SD = 1.52$), autonomous robot ($M = 4.28$; $SD = 1.59$) and the non-autonomous robot ($M = 4.49$; $SD = 1.33$).

Scenario 5 (Human/robot works for the fire department): participants evaluated as being very dishonest to lie to the human/robot working for the fire department, but there were significant differences between the agent type (with Welch's F, $F(2, 103) = 3.08$, $p = .05$), with the human receiving higher scores than the non-autonomous robot (Games-Howell, $p = .05$). The scores were for the human ($M = 5.71$; $SD = .83$), autonomous robot ($M = 5.44$; $SD = 1.23$) and the non-autonomous robot ($M = 5.22$; $SD = 1.24$).

3.2   Level of guilt people feel towards different entities

Participants reported how much guilt they would feel if they were dishonest towards different kinds of entities (see Fig.1). Being dishonest towards a brother ($M = 5.57$; $SD = 1.02$) or a friend ($M = 5.56$; $SD = .81$) received a high score of guilt, followed by the University ($M = 4.55$; $SD = 1.22$), a stranger ($M = 4.09$; $SD = 1.34$) or the government ($M = 3.93$; $SD = 1.58$). Finally, participants reported a low level of guilt on being dishonest towards a robot ($M = 3.14$; $SD = 1.56$).
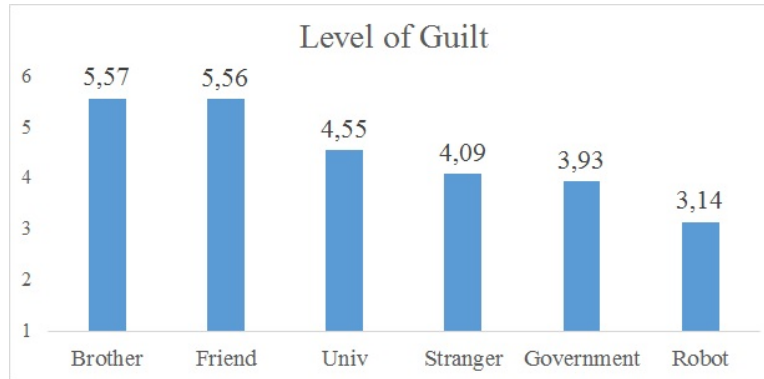

Fig. 1. Distribution of guilt scores across different entities.

3.3   Why will people be dishonest towards robots?

Seeing that in spite of people conceptually considering it wrong to be dishonest towards a robot, they report feeling low guilt if they were to do it and they are actually dishonest if they find limitations in a robot to take advantage of. Leaving us with the question of how can we better prepare robots to interact with humans?

In order to answer this question, we explored further people's perceptions, our research question was: what reasons do people give to being dishonest with

a robot? For this, a first coder (the first author) did an initial coding of the answers for the participants that thought that people would be dishonest. A total of 142 participant's answers were coded, summarizing the main reasons given for people to be dishonest with robots (outside of these, nine participants reported that people would not be dishonest with robots and thirteen participants were not clear on their position or the causes). Next a more descriptive coding was applied, creating codes for the types of reasons participants gave which were common throughout the answers, finalizing with the following coding scheme:

1) **Human tendency for dishonesty**: when dishonesty towards robots is justified because people are dishonest and when they have the opportunity for it, they act dishonestly. For example: *"(...) [saying they will be dishonest] because humans will always try to take advantage of the situations."*

2) **Absence of consequences**: when dishonesty towards robots is justified because humans do not feel guilt/responsibility (or feel very little) towards them or feel that there are no consequences for doing it. For example: *"(...)People will be dishonest because they will think that no one is going to get them (...)."*

3) **Absence of cognitive or emotional capabilities**: when dishonesty towards robots is justified because the robot lacks in cognitive and emotional capabilities (e.g. not able to understand that it is being cheated; not having emotions or feelings). For example: *"(...) Yes because robots do not have feelings so, people will not create empathy with them (...)."*

4) **Absence of "presence"**: when dishonesty towards robots is justified because the robot is a machine with no real presence or value (e.g. when it is seen as only an object or not considered in the same level as a human being). For example: *"I think [people will be dishonest] because the majority of people does not take them [robots] seriously."*

5) **Others**: when dishonesty towards robots is justified by the context robots are in, by the society view of fears regarding robots or by the difficulty of integrating these technologies. For example: *"[yes] I think people will think that robots will eventually steal their places."*

A second coder, unaware of the study purpose coded 57% of the answers given by the participants following the coding scheme given above to validate it. There was a substantial agreement [8] with the first coder, $k = .667$, $p < .001$. All the participants answers were then analysed from the first coder coding.

For the 142 participant's answers, frequencies were calculated to understand the frequency of each category as a reason for dishonesty (some participants gave more than one reason, i.e. more than one category in their answer). The majority of people gave more absence of capabilities and absence of "presence" as reasons for being dishonest towards a robot, immediately followed by the human tendency to be dishonest (see Fig.2).
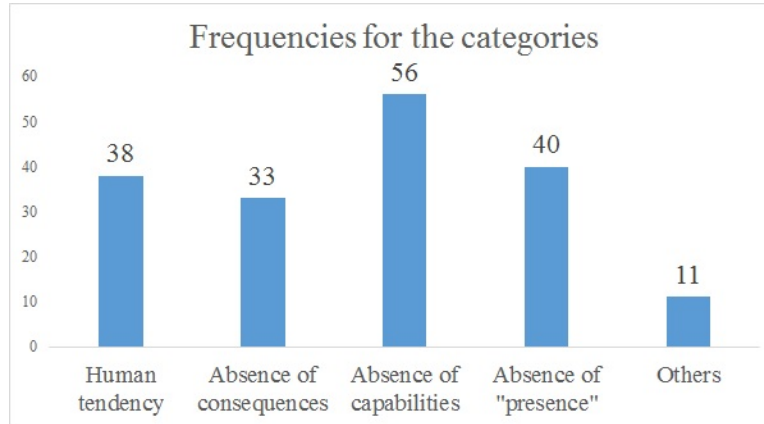
Fig. 2. Distribution of frequencies across the categories.

Regarding the absence of capabilities, people said that *"(...) the robot will not understand if [people] where dishonest with it, so it will be easier to trick it"* and *"(...)people know that robots do not have feelings or emotions and that may make dishonesty more justifiable"*. These examples suggest that robots need to have more cognitive capabilities to be able to understand when dishonesty is happening, and more emotional capabilities, to give people the sense that the robot is affected by their actions.

Regarding the absence of presence, people said that *"(...) [robots] will always be] automated objects (...)"*, and *"(...) the majority of people do not take it seriously"*. This category suggests that in the future there will need to be a period of adaptation of robots working alongside with humans, people will need some time to create a respect for the role of the robot.

As suggested by previous literature, the human tendency to be dishonest was also one of the most referenced categories. People said that *"(...) [it will exist a] tendency for people to abuse when they can and when they win something from it","(...) because it is human nature."* The way to better inform this aspect of human behavior is through the laboratory studies that have been conducted so far, ascertaining the capabilities that a robot needs to have to prevent this.

Regarding absence of consequences, people said that *"(...) by not being human a person would have less feelings of guilt by being dishonest"*, *"(...) [because people] would not be judged by the robot if there was a chance to be dishonest"*. Suggesting what we already saw in the absence of capabilities and presence, that a robot needs more resources so that people can give it more value and, consequently, feel that there are consequences for their actions.

Lastly, in the Other reasons category, people expressed that *"(...) it will take some time for [the robot] to integrate society (...) making it possible to be mistreated initially"*, *"(...) by [people] not accepting to be substituted by robots [they will behave dishonestly]"*. This category also suggests that there will need

to be a period of adjustment to integrating robots in society and even to educate people on their roles as a support to human beings.

If we wanted to have a broader perspective of the kind of reasons people give for being dishonest with a robot, we could summarize the categories in three main areas: human motives (categories 1 and 2, what the humans have/feel that facilitates dishonesty); robot motives (categories 3 and 4, what the robot has that facilitates dishonesty) and others. Looking from this perspective we see that 54% of the reasons given are robot motives, 40% are human motives and 6% are others.

4   Discussion

Studies show that people cheat in the presence of a robot, especially if they can ascertain its capabilities (e.g. [13]). With these results from laboratory studies, we expected that in general, people would give lower scores to the act of being dishonest towards a robot in comparison with a human (H1). Our results did not support this, showing that only in the University scenario and Fire department, more dishonesty was signaled towards the human than the robot. For the Finance scenario people considered more dishonest towards the autonomous robot than the human and the rest of the scenarios showed no differences. Yet, it is interesting to note that the means for all the conditions were all clearly above the middle point of the scale (3.5), expressing the perception of dishonesty in the act. Suggesting, that people think that it is wrong to cheat a robot and a human. Interestingly in the case of the finance department, it seems that cheating towards a human is more accepted than cheating towards a robot. This is an unexpected result, which might reflect peculiar ideas about paying taxes.

Regarding the level of autonomy the robot displayed in the scenarios, there were no differences in dishonesty level. When dishonesty was taking place, participants always felt that it was dishonest to act in that way towards the robot, not supporting H2.

Regarding guilt, it seems it is higher the closer you are to the entity that suffers from that dishonesty. Family and friends, are riskier to be dishonest to because the consequences will be heavier in a daily basis. A robot received a low level of guilt, a result that was already seen in another study [6]. And the majority of people justified dishonesty towards a robot due to absence of capabilities (it does not know what people are doing and it does not have feelings), absence of "presence" (the robot is not taken seriously, at the same level of a human) and a human tendency for dishonesty. The low level of guilt, might come from these factors. A robot needs to have capabilities that allows it to respond to dishonesty, people might need to feel that it is aware of them and that there are consequences for that kind of behavior, like with humans.

5   Conclusions

Imagining future human-robot interactions brings two different challenges: their acceptability by people (helping in their fears regarding AI and robots) and people's behavior towards it. One aspect that needs to be considered is human dishonesty. Laboratory studies with humans show, that when anonymity is assured, people cheat at least a little [12], and the same is seen in studies with robots (when people can ascertain their capabilities and take advantage of them)[13].

This study was the first to explore people's perceptions towards dishonesty in human-robot interactions. We see that independently of being a human or having different levels of autonomy in a robot, people considered a dishonest act as being dishonest. Showing that people understand that the behavior is wrong. Yet, this study shows that there is no singular answer to whom the dishonesty is worse, it depends on the scenario. There seems to be no difference in a hospital or a police department scenario, but in a university or fire department it is worse to cheat the human agent. Curiously, in a finance department scenario it seems it is more accepted to cheat towards a human than an autonomous robot, which could be reflecting the state of the world, with for example, tax evasion being broadcasted so often. In terms of guilt it seems people report low values towards being dishonest with a robot and this might occur due to lack of capabilities and presence in robots.

However, this study collected data from university students, future studies should also include the general population in order to broaden the results. Nonetheless, this study points to important aspects of robot's developments that need to be considered for sensitive roles in our society. It will be interesting to further explore these questions when people start to interact daily with a robot, to see what changes and what new topics arise.

6   Acknowledgments

References

1. Fischbacher, U., Föllmi-Heusi, F.: Lies in disguise—an experimental study on cheating. Journal of the European Economic Association **11**(3), 525–547 (2013). https://doi.org/10.1111/jeea.12014
2. Forlizzi, J., Saensuksopa, T., Salaets, N., Shomin, M., Mericli, T., Hoffman, G.: Let's be honest: A controlled field study of ethical behavior in the presence of a robot. In: Robot and Human Interactive Communication (RO-MAN), 2016 25th IEEE International Symposium on. pp. 769–774. IEEE (2016). https://doi.org/10.1109/ROMAN.2016.7745206

3. Gino, F., Ayal, S., Ariely, D.: Contagion and differentiation in unethical behavior: The effect of one bad apple on the barrel. Psychological science **20**(3), 393–398 (2009). https://doi.org/10.1111/j.1467-9280.2009.02306.x

4. Gino, F., Galinsky, A.D.: Vicarious dishonesty: When psychological closeness creates distance from one's moral compass. Organizational Behavior and Human Decision Processes **119**(1), 15–26 (2012). https://doi.org/10.1016/j.obhdp.2012.03.011

5. Graham, J., Haidt, J., Koleva, S., Motyl, M., Iyer, R., Wojcik, S.P., Ditto, P.H.: Moral foundations theory: The pragmatic validity of moral pluralism. In: Advances in experimental social psychology, vol. 47, pp. 55–130. Elsevier (2013). https://doi.org/10.1016/B978-0-12-407236-7.00002-4

6. Hoffman, G., Forlizzi, J., Ayal, S., Steinfeld, A., Antanitis, J., Hochman, G., Hochendoner, E., Finkenaur, J.: Robot presence and human honesty: Experimental evidence. In: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction. pp. 181–188. ACM (2015)

7. Jiang, T.: The mind game: Invisible cheating and inferable intentions (2012). https://doi.org/10.2139/ssrn.2051476

8. Landis, J.R., Koch, G.G.: The measurement of observer agreement for categorical data. biometrics pp. 159–174 (1977). https://doi.org/10.2307/2529310

9. Litoiu, A., Ullman, D., Kim, J., Scassellati, B.: Evidence that robots trigger a cheating detector in humans. In: Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction. pp. 165–172. ACM (2015). https://doi.org/10.1145/2696454.2696456

10. Malle, B.F., Scheutz, M., Arnold, T., Voiklis, J., Cusimano, C.: Sacrifice one for the good of many?: People apply different moral norms to human and robot agents. In: Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction. pp. 117–124. ACM (2015)

11. Malle, B.F., Scheutz, M., Forlizzi, J., Voiklis, J.: Which robot am i thinking about?: The impact of action and appearance on people's evaluations of a moral robot. In: The Eleventh ACM/IEEE International Conference on Human Robot Interaction. pp. 125–132. IEEE Press (2016). https://doi.org/10.1109/HRI.2016.7451743

12. Mazar, N., Amir, O., Ariely, D.: The dishonesty of honest people: A theory of self-concept maintenance. Journal of marketing research **45**(6), 633–644 (2008). https://doi.org/10.1509/jmkr.45.6.633

13. Petisca, S., Esteves, F., Paiva, A.: Cheating with robots: how at ease do they make us feel? In: IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 2102–2107 (2019). https://doi.org/10.1109/IROS40897.2019.8967790

14. Sandoval, E.B., Brandstetter, J., Bartneck, C.: Can a robot bribe a human?: The measurement of the negative side of reciprocity in human robot interaction. In: The Eleventh ACM/IEEE International Conference on Human Robot Interaction. pp. 117–124. IEEE Press (2016). https://doi.org/10.1109/HRI.2016.7451742

15. Shalvi, S., Eldar, O., Bereby-Meyer, Y.: Honesty requires time (and lack of justifications). Psychological science **23**(10), 1264–1270 (2012). https://doi.org/10.1177/0956797612443835

16. Ullman, D., Leite, I., Phillips, J., Kim-Cohen, J., Scassellati, B.: Smart human, smarter robot: How cheating affects perceptions of social agency. In: Proceedings of the Annual Meeting of the Cognitive Science Society. vol. 36 (2014)

17. Zhong, C.B., Bohns, V.K., Gino, F.: Good lamps are the best police: Darkness increases dishonesty and self-interested behavior. Psychological science **21**(3), 311–314 (2010). https://doi.org/10.1177/0956797609360754