

The Emotional Dilemma: Influence of a Human-like Robot on Trust and Cooperation

Dennis Becker*, Diana Rueda*, Felix Beese, Brenda Scarleth Gutierrez Torres†, Myriem Lafdili, Kyra Ahrens, Di Fu, Erik Strahl, Tom Weber, Stefan Wermter

Abstract—Increasing anthropomorphic robot behavioral design could affect trust and cooperation positively. However, studies have shown contradicting results and suggest a task-dependent relationship between robots that display emotions and trust. Therefore, this study analyzes the effect of robots that display human-like emotions on trust, cooperation, and participants’ emotions. In the between-group study, participants play the coin entrustment game with an emotional and a non-emotional robot. The results show that the robot that displays emotions induces more anxiety than the neutral robot. Accordingly, the participants trust the emotional robot less and are less likely to cooperate. Furthermore, the perceived intelligence of a robot increases trust, while a desire to out-compete the robot can reduce trust and cooperation. Thus, the design of robots expressing emotions should be task dependent to avoid adverse effects that reduce trust and cooperation.

I. INTRODUCTION

Trust and cooperation affect humans’ willingness to interact with robots [1], and are essential to establish successful human-robot interaction [2], [3], [4]. Human-robot trust is composed of human-related, robot-related, and environment-related factors [5], [6]. To the human-related factors also belong emotions, which can influence trust [7] and have been suggested as an integral part of forming trust [8]. Humans’ emotional states have been shown to affect trust in human-robot interaction [9], where positive emotions such as happiness increase trust [10] and negative emotions such as anxiety reduce trust [11]. Likewise, trust in a robot influences the willingness to cooperate with the robot [12], which can be more pronounced in cooperative tasks [13]. Cooperation is frequently associated with an initial investment and trust in the collaborator in expectation of a greater mutual benefit [14].

While robots are becoming steadily human-like, research has been conducted on how anthropomorphic a robot should be designed for a specific task [15], [16]. Studies suggest that robots that act more human-like by expressing emotions are perceived as more trustworthy [17], [18] and encourage cooperation [19], [20]. Specifically, trust is fostered when the robot’s expressed emotions are congruent and engages in social conversation [21]. Contrary, to this relationship of

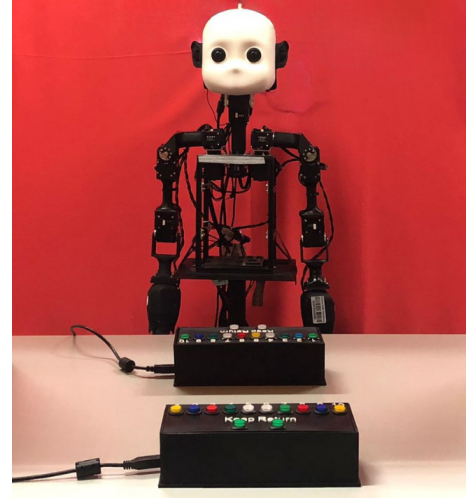


Fig. 1: Coin entrustment game to measure the difference in trust and cooperation with an emotional and non-emotional robot.

increased anthropomorphism and trust, recent findings suggest a task dependency [22], [23], where a robot expressing social cues for critical tasks receives less trust than a neutral robot [24]. Thus, research on human-like robots and their effect on trust and cooperation with respect to emotions is required.

For measuring trust and cooperation in human-robot interaction, we adopt a variation of the prisoner’s dilemma, the coin entrustment game [25]. The coin entrustment game allows participants to decide on the amount of trust in the robot in expectation of an increased reward. In the conducted between-group experiment, an emotional robot is compared to a non-emotional robot. While the non-emotional robot portrays neutral emotions, the emotional robot displays happiness or sadness, depending on the participant’s cooperation. The emotions are conveyed utilizing multiple modalities, such as speech interjections that deploy emotional prosody [26], body gestures [27], and facial expressions [28].

II. RELATED WORK

The prisoner’s dilemma is a game theory-based [29] social dilemma that has frequently been used to study trust and cooperation in social situations [30]. The initial definition accounts for one round, where defecting is the preferred strategy to avoid punishment. However, experiments could not reliably explain the emergence of participants coop-

*These authors contributed equally.

†Stipendiary from CONACyT and DAAD

The authors gratefully acknowledge partial support from the German Research Foundation DFG under project CML, TRR169 and LeCareBot.

All of the authors are with the Knowledge Technology Group, Department of Informatics, Universität Hamburg, Vogt-Kölln-Straße 30, Hamburg D-22527, Germany.

Many thanks to Moritz Lahann, Ramtin Nouri, Jose Angel Sanchez Castro, Sebastian Stelter, and Denyse Uwase.

erating [31]. Thus, the repeated prisoner's dilemma [32] was suggested, where the game is played over consecutive rounds. With an undisclosed number of rounds, a cooperative strategy is encouraged and leads to a superior outcome for both players [33]. Since the prisoner's dilemma has a predefined payout matrix, a coin entrustment variant of the game was proposed [25]. In the coin entrustment game, a player's pay off and risk are dependent on the amount of entrusted coins, which allows measuring trust in the other player to cooperate and own cooperation separately.

Research on social dilemmas suggests that humans act more selfishly and less emotional when interacting with artificial agents and robots [34], [35], [36]. When facing an artificial agent, humans tend to derive decisions more rational [37], and the influence of emotions on decision-making is mitigated [38]. These findings are supported by physiological measures of skin conductivity and heart rate in a betting scenario against a computerized agent and a human [39]. Studies on the prisoner's dilemma in human-robot interaction suggest that cooperative behavior with humans tends to be higher than with robots [40], [41]. When cooperating with a robot, self-reported measures indicate that decision-making is less emotionally driven and imply a notable lack of empathy that could result in a more selfish behavior [42]. However, when the robot is facing consequences for losing in the prisoner's dilemma, such as erasing its memory, individuals' game-play becomes more empathetic towards the robot [43]. Similarly, a robot expressing moral values leads to less competitive game-play, while a robot expressing emotions can increase competitive behaviour [44]. Contradicting these cooperative preferences, research on the n-player prisoner's dilemma with human and robotic players shows that cooperation with robotic players is higher, potentially due to the unpredictability of human players' strategy [45]. Correspondingly, for the coin entrustment game, more trust towards robotic players is reported [46].

Non-verbal communication signals and indicates willingness to cooperate in social dilemma and trust games [47], [48]. People who are likely to cooperate are emotionally more expressive than people that are inclined to defect [49]. Additionally, research has shown that humans are sensitive to emotional display and facial expressions when evaluating a person's cooperativeness [50]. Specifically, smiling can evoke and indicate cooperation and is perceived as an indicator of the other's trustworthiness [51], whereas contemptuous behavior is associated with defect [52]. Likewise, body motions [53] and gestures [54] of non-verbal communication are utilized to judge the partner's tendency to cooperate. Moreover, vocal features affect trustworthiness and cooperation [55], where emotional expressiveness [56] and a happy-sounding voice can foster trust [57]. Further, congruence between a robot's behavior and its voice influences trust [58] and convergence of a robot's speech influences cooperation [59]. Despite the effect of non-verbal communication, studies on the prisoner's dilemma with a virtual agent suggest, that nonverbal behavior might be too

subtle to be recognized [60]. Subsequently, no difference in the cooperation between a robot expressing sad and angry emotions could be shown; however, recognizing the robot's emotions can reduce the participants' cooperation [61].

III. METHODOLOGY

A. Participants

After receiving a positive response from the Ethics Commission of the Department of Informatics at the University of Hamburg, participants were recruited from the university's campus, and the experiment was conducted with 47 participants. Out of these participants, three were excluded due to technical issues, and an additional three participants were discarded based on the control question. As control question, the participants rated the robot's emotionality on a 5-point scale ranging from Emotional to Non-Emotional. Participants who considered the non-emotional robot as Emotional, and participants that labeled the emotional robot as Non-Emotional were excluded. This results in a data set of 41 participants for analysis, consisting of 20 participants in the emotional and 21 participants in the non-emotional group.

Of these participants, 41.5% were female and 58.5% were male. The participants' age distribution was 78% between 18–29 years, 19.5% between 30–39 years, and 2.5% between 40–49 years. Four of the participants self-reported a high familiarity with humanoid robots, two participants stated a high familiarity with negotiation games, and one participant self-reported a high familiarity with both humanoid robots and negotiation games.

B. Experiment Design

The experiment implements the coin entrustment game, and the participants play a total of 16 rounds, where each round consists of two stages. During the first stage, both the participant and the robot secretly entrust between one and ten coins to the other player. For the robot, the number of entrusted coins follows the design of the experiment conducted in [46]. Specifically, the number of entrusted coins depends on the payoff from the previous round and at least one coin is entrusted. The number of entrusted coins is expressed as follows:

$$E(p) = \begin{cases} \lceil \min(10 + \frac{p-10}{1.5}, 10) \rceil & \text{if } p > 0, \\ 1 & \text{if } p \leq 0, \end{cases}$$

where p is the payoff of the previous round and the initial entrustment in the first round is three coins.

In the second stage, the number of entrusted coins is revealed, and both players decide whether to keep or return the coins entrusted to them. If the entrusted coins are returned, the other player receives double the amount of the entrustment. If the player instead decides to keep the coins, these coins are added to their coins.

To analyze the effect of an emotional robot on trust and cooperation, the robot always returns the entrusted coins, except for round eight. The robot's cooperative strategy was



Fig. 2: Interaction with the robot in the experiment.

chosen since research suggests that a strict strategy can hinder cooperation [62] and that more generous game-play encourages cooperation [63].

After round eight, trust and cooperation with the robot has to be re-established. The robot encourages regaining trust by acknowledging the trust violation, “Perhaps I tried too hard to maximize my coins. I should not do that again.” Hereby, the robot attempts to restore the broken trust. However, the statement is deliberately left open as promises have shown a strong effect on trust repair [64], [65].

To encourage cooperative game-play, the participants are unaware of the number of rounds to be played. Further, the participants are instructed to score the highest number of coins out of all participants, rather than competing against the robot. As an incentive for participants, the number of obtained coins was anonymously placed on a public leader board.

C. Experiment Setup

To avoid recognition of the study objective, the experiment is presented as a carnival attraction. The general setup is illustrated in Figure 1, and the interaction with the robot during the experiment is shown in Figure 2.

For a detailed illustration of the experiment setup, a schematic overview is provided in Figure 3. During the experiment, the participant is seated in front of a table and faces the robot. Both players (Participant and Robot) use a controller (Controller 1 and Controller 2) to input their decisions during the game. Each controller has two rows of light-up buttons. The first row allows the participant to select the number of coins to entrust by pressing one of the buttons labeled from 1 to 10. Similarly, with the second row of buttons, the participant can decide whether to keep or return the other player’s entrustment. During both game stages, the robot utilizes a button-press animation, where the robot moves its hand close to the buttons and focuses its gaze on the controller. Afterward, the lighted button of the robot’s controller provides visual confirmation to the participant about the robot’s decision.

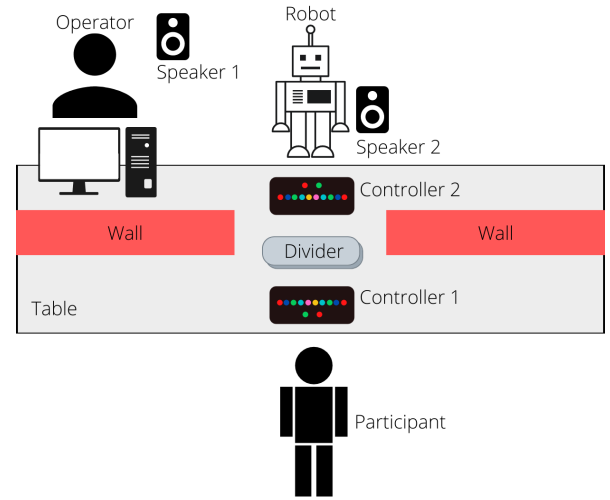


Fig. 3: Schematic illustration of the experiment setup. The human operator is not visible to the participant.

To block the players’ view of the other player’s controller during the decision-making of both game stages, a motorized divider is used. Additionally, the stages of each round are clearly separated by a narrator’s voice that announces the decisions of each player.

The difference between the narrator and the robot is emphasized by utilising two spatially separated loudspeakers (Speaker 1 and Speaker 2). The first loudspeaker is located close to the experiment operator and is used by the narrator for game-related information and announcements. The second loudspeaker is used for the robot’s voice and is located inside the robot’s body frame.

Before the experiment, a pilot study with eight participants was conducted. The data evaluation of the pilot study indicated that the participants understood the provided instructions and chose cooperative strategies instead of exploiting the robot’s cooperative behavior. Further, it was suggested that the participants noticed the difference in emotions between both robots.

D. Emotional Expressions

In the experiment, the Neuro-Inspired COmpanion (NICO) [66] robot was used, which was designed to combine neurorobotics with human-robot interaction. The software for the experiment is implemented in the Robot Operating System (ROS) [67] and is composed of multiple modules to control the robots’ interactions during the game, speech, gestures, and facial expressions.

1) *Speech*: The speech generation module utilizes Google Text-to-Speech¹ (gTTS) to synthesize speech from text for the robot’s voice. Spoken sentences comprise instructions for the beginning of the game, the game-play, and the final interaction. For the robot’s responses after each game round, sentence variations are included to maintain the perceived robot’s animacy. Depending on the participant’s choice, 18

¹<https://github.com/pndurette/gTTS>



Fig. 4: Facial expressions of the NICO robot. From left to right: happy, sad, neutral.

unique sentences for defect and 38 sentences for cooperation are implemented. To differentiate between the robots, the emotional robot uses interjections to express emotions [68]. When the participant returns the coins, the emotional robot might say: “*Hooray!* You gave the coins back”, while the non-emotional robot limits the answer to: “You gave the coins back”. In contrast, when the participant keeps the coins, the answer “I see you kept the coins” for the non-emotional robot, is adjusted to “*Owww*, I see you kept the coins” for the emotional robot.

The emotional interjections have specific prosody to convey either happy emotions when the participant cooperates, or sad emotions in the case of defect. The intonation of these interjections was simulated by applying a transfer learning text-to-speech model [69] to generate the utterances from audio references expressing either happiness or sadness. Afterward, the generated utterances were post-processed for artifact filtering, pitch, and tempo to match the robot’s voice. For the experiment, six different happiness-conveying interjections and five sadness-conveying interjections are utilized.

2) *Gestures and Facial Expressions*: The robot’s gestures are used to increase liveliness and display emotional behavior. For the non-emotional robot, three neutral gestures (looking in a direction, pointing, and a hand gesture) are implemented [68]. The happy and sad gestures of the emotional robot are designed to express emotions depending on the participant’s decision to cooperate or defect. For instance, gestures that include lowered arms and head are used to express sadness, while opening the arms is associated with happiness [70]. For each emotion, three different happy and sad gestures are implemented and randomly displayed after each round.

In addition to the emotion-conveying gestures and interjections, the robot possesses LED arrays under its translucent face-plate in the mouth and eye area, which allows for displaying seven universal emotions [71]. Depending on the participant’s decision to cooperate or defect, the robot shows either a happy or sad facial expression. During the remainder of the game and in the non-emotional group, a neutral facial expression is shown. The facial expressions of the NICO robot are depicted in Figure 4.

E. Questionnaires and Measurements

For the evaluation of an emotional and non-emotional robot, the Godspeed [72] and Discrete Emotions [73] questionnaires are assessed, and the participants’ entrusted coins and decisions to cooperate or defect are recorded.

The Godspeed questionnaire provides standardized metrics about the perceived robot’s anthropomorphism, animacy, likeability, intelligence, and safety. Since the conducted experiment does not pose safety concerns, the category of perceived safety was omitted.

The Discrete Emotions questionnaire provides insight into the participants’ emotions during the experiment. Specifically, the questionnaire measures eight discrete emotions: anger, anxiety, desire, disgust, fear, happiness, relaxation, and sadness.

In each game round, the number of entrusted coins is a direct measure of the participant’s trust in the robot. Likewise, the decision to keep or return the robot’s coins is a direct measure of cooperation.

F. Study Design and Procedure

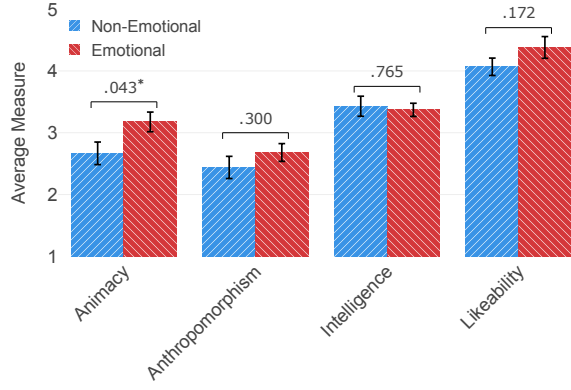
The experiment is conducted in a between-subject design with two groups. The robot in the emotional group shows happy and sad emotions through facial expressions, gestures, and utterances. These displays of emotion are shown at the end of each round. The robot in the non-emotional group also demonstrates facial expressions, speech, and gestures. However, the non-emotional robot’s gestures are replaced with neutral gestures and a neutral facial expression is used.

The participants signed a consent form agreeing to participate in the experiment, and were randomly assigned to one of the two experimental conditions. After the participant’s demographics are assessed, the participant is handed a scenario description with an introduction to the experiment. Then, the participant is brought to a neighboring room where the experiment is conducted and seated in front of the robot. An introduction to the controller is provided, and a trial round introduces the experiment procedure. Afterward, the experiment begins. After the experiment, the participant is guided back to the initial room and the questionnaires are assessed.

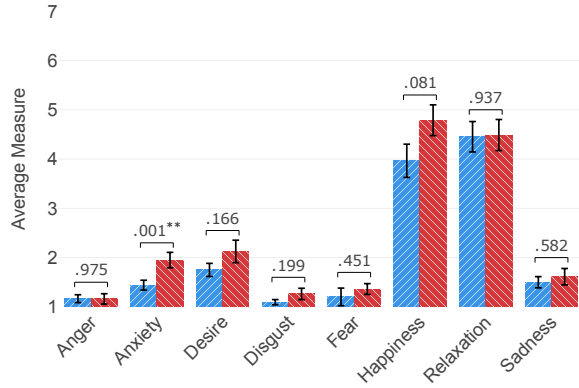
IV. RESULTS

An illustration of the Godspeed and Discrete Emotions questionnaire items with standard errors is shown in Figure 5. A Student’s t -test [74] of the Godspeed questionnaire items shows a significant difference in animacy ($p = .043$) of the emotional robot ($M = 3.18$, $SD = 0.71$) in contrast to the non-emotional robot ($M = 2.66$, $SD = 0.84$). Regarding the participants’ emotions, the Student’s t -test suggests a significant difference in the anxiety item ($p = .001$) between the emotional group ($M = 1.95$, $SD = 0.70$) and non-emotional group ($M = 1.44$, $SD = 0.46$).

Over the course of the experiment, the participants in both groups entrusted coins to the robot and decided to cooperate or defect. The on-average entrusted coins and cooperation rate with standard errors are shown in Figure 6. A Mann–Whitney U test [75] shows that the non-emotional group ($M = 6.88$, $SD = 2.94$) entrusted significantly more coins to the robot over the course of the experiment ($p = .015$) than the emotional group ($M = 6.38$, $SD = 2.85$). For the cooperation rate, a Fisher’s exact test [76] shows that the



(a)



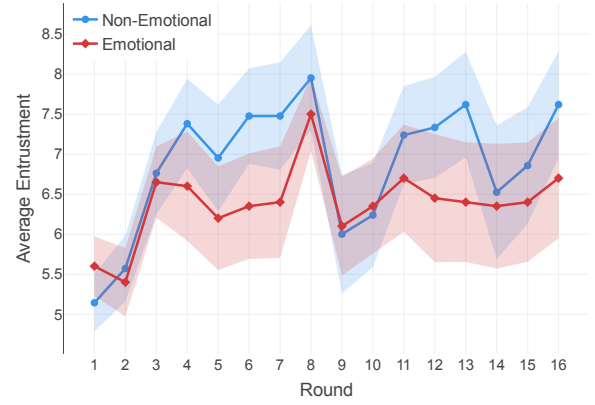
(b)

Fig. 5: Godspeed (a) and Discrete Emotions (b) items for the emotional and non-emotional group. Values indicate p value, $*p < .05$, $**p < .01$

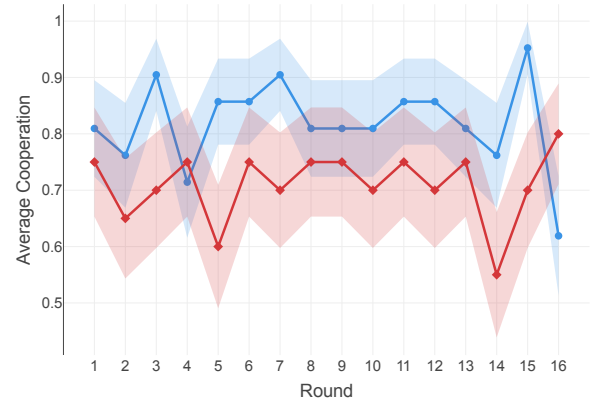
non-emotional group ($M = 0.82$, $SD = 0.39$) was significantly more likely to cooperate ($p = .001$) than the emotional group ($M = 0.71$, $SD = 0.45$).

Analyzing the difference of the average entrustment for both groups over two consecutive rounds, a Mann–Whitney U test suggests that the emotional group entrusted significantly more coins ($p = .038$) in the third round ($M = 6.65$, $SD = 1.98$) in comparison to the second round ($M = 5.40$, $SD = 1.90$). In the ninth round, after the robot defected, a noticeable decrease in the average entrustment can be noticed. Especially, the average entrustment of the non-emotional group significantly ($p = .039$) decreased from the eighth round ($M = 7.95$, $SD = 3.01$) to the ninth round ($M = 6.00$, $SD = 3.39$).

For the cooperation rate, a Fisher’s exact test shows a significant difference ($p = .045$) between the emotional group ($M = 0.70$, $SD = 0.46$) and non-emotional group ($M = 0.95$, $SD = 0.21$) in round 15, and a difference in cooperation for the non-emotional group ($p = .020$) between round 15 ($M = 0.95$, $SD = 0.21$) and round 16 ($M = 0.62$, $SD = 0.49$). Although the robot’s defect noticeably reduced the entrustment in the ninth round, the cooperation rate remained



(a)



(b)

Fig. 6: Average entrustment (a) and cooperation rate (b) over the consecutive rounds.

unaffected. Further, the participants exhibited a variety of entrustment strategies as indicated by the standard error.

To estimate the relationship between the assessed items of the Godspeed and Discrete Emotions questionnaire, the Spearman correlation [77] of the on-average entrusted coins and cooperation rate is estimated and shown in Table I. The estimated correlations show a relationship between the entrusted coins and the perceived intelligence of the robot, where a robot that is perceived as more intelligent receives a larger entrustment. On the contrary, high positive anticipation as measured with the Desire item results in fewer entrusted coins and can decrease the likelihood of cooperating. The amount of coins entrusted to the robot exhibits a strong correlation with the likelihood of cooperation.

V. DISCUSSION

The assessed Godspeed and Discrete Emotions questionnaires show significant differences between both experiment groups. Although both robots utilize gestures, facial expressions, and voice, the emotional robot was perceived as more animated than the non-emotional robot. The robot’s emotional reaction to the participant’s choice to either cooperate or defect, might have led the participants to perceive the

Entrustment													
Item	Godspeed				Discrete Emotions								
	Animacy	Likeability	Anthropomorphism	Intelligence	Anger	Fear	Disgust	Sadness	Desire	Happiness	Relaxation	Anxiety	
ρ	-.210	-.099	-.229	.332	.168	-.055	-.118	-.052	-.352	-.221	.084	-.029	
p value	.189	.536	.149	.034	.294	.734	.463	.746	.024	.166	.600	.858	

Cooperation													
Item	Godspeed				Discrete Emotions								
	Animacy	Likeability	Anthropomorphism	Intelligence	Anger	Fear	Disgust	Sadness	Desire	Happiness	Relaxation	Anxiety	Entrustment
ρ	-.228	-.196	-.155	.150	.222	-.078	-.217	-.035	-.418	-.171	.250	-.129	.733
p value	.151	.218	.332	.349	.163	.626	.173	.826	.007	.284	.115	.421	<.001

TABLE I: Correlation between the entrustment, the cooperation rate, and the measured questionnaire items.

emotional robot as more lively and animated. Further, the participants in the emotional group experienced stronger emotions toward the emotional robot. Specifically, participants in the emotional experiment group exhibited a greater level of anxiety than the non-emotional group.

Evidently, the display of human-like emotions by the robot caused the participants more discomfort in the researched scenario. In the social dilemma that requires cooperation and reliance on the other player, participants' negative emotions reduced trust and cooperation. This indicates that the participants associated a more human-like behavior with a higher unpredictability of the robot's strategy, whereas a more neutral robot could be associated with pre-programmed and predictable actions to return the participant's entrustment. Likewise, the display of human emotions encouraged a more competitive game-play, as shown by the lower cooperation rate with the emotional robot.

The analysis of changes in trust and cooperation over two consecutive rounds shows that both experiment groups behave similarly and that the robot defecting results in a significant trust loss for the non-emotional group. However, the cooperation rate remains unaffected. This might be attributed to the robot's trust repair attempt by acknowledging the trust violation. Afterward, the robot continues to cooperate, which rebuilds the participants' trust.

The correlation analysis of the entrusted coins and the measured questionnaire items shows that the participants are likely to entrust more coins to a robot that is perceived as intelligent. This suggests that an as intelligent perceived robot could positively influence the establishment of trust and willingness to cooperate. The negative correlation between the Desire item of the Discrete Emotions questionnaire and the entrusted coins indicates that a participant's desire to succeed can lead to a more cautious entrustment. Since this task requires cooperation, a strong desire to succeed might result in smaller entrustments to assess the partner's willingness to cooperate. For the cooperation rate with the robot, the negative correlation of the desire to outcompete the other player could indicate that the participants initially establish cooperation with the robot, which results in larger entrustments from the robot, with the intent to keep the robot's coins if the amount appears profitable. These results underline that emotions directly affect trust and cooperation in the coin entrustment game, and suggest that negative emotions towards the robot might hinder the formation of trust and cooperation.

Contrary to the desire to accumulate coins by defecting the robot, the amount of coins entrusted to the robot exhibits a strong correlation with cooperation. Since the coins represent trust, it appears plausible that a larger entrustment reflects the participant's willingness to cooperate. However, the correlation is symmetrical and does not show that trust is independent of cooperation. The willingness to cooperate could affect the extent of trust in the robot.

VI. LIMITATIONS

Despite the difference in trust and cooperation between an emotional and non-emotional robot, the findings have some limitations that should be addressed in future work. The sample size of the experiment could be increased, which might lead to the identification of additional effects and emotions that influence trust and willingness to cooperate with a robot in a social dilemma. Personality traits and trust disposition should be considered to estimate the effect and interplay between emotions and trust more accurately. Additionally, the experiment was conducted in a laboratory environment, and the effects on trust and cooperation could be more pronounced in a real-world environment. Likewise, a more realistic-looking humanoid robot could have a stronger effect.

VII. CONCLUSION

This study investigated the effects of an emotional robot on trust and cooperation in the coin entrustment game. Specifically, a robot that conveys emotions by utilizing prosodic speech interjections, facial expressions, and body gestures was compared to a robot that portrays neutral emotions. The robot was programmed to cooperate throughout the experiment, except for the experiment's midpoint, to encourage the participants' trust and cooperation. The participants' entrusted coins, cooperation, perception of the robot, and emotions during the experiment were evaluated.

The results show that the participants experienced more anxiety during the interaction with the emotional robot. Accordingly, the emotional robot received less trust and participants were less likely to cooperate. The perceived robot's intelligence affected trust, and robots that exhibit emotions might be perceived as less suitable and competent for a cooperative task. However, emotional robots could be more resilient to breaches of trust, which points towards differences in interaction with and perception of emotional robots. These differences might be based on humans' pre-conceived assumptions about robots and how they should

operate. Furthermore, the results are consistent with findings in the literature that suggest that trust in anthropomorphic robots is task-dependent, and provide evidence that depending on the task, trust and cooperation with a neutral robot can be higher than with a robot that displays human-like emotions.

Subsequent experiments could research the specific effects of robots displaying either negative or positive emotions on trust and cooperation. Further, the granularity of the displayed emotions could be considered, to provide a better understanding of the impact of emotions in human-robot interaction. Consequently, leading to guidelines and insights on applications for which emotional robots could be beneficial or have adverse effects.

This research underlines the importance of social interaction between robots and humans, where an anthropomorphic robot can reduce trust. Especially, for critical or safety-related tasks, where cooperation is essential and mistakes can result in consequences, the relationship between emotions and the effect on trust and cooperation has to be considered.

REFERENCES

- [1] A. Freedy, E. DeVisser, G. Weltman, and N. Coeyman, "Measurement of trust in human-robot collaboration," *2007 International Symposium on Collaborative Technologies and Systems*, pp. 106–114, 2007.
- [2] R. E. Yagoda and D. J. Gillan, "You want me to trust a ROBOT? The development of a human-robot interaction trust scale," *International Journal of Social Robotics*, vol. 4, no. 3, pp. 235–248, 2012.
- [3] Z. R. Khavas, "A review on trust in human-robot interaction," pp. 1–10, 2021.
- [4] M. Lewis, K. Sycara, and P. Walker, "The role of trust in human-robot interaction," in *Studies in Systems, Decision and Control*, vol. 117, pp. 135–159, 2018.
- [5] P. A. Hancock, D. R. Billings, K. E. Schaefer, J. Y. Chen, E. J. De Visser, and R. Parasuraman, "A meta-analysis of factors affecting trust in human-robot interaction," *Human Factors*, vol. 53, no. 5, pp. 517–527, 2011.
- [6] K. E. Schaefer, "Measuring trust in human robot interactions: Development of the "trust perception scale-HRI"," in *Robust Intelligence and Trust in Autonomous Systems*, pp. 191–218, Boston, MA: Springer US, April 2016.
- [7] W. S. Lee and M. Selart, "The impact of emotions on trust decisions," in *Handbook on Psychology of Decision-Making: New Research*, pp. 235–248, 2012.
- [8] L. Miller, J. Kraus, F. Babel, and M. Baumann, "More than a feeling—Interrelation of trust layers in human-robot interaction and the role of user dispositions and state anxiety," *Frontiers in Psychology*, vol. 12, pp. 1–18, April 2021.
- [9] R. Stock-Homburg, "Survey of emotions in human-robot interactions: Perspectives from robotic psychology on 20 years of research," *International Journal of Social Robotics*, vol. 14, no. 2, pp. 389–411, 2022.
- [10] C. K. Stokes, J. B. Lyons, K. Littlejohn, J. Natarian, E. Case, and N. Speranza, "Accounting for the human in cyberspace: Effects of mood on trust in automation," in *2010 International Symposium on Collaborative Technologies and Systems*, pp. 180–187, 2010.
- [11] J. Kraus, D. Scholz, E. M. Messner, M. Messner, and M. Baumann, "Scared to trust? – Predicting trust in highly automated driving by depressiveness, negative self-evaluations and state anxiety," *Frontiers in Psychology*, vol. 10, pp. 1–15, January 2020.
- [12] G. R. Jones and J. M. George, "The experience and evolution of trust: Implications for cooperation and teamwork," *Academy of Management Review*, vol. 23, no. 3, pp. 531–546, 1998.
- [13] A. Weidemann and N. Rußwinkel, "The role of frustration in human-robot interaction – What is needed for a successful collaboration?," *Frontiers in Psychology*, vol. 12, pp. 1–17, March 2021.
- [14] D. G. Rand and M. A. Nowak, "Human cooperation," *Trends in Cognitive Sciences*, vol. 17, pp. 413–425, August 2013.
- [15] K. S. Haring, Y. Matsumoto, and K. Watanabe, "How do people perceive and trust a lifelike robot," in *Lecture Notes in Engineering and Computer Science*, vol. 1, pp. 425–430, October 2013.
- [16] M. Desai, K. Stubbs, A. Steinfeld, and H. Yanco, "Creating trustworthy robots: Lessons and inspirations from automated systems," in *Adaptive and Emergent Behaviour and Complex Systems - Proceedings of the 23rd Convention of the Society for the Study of Artificial Intelligence and Simulation of Behaviour, AISB 2009*, pp. 49–56, 2009.
- [17] N. Martelaro, V. C. Nneji, W. Ju, and P. Hinds, "Tell me more: Designing HRI to encourage more trust, disclosure, and companionship," in *ACM/IEEE International Conference on Human-Robot Interaction*, pp. 181–188, April 2016.
- [18] M. Hashemian, R. Paradedda, C. Guerra, and A. Paiva, "Do you trust me? Investigating the formation of trust in social robots," in *Lecture Notes in Computer Science*, vol. 11805, pp. 357–369, Springer International Publishing, 2019.
- [19] Y. Takahashi, Y. Kayukawa, K. Terada, and H. Inoue, "Emotional expressions of real humanoid robots and their influence on human decision-making in a finite iterated prisoner's dilemma game," *International Journal of Social Robotics*, vol. 13, no. 7, pp. 1777–1786, 2021.
- [20] A. Andriella, H. Siqueira, D. Fu, S. Magg, P. Barros, S. Wermter, C. Torras, and G. Alenyà, "Do I have a personality? Endowing care robots with context-dependent personality traits," *International Journal of Social Robotics*, vol. 13, no. 8, pp. 2081–2102, 2021.
- [21] R. B. Paradedda, M. Hashemian, R. A. Rodrigues, and A. Paiva, "How facial expressions and small talk may influence trust in a robot," in *Lecture Notes in Computer Science*, vol. 9979, pp. 169–178, 2016.
- [22] L. Onnasch and C. L. Hildebrandt, "Impact of anthropomorphic robot design on trust and attention in industrial human-robot interaction," *ACM Transactions on Human-Robot Interaction*, vol. 11, pp. 1–24, March 2022.
- [23] E. Roesler, D. Manzey, and L. Onnasch, "A meta-analysis on the effectiveness of anthropomorphism in human-robot interaction," *Science Robotics*, vol. 6, no. 58, 2021.
- [24] A. Rossi, K. Dautenhahn, K. Lee Koay, and M. L. Walters, "How social robots influence people's trust in critical situations," in *29th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2020*, pp. 1020–1025, 2020.
- [25] T. Yamagishi, S. Kanazawa, R. Mashima, and S. Terai, "Separating trust from cooperation in a dynamic relationship: Prisoner's dilemma with variable dependence," *Rationality and Society*, vol. 17, no. 3, pp. 275–308, 2005.
- [26] R. Savery, L. Zahray, and G. Weinberg, "Emotional musical prosody for the enhancement of trust: Audio design for robotic arm communication," *Paladyn*, vol. 12, pp. 454–467, December 2021.
- [27] S. Saunderson and G. Nejat, "How robots influence humans: A survey of nonverbal communication in social human-robot interaction," *International Journal of Social Robotics*, vol. 11, no. 4, pp. 575–608, 2019.
- [28] N. Rawal and R. M. Stock-Homburg, "Facial emotion expressions in human-robot interaction: A survey," *International Journal of Social Robotics*, vol. 14, no. 7, pp. 1583–1604, 2022.
- [29] R. Axelrod, *The evolution of cooperation*. New York: Basic, 1984.
- [30] P. A. Van Lange, J. Joireman, C. D. Parks, and E. Van Dijk, "The psychology of social dilemmas: A review," *Organizational Behavior and Human Decision Processes*, vol. 120, no. 2, pp. 125–141, 2013.
- [31] J. Grandgirard, D. Poinot, L. Krespi, J. P. Nénon, and A. M. Cortesero, "Costs of secondary parasitism in the facultative hyperparasitoid *Pachycrepoideus dubius*: Does host size matter?," *Entomologia Experimentalis et Applicata*, vol. 103, no. 3, pp. 239–248, 2002.
- [32] A. Rapoport, A. Chammah, and C. Orwant, *Prisoner's Dilemma: A study in conflict and cooperation*. Ann Arbor paperbacks, University of Michigan Press, 1965.
- [33] J. García and M. van Veelen, "No strategy can win in the repeated prisoner's dilemma: Linking game theory and computer simulations," *Frontiers Robotics AI*, vol. 5, pp. 1–14, August 2018.
- [34] M. Chugunova and D. Sele, "We and it: An interdisciplinary review of the experimental evidence on how humans interact with machines," *Journal of Behavioral and Experimental Economics*, vol. 99, no. 20, 2022.
- [35] C. March, "Strategic interactions between humans and artificial intelligence: Lessons from experiments with computer players," *Journal of Economic Psychology*, vol. 87, 2021.

- [36] R. Hoegen, G. Stratou, G. M. Lucas, and J. Gratch, "Comparing behavior towards humans and virtual humans in a social dilemma," *Lecture Notes in Computer Science*, vol. 9238, pp. 452–460, 2015.
- [37] D. Houser and R. Kurzban, "Revisiting kindness and confusion in public goods experiments," *American Economic Review*, vol. 92, no. 4, pp. 1062–1069, 2002.
- [38] E. Schniter, T. W. Shields, and D. Sznycer, "Trust in humans and robots: Economically similar but emotionally different," *Journal of Economic Psychology*, vol. 78, p. 102253, 2020.
- [39] T. Teubner, M. Adam, and R. Riordan, "The impact of computerized agents on immediate emotions, overall arousal and bidding behavior in electronic auctions," *Journal of the Association for Information Systems*, vol. 16, no. 10, pp. 838–879, 2015.
- [40] E. B. Sandoval, J. Brandstetter, M. Obaid, and C. Bartneck, "Reciprocity in human-robot interaction: A quantitative approach through the prisoner's dilemma and the ultimatum game," *International Journal of Social Robotics*, vol. 8, no. 2, pp. 303–317, 2016.
- [41] M. A. Maggioni and D. Rossignoli, "If it looks like a human and speaks like a human ... communication and cooperation in strategic human-robot interactions," *SSRN Electronic Journal*, 2022.
- [42] I. Barreda-Tarrazona, A. Jaramillo-Gutiérrez, M. Pavan, and G. Sabater-Grande, "The 'human factor' in prisoner's dilemma cooperation," Working Papers 10, Economics Department, Universitat Jaume I, Castellón (Spain), 2021.
- [43] T.-Y. Hsieh, B. Chaudhury, and E. S. Cross, "Human-robot cooperation in economic games: People show strong reciprocity but conditional prosociality toward robots," Available at SSRN: <https://ssrn.com/abstract=3917169> or <http://dx.doi.org/10.2139/ssrn.3917169>, 2020.
- [44] J. E. Plaks, L. Bustos Rodriguez, and R. Ayad, "Identifying psychological features of robots that encourage and discourage trust," *Computers in Human Behavior*, vol. 134, p. 107301, 2022.
- [45] J. Duffy and H. Xie, "Group size and cooperation among strangers," *Journal of Economic Behavior and Organization*, vol. 126, pp. 55–74, 2016.
- [46] J. Wu, E. Paeng, K. Linder, P. Valdesolo, and J. C. Boerkoel, "Trust and cooperation in human-robot decision making," in *AAAI Fall Symposium - Technical Report*, vol. FS-16-01, pp. 110–116, 2016.
- [47] J. Jahng, J. D. Kralik, D. U. Hwang, and J. Jeong, "Neural dynamics of two players when using nonverbal cues to gauge intentions to cooperate during the prisoner's dilemma game," *NeuroImage*, vol. 157, pp. 263–274, 2017.
- [48] W. Mou, M. Ruocco, D. Zanatto, and A. Cangelosi, "When would you trust a robot? A study on trust and theory of mind in human-robot interactions," in *29th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2020*, pp. 956–962, 2020.
- [49] J. Schug, D. Matsumoto, Y. Horita, T. Yamagishi, and K. Bonnet, "Emotional expressivity as a signal of cooperation," *Evolution and Human Behavior*, vol. 31, no. 2, pp. 87–94, 2010.
- [50] E. Krumhuber, A. S. Manstead, D. Cosker, D. Marshall, P. L. Rosin, and A. Kappas, "Facial dynamics as indicators of trustworthiness and cooperative behavior," *Emotion*, vol. 7, no. 4, pp. 730–735, 2007.
- [51] W. M. Brown and C. Moore, "Smile asymmetries and reputation as reliable indicators of likelihood to cooperate: An evolutionary analysis," *Advances in Psychology Research*, vol. 11, no. 3, pp. 59–78, 2002.
- [52] L. I. Reed, K. N. Zeglen, and K. L. Schmidt, "Facial expressions as honest signals of cooperative intent in a one-shot anonymous prisoner's dilemma game," *Evolution and Human Behavior*, vol. 33, no. 3, pp. 200–209, 2012.
- [53] C. Straßmann, A. R. Von Der Pütten, R. Yaghoubzadeh, R. Kaminski, and N. Krämer, "The effect of an intelligent virtual agent's nonverbal behavior with regard to dominance and cooperativity," in *Lecture Notes in Computer Science*, vol. 10011, pp. 15–28, 2016.
- [54] S. Zörner, E. Arts, B. Vasiljevic, A. Srivastava, F. Schmalzl, G. Mir, K. Bhatia, E. Strahl, A. Peters, T. Alpay, and S. Wermter, "An immersive investment game to study human-robot trust," *Frontiers in Robotics and AI*, vol. 8, pp. 1–16, June 2021.
- [55] I. Torre and L. White, "Trust in vocal human–robot interaction: Implications for robot voice design," pp. 299–316, 2021.
- [56] J. Schug, D. Matsumoto, Y. Horita, T. Yamagishi, and K. Bonnet, "Emotional expressivity as a signal of cooperation," *Evolution and Human Behavior*, vol. 31, no. 2, pp. 87–94, 2010.
- [57] I. Torre, J. Goslin, and L. White, "If your device could smile: People trust happy-sounding artificial agents more," *Computers in Human Behavior*, vol. 105, p. 106215, April 2020.
- [58] I. Torre, J. Goslin, L. White, and D. Zanatto, "Trust in artificial voices: A 'congruency effect' of first impressions and behavioural experience," in *ACM International Conference Proceeding Series*, April 2018.
- [59] J. H. Manson, G. A. Bryant, M. M. Gervais, and M. A. Kline, "Convergence of speech rate in conversation predicts cooperation," *Evolution and Human Behavior*, vol. 34, no. 6, pp. 419–426, 2013.
- [60] C. Straßmann, A. M. Rosenthal-Von Der Pütten, and N. C. Krämer, "With or against each other? The influence of a virtual agent's (non)cooperative behavior on user's cooperation behavior in the prisoners' dilemma," *Advances in Human-Computer Interaction*, vol. 2018, 2018.
- [61] T. Y. Hsieh and E. S. Cross, "People's dispositional cooperative tendencies towards robots are unaffected by robots' negative emotional displays in prisoner's dilemma games," *Cognition and Emotion*, vol. 36, no. 5, pp. 995–1019, 2022.
- [62] A. Klapwijk and P. A. Van Lange, "Promoting cooperation and trust in 'noisy' situations: The power of generosity," *Journal of Personality and Social Psychology*, vol. 96, no. 1, pp. 83–103, 2009.
- [63] J. M. Weber and J. K. Murnighan, "Suckers or saviors? Consistent contributors in social dilemmas," *Journal of Personality and Social Psychology*, vol. 95, no. 6, pp. 1340–1353, 2008.
- [64] C. Esterwood and L. P. Robert, "Do you still trust me? Human-robot trust repair strategies," in *2021 30th IEEE International Conference on Robot and Human Interactive Communication, RO-MAN 2021*, pp. 183–188, Institute of Electrical and Electronics Engineers Inc., 2021.
- [65] L. Cominelli, F. Feri, R. Garofalo, C. Giannetti, M. A. Meléndez-Jiménez, A. Greco, M. Nardelli, E. P. Scilingo, and O. Kirchkamp, "Promises and trust in human–robot interaction," *Scientific Reports*, vol. 11, no. 1, 2021.
- [66] M. Kerzel, E. Strahl, S. Magg, N. Navarro-Guerrero, S. Heinrich, and S. Wermter, "NICO-Neuro-inspired companion: A developmental humanoid robot platform for multimodal interaction," in *RO-MAN 2017 - 26th IEEE International Symposium on Robot and Human Interactive Communication*, pp. 113–120, January 2017.
- [67] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, A. Y. Ng, et al., "ROS: An open-source robot operating system," in *ICRA workshop on open source software*, vol. 3, p. 5, Kobe, Japan, 2009.
- [68] C. Tsiourti, A. Weiss, K. Wac, and M. Vincze, "Multimodal integration of emotional signals from voice, body, and context: Effects of (in)congruence on emotion recognition and attitudes towards robots," *International Journal of Social Robotics*, vol. 11, no. 4, pp. 555–573, 2019.
- [69] Y. Jia, Y. Zhang, R. J. Weiss, Q. Wang, J. Shen, F. Ren, Z. Chen, P. Nguyen, R. Pang, I. L. Moreno, and Y. Wu, "Transfer learning from speaker verification to multispeaker text-to-speech synthesis," in *Advances in Neural Information Processing Systems*, no. NeurIPS, pp. 4480–4490, December 2018.
- [70] P. R. De Silva and N. Bianchi-Berthouze, "Modeling human affective postures: An information theoretic characterization of posture features," in *Computer Animation and Virtual Worlds*, vol. 15, pp. 269–276, 2004.
- [71] P. Ekman and D. Keltner, "Universal facial expressions of emotion," *Segerstrale U. P. Molnar P. eds. Nonverbal communication: Where nature meets culture*, vol. 27, p. 46, 1997.
- [72] C. Bartneck, D. Kulić, E. Croft, and S. Zoghbi, "Measurement instruments for the anthropomorphism, animacy, likeability, perceived intelligence, and perceived safety of robots," *International Journal of Social Robotics*, vol. 1, no. 1, pp. 71–81, 2009.
- [73] C. Harmon-Jones, B. Bastian, and E. Harmon-Jones, "The discrete emotions questionnaire: A new tool for measuring state self-reported emotions," *PLoS ONE*, vol. 11, no. 8, pp. 1–25, 2016.
- [74] Student, "The probable error of a mean," *Biometrika*, pp. 1–25, 1908.
- [75] P. E. McKnight and J. Najab, "Mann-whitney u test," *The Corsini encyclopedia of psychology*, pp. 1–1, 2010.
- [76] R. A. Fisher, "Statistical methods for research workers," in *Breakthroughs in statistics*, pp. 66–70, Springer, 1992.
- [77] *Spearman rank correlation coefficient*, pp. 502–505. New York, NY: Springer New York, 2008.