Honesty in the Digital Era:
Preliminary results on the Social Interactions with Robots, Avatars, and Humans

Fabiola Diana[1,2], Ruud Hortensius[3], Mariska Kret[1,2]

[1]Comparative Psychology and Affective Neuroscience Lab, Cognitive Psychology Unit, Leiden University, Netherlands
[2]Leiden Institute for Brain and Cognition (LIBC), Leiden University, Netherlands
[3] Department of Psychology, Utrecht University, Utrecht, The Netherlands

ABSTRACT

People lie daily. To lie or not is a cognitive decision that depends on a variety of factors, such as the potential personal gain, the repercussions for other people, or the distinct characteristics of the person we are lying to. One subtle social cue that can influence (dis)honesty is the interaction's partner pupil size. Van Breen and colleagues (2018) found that participants were more honest while watching a video of a partner with dilating pupils, compared with constructing ones. The past decades witnessed a huge growth in the use of artifical agents in our daily life. Therefore, our study aims to assess whether this effect of pupil size on dishonesty is modulated also by interacting with non-human agents. Participants (n=15) played multiple rounds of a lie-inducing coin toss game with three agents with either large or small pupils: an avatar on a screen, a small robot, and a human with customized contact lenses. Our preliminary results showed that neither pupil size nor the type of agent affected the degree of honesty. Regardless of the limitations, comparable results among the agents have promising implications from the perspective of an integrated society with artificial agents.

Keywords: human-robot interaction; pupil size; affective cues; moral cognition; social interaction;

1. INTRODUCTION

Human social interaction is an elegant expression of the interplay between cognition and emotions. Moral cognition, a crucial aspect of human social functioning, is a good example of this complex interplay: roughly, it allows us to evaluate basic harmful actions, and act n them (Buon et al., 2016). Deciding to act prosocially or not, in the case of honesty, is always a dilemma between what is good for oneself and what is beneficial for the other individual. This decision is often the result of an affect-based unconscious evaluation (Kret & De Dreu, 2013). In humans, this evaluation is based on a variety of emotional sources, such as posture (Grant & Mackintosh, 1963; Kret et al., 2020), smell (de Groot et al., 2012), and vocalization (Ross et al., 2010). Especially, facial signals play a critical role. Across culture, facial expressions of emotion play an important role to inform us on other people's intentions and goals (Ekman & Rosenberg, 2012). Different emotional expressions have different evaluative consequences. For instance, the existing literature on smiling suggests that this signal might be related to prosocial intent, as we tend to like and trust more people expressing happiness compared to neutral and negative expressions (Tortosa et al., 2013). In contrast to smiling, cooperation was found to be lower when participants were playing with angry partners and the lowest when angry individuals were paired with other angry individuals (Motro et al., 2016). The fact that facial expressions are salient stimuli during social interactions is unquestionable. Nonetheless, not all facial expressions can be considered *genuine* signals of a person's emotions and intentions. For instance, a smile represents the perfect example of an easily fakable expression: we can just display a smile to appear trustworthy and gain access to funds that would otherwise be denied (Krumhuber et al., 2014).

Since facial expressions are partially under cognitive control, relying solely on them to investigate unconscious moral decisions is reductive. In fact, we are exchanging numerous autonomic cues, outside our conscious control, that are slowly starting to receive the attention of the scientific field of emotion perception (Kret, 2015). One example of an unconscious autonomic signal is blushing: when we are experiencing a strong affective state, our skin gets perfused with oxygenated blood, which results in increased redness of the face (Prochazkova et al., 2018). How this signal is interpreted depends on the intensity of the facial redness. In general, people with a redder face are perceived as healthier and more attractive (Stephen et al., 2012). These findings were replicated also with virtual agents, as it has been discovered that higher levels of blushing in a virtual partner were associated with higher investment and higher self-report trust compared to the non-blushing partner (Dijk et al., 2011). Among various autonomic cues, pupil size has been suggested to have a substantial effect on the way we perceive others (Kret, 2018; van Breen et al., 2018). Dilated pupils are generally perceived positively and constricted pupils are negative. Notably, faces with large pupils are perceived as more attractive (Tombs & Silverman, 2004) and more trustworthy (Kret & De Dreu, 2017). In their recent experiment, van Breen and colleagues (2018) investigated whether this effect of pupil size could also apply to moral cognition. Participants were asked to play a coin-toss game in which they privately predicted the outcome of the coin flip and simply report whether they had won or not. For each reported win, participants earned money, tempting them to over-report correct guesses to increase personal gains. Before reporting the (in)correctness of their prediction, participants were exposed to a short video of either dilated or constricted pupils of the assumed interaction partner. Every time the participants won, the partner lost money. Results showed that people were more honest and reported fewer wins when watching a video of dilating pupils, compared to constricting (van Breen et al., 2018).

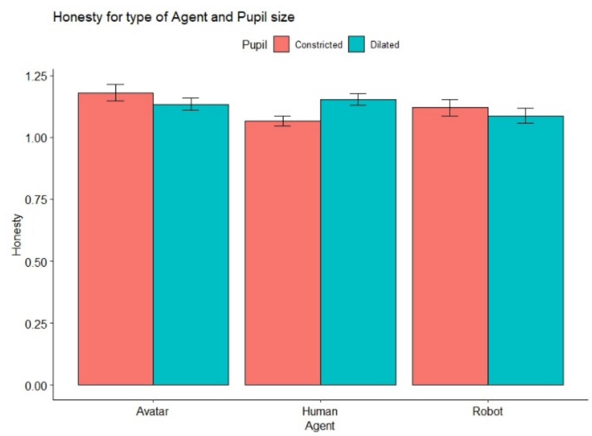
To what extent do social signals influence interactions/perceptions with artificial agents to same extent as human interaction? our study, we applied the coin-toss manipulation of van Breen and colleagues (2018) in real-life interaction with an avatar on a screen, a social robot, and a human confederate, all with manipulated pupil size (constricted or dilated). Assessing pupil size over other signals has to two main advantages: on one side, it is an unconscious cue and it can't be controlled by the participant; on the other side, since the experiment involves interaction with a machine-like robot, an avatar on a screen, and a human confederate, pupil size is easy to manipulate across the agents and manipulating it in a human being is a unique of an opportunity. Based on van Breen and colleagues (2018), we expect the participants to be more honest towards the confederate with dilated pupils than with constricted pupils. Previous research found that the presence of either a robot or a human evokes the same level of honesty (Hoffman et al., 2015). Pupil responses across emotions were similar between the pictures of the robots and those of humans (Reuten et al., 2018). Therefore, we predict that participants will lie to the same extent the robot and the human, but we still expect more honesty towards the robot with dilated pupils rather than constricted. Finally, we expect the people to lie most to the avatar, regardless of the pupil size, due to the uncanny valley effect (Mori et al., 2012).

2. MATERIALS AND METHODS

Fifteen female participants, with a mean age of 22.73(±2.54) played multiple rounds of a manipulated coin-toss game against either a female avatar on a screen, the small robot Cozmo by Anki, and a female confederates. They played a total of 18 rounds per agent (54 trials). For each win, participants were earning € 0.10 and the partners were losing € 0.10 or, for the robot, battery points. We manipulated the pupil size of the avatar and Cozmo with custom-made scripts, while for the confederate we used two pairs of tailored contact lenses with dilated (8mm) and constricted (3mm) pupils. These lenses made the confederate blind, ensuring that no subtle social cues were unconsciously mimicked from the confederate. Moreover, we assessed pupil dilation, skin conductance, and heart rate, but they will not be further analyzed in these preliminary results.

3. PRELIMINARY ANALYSIS AND RESULTS

We ran a multilevel linear model using Honesty as the dependent variable, Pupil Size, Type of Agent, and the interaction Pupil Size*Agent as fixed effect and Individuals as a random effect. The amount of money earned was taken as a measure of honesty. We did not found any effect of pupil size on the participants' honesty (β = -0.046, SE = 0.069, p=.506), neither of the type of agent (β = -0.113, SE = 0.697, p=.109). We did not find an interaction effect between Pupil Size and Type of Agent (β = -0.013, SE = 0.098, p=.181).



4. DISCUSSION

The preliminary results of this study suggest that neither pupil size nor the type of interaction partner affect the degree of reported wins, and therefore honesty. This was expected only for the avatar agent, but not for Cozmo and the confederate. This outcome might depend on the fact that pupil size in this study was static. Particularly for the human agent, it is likely that a static pupil does not evoke the same effect as a moving one, which would partially explain the lack of an effect in these regards. Besides, the static pupil size of the confederate enabled a perceptual mismatch between the unnatural pupil size and expected human features, an uncanny valley effect could have been created (Kätsyri et al., 2015). Furthermore, the confederate tended to have an averted gaze as the blindness induced by the contact lenses hindered the orientation of gaze. Averted gaze increases dishonesty regardless of pupil size as it is interpreted as reduced attention and interest from the interaction partner (Van Breen et al., 2018). An obvious limitation of the study is the very small sample size: different results may arise at the end of the data collection. Regardless of the limitation, the comparable results on the interaction with humans and robots are in line with previous research (Hoffman et al., 2015) and represent promising results towards a society where we share the environment with social robots.

ACKNOWLEDGEMENT

Funding research was supported by the European Research Council (ERC) to Mariska Kret. We would like to thank S. Lehtonen for the fruitful discussions, and Bas van Hal for customizing the contact lenses.

REFERENCES

Buon, M., Seara-Cardoso, A., & Viding, E. (2016). Why (and how) should we study the interplay between emotional arousal, Theory of Mind, and inhibitory control to understand moral cognition? *Psychonomic Bulletin and Review*, *23*(6), 1660–1680. https://doi.org/10.3758/s13423-016-1042-5

de Groot, J. H. B., Smeets, M. A. M., Kaldewaij, A., Duijndam, M. J. A., & Semin, G. R. (2012). Chemosignals Communicate Human Emotions. *Psychological Science*, *23*(11), 1417–1424. https://doi.org/10.1177/0956797612445317

Dijk, C., Koenig, B., Ketelaar, T., & de Jong, P. J. (2011). Saved by the blush: Being trusted despite defecting. *Emotion*, *11*(2), 313–319. https://doi.org/10.1037/a0022774

Ekman, P., & Rosenberg, E. L. (2012). What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS). In *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. https://doi.org/10.1093/acprof:oso/9780195179644.001.0001

Grant, E. C., & Mackintosh, J. H. (1963). A comparison of the social postures of some common laboratory rodents. *Cement and Concrete Research*, *2014*(860), 645–655. https://repositories.lib.utexas.edu/handle/2152/39127%0Ahttps://cris.brighton.ac.uk/ws/portalfiles/portal/4755978/Julius+Ojebode%27s+Thesis.pdf%0Ausir.salford.ac.uk/29369/1/Angela_Darvill_thesis_esubmission.pdf%0Ahttps://dspace.lboro.ac.uk/dspace-jspui/ha

Hoffman, G., Forlizzi, J., Ayal, S., Steinfeld, A., Antanitis, J., Hochman, G., Hochendoner, E., & Finkenaur, J. (2015). Robot Presence and Human Honesty: Experimental Evidence. *ACM/IEEE International Conference on Human-Robot Interaction*, *2015-March*(March), 181–188. https://doi.org/10.1145/2696454.2696487

Kret, M. (2015). Emotional expressions beyond facial muscle actions. A call for studying autonomic signals and their impact on social perception. *Frontiers in Psychology*, *6*(May), 1–10. https://doi.org/10.3389/fpsyg.2015.00711

Kret, M.E., & De Dreu, C. K. W. (2013). Oxytocin-motivated ally selection is moderated by fetal testosterone exposure and empathic concern. *Frontiers in Neuroscience*, *7*(7 JAN), 1–9. https://doi.org/10.3389/fnins.2013.00001

Kret, M.E., Prochazkova, E., Sterck, E. H. M., & Clay, Z. (2020). Emotional expressions in human and non-human great apes. *Neuroscience and Biobehavioral Reviews*, *115*(January), 378–395. https://doi.org/10.1016/j.neubiorev.2020.01.027

Kret, M.E. (2018). The role of pupil size in communication. Is there room for learning? *Cognition and Emotion*, *32*(5), 1139–1145. https://doi.org/10.1080/02699931.2017.1370417

Kret, Mariska E., & De Dreu, C. K. W. (2017). Pupil-mimicry conditions trust in partners: Moderation by oxytocin and group membership. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1850), 1–10. https://doi.org/10.1098/rspb.2016.2554

Krumhuber, E. G., Likowski, K. U., & Weyers, P. (2014). Facial Mimicry of Spontaneous and Deliberate Duchenne and Non-Duchenne Smiles. *Journal of Nonverbal Behavior*, *38*(1), 1–11. https://doi.org/10.1007/s10919-013-0167-8

Mori, M., MacDorman, K. F., & Kageki, N. (2012). The uncanny valley. *IEEE Robotics and Automation Magazine*, *19*(2), 98–100. https://doi.org/10.1109/MRA.2012.2192811

Motro, D., Kugler, T., & Connolly, T. (2016). Back to the basics: how feelings of anger affect cooperation. *International Journal of Conflict Management*, *27*(4), 523–546. https://doi.org/10.1108/IJCMA-10-2015-0068

Reuten, A., van Dam, M., & Naber, M. (2018). Pupillary responses to robotic and human emotions: The uncanny valley and media equation confirmed. *Frontiers in Psychology*, *9*(MAY), 1–12. https://doi.org/10.3389/fpsyg.2018.00774

Ross, M. D., Owren, M. J., & Zimmermann, E. (2010). The evolution of laughter in great apes and humans. *Communicative & Integrative Biology*, *3*(2), 191–194. https://doi.org/10.4161/cib.3.2.10944

Tombs, S., & Silverman, I. (2004). Pupillometry - A sexual selection approach. *Evolution and Human Behavior*, *25*(4), 221–228. https://doi.org/10.1016/j.evolhumbehav.2004.05.001

Tortosa, M. I., Lupiáñez, J., & Ruz, M. (2013). Race, emotion and trust: An ERP study. *Brain Research*, *1494*(May 2018), 44–55. https://doi.org/10.1016/j.brainres.2012.11.037

van Breen, J. A., De Dreu, C. K. W., & Kret, M. E. (2018). Pupil to pupil: The effect of a partner's pupil size on (dis)honest behavior. *Journal of Experimental Social Psychology*, *74*(November 2017), 231–245. https://doi.org/10.1016/j.jesp.2017.09.009