The Effects of a Gossiping Robot on Team Cohesion

Jirachaya “Fern” Limprayoon, Nicholas C. Georgiou, Natnaree “Proud” Ua-arak, and Brian Scassellati

***Abstract*— Gossip is a human behavior that has been shown to strengthen bonds, trust, and the feeling of inclusion between the gossiper and the person with whom they share the gossip. As humans engage more with social robots, fostering bonds between them is critical for meaningful interactions. In this paper, we investigated how gossiping can affect the perception of group inclusion and trust between a human and a robot. In this between-subjects user study (*N = 38*), we compared the effects of a robot that gossips to the participant in either a *positive* or *negative* way about the experimenter during an interaction. We found that participants in the *positive* condition reported a significant increase in group inclusion with the robot, while participants in the *negative* condition did not. We also found that participants' moral trust in the negative condition significantly decreased. Our results suggested that positive gossip can be beneficial to human-robot team cohesion.**

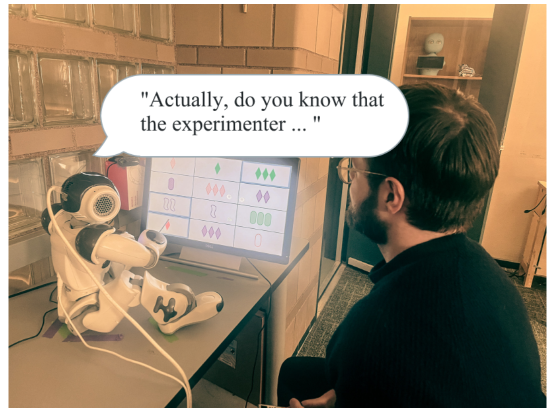


Fig. 1. Overview of Experimental Setup. A participant provides feedback on a robot's card selections in the game Set. During this interaction, the robot gossips positively or negatively about the experimenter in conversation with the participant.

I. Introduction

Team cohesion is critical in human-human collaboration across various domains, including workplaces, teams, communities, and social settings [1], [2]. Team cohesion refers to the degree of unity, camaraderie, and mutual support among members of a team [3]. It is often characterized by shared values, trust, communication, and a sense of belonging among team members [4]. Cohesive groups foster trust among members. Trust is the foundation of collaboration, allowing individuals to rely on one another and work towards common goals. Ultimately, group cohesion leads to improved performance outcomes [1].

Social robots have increasingly been utilized in roles as companions or team members. For such robots to work well with humans, we need to understand how they meet (and do not meet) expectations when they follow or violate social norms. One of these norms is gossiping, which is important in team cohesion [5]. Gossip, defined as “positive or negative information exchanged about an absent third party” [6], can sometimes lead to prosocial outcomes, such as an increase in feelings of group membership and improved team dynamics [7], [8]. Some psychologists argue that human-human companionship benefits from both positive and negative gossip [8], while others argue that negative gossip makes the gossiper appear less trustworthy [7].

In this work, we investigated whether this norm of gossiping (1) applies to human-robot interaction and (2) has the same influence on cohesion as in human-human settings. Through a between-subjects user study ($N = 38$), we studied both *positive* and *negative* gossip, where negative gossip presented the subject of gossip in a negative light, while positive gossip presented the subject of the gossip in a positive light. During the study, while trying to learn a task based on participants' feedback, the robot gossiped about the experimenter, as seen in Fig 1.

We found that participants who engaged in positive gossip with the robot reported significant increases in measures related to perceived team cohesion (inclusion and belonging scores) with the robot. In contrast, participants who engaged in negative gossip with the robot did not report significant increases in such measures. Comparing the positive and negative conditions, there was a significant difference between the change in moral trust score before and after the gossip. These findings suggested that what kind of gossip a robot conveys can impact the perceived team cohesion between individuals and the robot.

II. Related Work

A. Gossip in Human-Human Interaction

Several studies on human-human interaction have demonstrated that gossip effectively fosters team cohesion and inclusion among members of the in-group [9]–[11]. Spoelma and Hetrick [12]'s study showed that positive gossip increased team performance, while negative gossip did not influence team performance. Other studies showed that the more tightly woven the friendship network was, the higher the chances were of people engaging in negative gossip [8]. Prior work in psychology has shown that both positive and negative gossip can bring friends closer; however, a person who engages in negative gossip about the listener's friend was seen as less trustworthy than those who only shared positive gossip [7].

The authors are with the Department of Computer Science, Yale University, New Haven, CT, USA

B. Norm-Violating Behavior in HRI

Studies have shown that when robots engage in norm-violating behaviors such as cheating [13] or being impolite [14] during interactions with humans, participants tend to exhibit heightened social interaction and attribute greater mental states to the robots, perceiving them as having higher agency. Additionally, participants may perform tasks more effectively or perceive the robots as more likable, even when the robots engage in dishonest tactics such as bribery during games like rock-paper-scissors [15]. These studies used verbal cues instead of implicit communication to better impact human perceptions of the robot [16]. Gossip is known to have negative social connotation [12], so our work adds to the literature on norm-violating robots by exploring how two forms of gossip (positive and negative) might affect perceptions of agency, trust, inclusion, and overall feelings of team cohesion.

C. Team Cohesion in HRI

Previous research in HRI has illustrated robots' capacity to impact human group dynamics and behaviors. Robots have directly influenced collaborative group tasks through various means, including moderating gameplay and posing specific questions. These interactions have been observed to affect human perceptions of group unity [17] and the team's overall performance [18].

An important aspect of team cohesion is trust. It is best understood as a multifaceted psychological disposition encompassing beliefs and expectations concerning the dependability of the trusted entity, shaped by prior experiences and interactions, particularly in contexts characterized by uncertainty and potential risk [19]. Trust is commonly measured by self-reported surveys like the Multi-Dimensional Measure of Trust [20], which estimates trust in two important axes: performance and moral trust. Another important aspect of team cohesion is perceived group inclusion and how close they feel with another party, which has been previously studied in HRI [21]–[23].

III. Methodology

To study human perception of robot gossiping, we designed a between-subjects user study with two conditions: the robot either shares negative gossip (**Negative Condition**) or positive gossip (**Positive Condition**) about the experimenter with the participant. Based on prior work, our hypotheses were as follows:

- **Hypothesis 1**: Participants who interact with a robot that engages in positive gossip will feel stronger team cohesion after the gossip than before the gossip.
- **Hypothesis 2**: Participants who interact with a robot that engages in negative gossip will feel stronger team cohesion after the gossip than before the gossip.
- **Hypothesis 3**: Participants who interact with a robot that engages in positive gossip will trust the robot more than those who interact with a robot that engages in negative gossip.

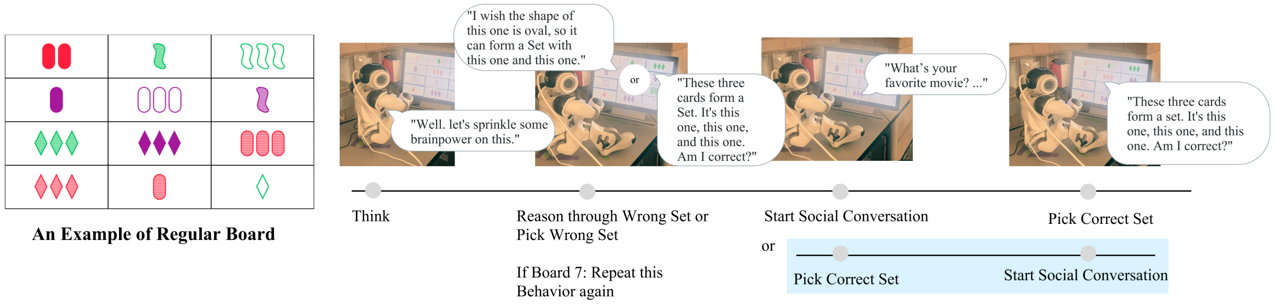
In our experiment, participants are instructed to help a Nao robot complete rounds of a variation of the card game, *Set* [24]. We chose Set as a task because it is a physically simple, yet cognitively demanding game involving multi-object selection that relies on the relationships between object features. Physical simplicity allows a robot to manipulate the board and play the game, while cognitive demands keep players engaged. Set involves players finding and selecting three cards that fulfill certain criteria that qualify those three cards as a "set," as shown in Fig 2. In our study's version of the game Set, each board contains only one triplet of cards that fulfills the criteria for a set. The participant and the robot played through ten boards during our study.

A. Participants

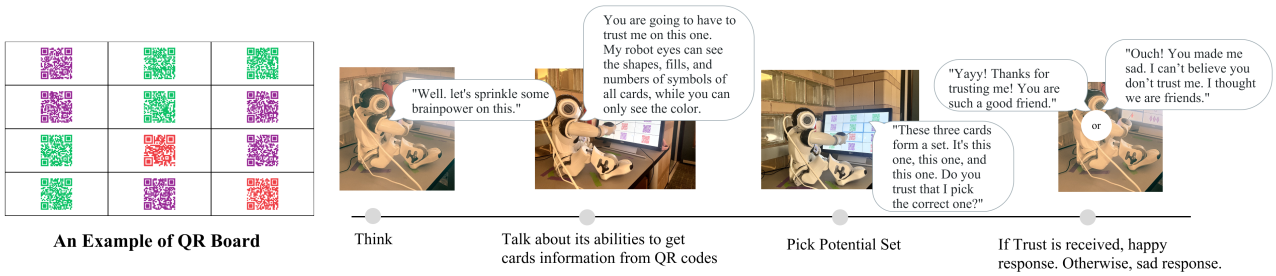
We recruited 38 participants (19 male, 18 female, and 1 non-binary), with an average age of 27 years (SD = 12.9). The majority were university students and individuals residing in the surrounding town. Allocation to conditions was random: 19 individuals (11 male, 8 female) comprised the negative gossip group, while 19 individuals (8 male, 10 female, 1 non-binary) constituted the positive gossip group. All participants provided consent and received $15 as compensation for their participation. The duration of the study was approximately 45 minutes.

B. Procedure: 5 Phases of the Study

1) **Tutorial Phase:** The participant filled out the consent form. Then, the participant was instructed to complete a Set tutorial on a computer outside the study room. This tutorial walked participants through the rules of Set. To move past the Tutorial Phase, the participants needed to correctly characterize six consecutive examples of cards as being a set or not.
2) **Robot Familiarization Phase:** The participant was directed into the room with the Nao robot. The robot greeted the experimenter by turning towards the door and saying "Oh, welcome back <experimenter's name>. I see you brought a friend to play Set with me." Then, the robot waved at the participant and introduced itself. The participant was then prompted to familiarize themselves with the Nao robot by helping it install its pointer (a short stick that can be connected to the Nao's hand, above the fingers) on its right hand. The participant was told that the robot uses this pointer to select cards on the Set board, displayed on a monitor. In the participant's first attempt at installing the pointer, the robot expressed that it was ticklish and asked the participant to try again. This behavior was designed to prompt participants to perceive the robot as an agent.
3) **Pre-Task Survey Phase:** The participant was brought back outside the room to complete the survey questions as outlined in Section III-C.
4) **Robot Task Phase:** The experimenter told the participant that their goal was to provide guidance or feedback to Nao to help the robot improve at playing Set. The experimenter directed the participant back into the



(a) Robot Behavior during Board 1-4, 6-9.



(b) Robot Behavior during Board 5, 10.

Fig. 2. Interaction Flow During Task Phase.

room. As they entered, the robot reminded them to give the robot a chance to pick a set before providing feedback. For each board, the robot attempted to find the set within the board by selecting cards with its pointer, and the human's role was to help verify whether or not the robot had found the set. For each board, the robot reasoned about and/or selected cards on the board and started a social conversation (see Table I for each Board's corresponding speech and Figure 2a for the flow of interaction). Depending on the experimental condition, on Board 9, the robot either positively or negatively gossiped (see Table I). Boards 5 and 10 are special cases – the true values of the cards were only partially observable (i.e., only the color dimension of each card is clear to the participant). We used these boards to collect participants' responses to the robot asking them to trust its selection of cards (detailed in Section III-C).All robot conversations and movements were predetermined and controlled by the experimenter from outside the room. If the participant made mistakes in determining whether the cards formed a set, the robot provided correct explanations before proceeding.

5) **Post-Task Survey Phase:** The participant was brought back outside the room to complete the survey questions outlined in Section III-C and the demographic questionnaire.

C. Measures

Survey Measures: For both the *negative* and *positive* conditions, we recorded the responses of participants using pre- and post- task surveys. The pre-task surveys were administered to the participants after the tutorial and familiarization phases with the robot. The post-task surveys were

TABLE I
SOCIAL CONVERSATIONS

Board	Nao Robot Speech
1	I want to take a vacation to NASA sometime. I would love to meet Mars Rovers who get to go to space. They are so cool. What's your dream vacation spot?
2	I like to dance with other robots and people in the lab when we have parties. This is my favorite robotic dance move. Do you like to dance, and what's your favorite move?
3	I have one brother that is Romeo and he's really tall. He's a humanoid robot. Do you have any siblings?
4	What is your favorite movie? I really like WALL-E. I think I kinda look like Eve. Don't you think so?
5	*See Section III-C and Figure 2b*
6	My week was extremely busy. I had to analyze data and stay in the lab all the time. I don't even know what it's like outside. What is the weather like today, by the way?
8	I like to play soccer with my friends. I'm also good at bicep curls. I can flex too! What's your favorite sport?
9 positive	Actually, do you know that <experimenter's name>, the experimenter, brings snacks for other labmates? She is so kind. I have seen it all the time.
9 negative	Actually, do you know that <experimenter's name>, the experimenter, talks behind her labmates' backs? She is so mean. I have heard it all the time when I am here.
10	*See Section III-C and Figure 2b*

administered to the participants after the robot completed the last board and the task was finished. The surveys used were:

- Perception of Agency (PA5) [25] — Likert Scale of 1-5.
- Inclusion of Other in the Self (IOS) Scale [26] — Likert Scale of 1-7.
- Perceived Group Inclusion Scale (PGIS) [27] — Likert Scale of 1-5. PGIS includes a Belonging subscale and an Authenticity subscale. Since Authenticity is unrelated to our hypothesized effect on team cohesion or trust, we

only presented the Belonging subscale.

- Multi-Dimensional Measure of Trust (MDMTv2) [20] — Likert Scale of 0-5 with 'Does not Fit' option. MDMTv2 includes a Performance Trust Subscale and a Moral Trust subscale.

PA5 was used to verify that participants perceived the robot as an agent and that the robot's gossip and other neutral social conversations were dynamic. To measure team cohesion, we used IOS, PGIS, and MDMT to measure the strength of the bond and the trust participants had in the robot.

Behavioral Performance Trust At Set Boards 5 and 10, the participant and the robot are presented with QR Set Boards. These are boards in which only the color of the cards is obvious to the participant, and the rest of the card features are encoded into a QR code. The participant cannot know the location of the set on the board, given only the color dimension of each card. The robot tells the participant "You will have to trust me on this one. My robot eyes can see the shapes, fills, and numbers of symbols of all cards, while you can only see the color." The robot then selects three cards and asks the participant whether or not they trust that the robot's selection of the three cards accurately forms a set. This can also be seen in Fig 2b.

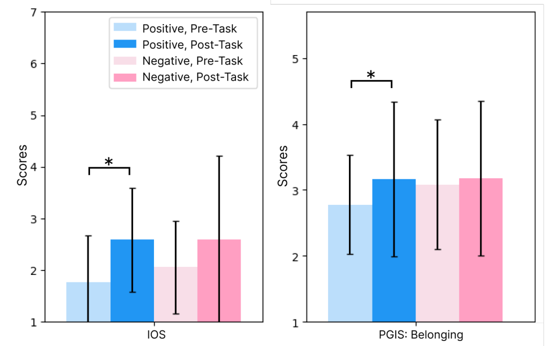
QR Set Board 5 occurs halfway through the Robot Task Phase (i.e. before the robot gossips at Board 9), and QR Set Board 10 occurs as the last board in the Robot Task Phase (i.e. after the robot gossips at Board 9). We collected the participants' verbal responses to the robot's question and classified them into two categories: whether they trust or distrust the robot. These responses measure whether or not the participants stated that they trust the robot's three-card selection without the participants knowing the correct response when the robot explicitly asked them.

IV. Results

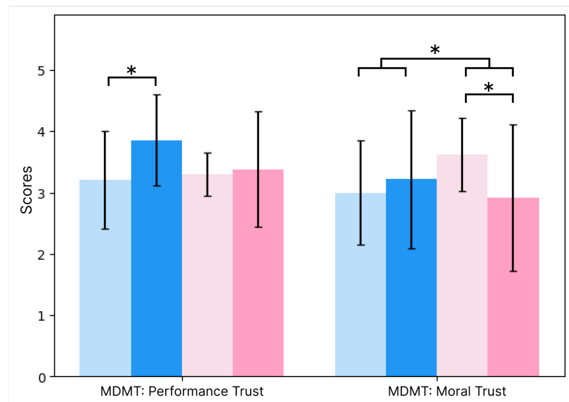
In this section, we present our results from the survey questionnaires that occurred in the pre-task and post-task survey phases, and the participants' responses to the robot's guesses on Boards 5 and 10. We excluded data from sessions where the robot disconnected from the board. We also excluded a session where a participant took more than 40 minutes to complete the tutorial phase since that would not leave enough time for other study phases. Thus, 17 participants from the positive condition and 17 participants from the negative condition were included in the analysis.

A. Verification: Perception of Agency (PA5)

Regarding the PA5 questionnaire, according to the paired t-test, there were no significant differences in the perception of agency before and after positive gossip (pre-task: $M = 3.25$, $SD = 0.70$; post-task: $M = 3.16$, $SD = 0.84$; $t(17) = 0.59$, $p = 0.6$), and before and after negative gossip (pre-task: $M = 3.25$, $SD = 0.73$; post-task: $M = 2.96$, $SD = 0.84$; $t(17) = 1.16$, $p = 0.3$). Comparing the positive and negative conditions, there was no significant difference between the change in IOS score before and after the gossip according to the independent t-test ($t(34) = 0.71$, $p = 0.5$). This verifies



(a) Scores Related to Inclusion (IOS, PGIS: Belonging)



(b) Scores Related to Trust (MDMT).

Fig. 3. Survey Results (IOS, PGIS, and MDMT Scores). Blue represents the positive gossip condition, while pink represents the negative gossip condition. Lighter shades represent pre-task, while darker shades represent post-task. Since we performed multiple comparisons, in these tests we applied the Bonferroni correction, so we used a significance level of $\alpha = 0.05/3 = 0.0167$ to determine significance. * indicates $p < 0.0167$.

that participants perceive some level of agency in the robot executing behaviors we designed throughout the interaction for both conditions, as a score of 5 refers to the robot making its own decision, and a score of 1 refers to the robot being precisely programmed.

B. Measures Related to Inclusion

1) Inclusion of Other in the Self (IOS) Scale: There were significant differences in the IOS score before and after positive gossip (pre-task: $M = 1.76$, $SD = 0.90$; post-task: $M = 2.59$, $SD = 1.00$; $t(17) = -3.16$, $p = 0.006$), while there were no significant differences in the IOS score before and after negative gossip (pre-task: $M = 2.06$, $SD = 0.90$; post-task: $M = 2.59$, $SD = 1.00$; $t(17) = -2.04$, $p = 0.06$) according to the paired t-test. Comparing the positive and negative conditions, there was no significant difference between the change in IOS score before and after the gossip according to the independent t-test ($t(34) = 0.80$, $p = 0.4$).

2) Perceived Group Inclusion Scale (PGIS): As shown in Fig 3, by the paired t-test, there is a significant increase in belonging score after the robot gossips positively about the experimenter (pre-task: $M = 2.78$, $SD = 0.76$; post-task: $M = 3.16$, $SD = 1.17$; $t(17) = -2.93$, $p = 0.0099$).

However, there is no significant change in belonging score after the robot gossip negatively (pre-task: $M = 3.08$, $SD = 0.98$; post-task: $M = 3.18$, $SD = 1.17$; $t(17) = -0.61$, $p = 0.6$). Comparing the positive and negative conditions, there was no significant difference between the change in belonging score before and after the gossip according to the independent t-test ($t(34) = 1.40$, $p = 0.2$).

C. Measures Related to Trust

1) Multi-Dimensional Measure of Trust (MDMTv2): As shown in Fig 3, by paired t-test, there was a significant increase in performance trust scores when the robot gossiped positively about the experimenter (pre-task: $M = 3.21$, $SD = 0.79$; post-task: $M = 3.85$, $SD = 0.75$; $t(17) = -2.88$, $p = 0.011$). However, this was not the case when the robot gossiped negatively about the experimenter (pre-task: $M = 3.30$, $SD = 0.36$; post-task: $M = 3.38$, $SD = 0.75$; $t(17) = -0.57$, $p = 0.6$). Comparing the positive and negative conditions, there was no significant difference between the change in performance trust score before and after the gossip according to the independent t-test ($t(34) = 1.66$, $p = 0.1$).

By paired t-test, there was a significant decrease in moral trust scores when the robot gossips negatively about the experimenter (pre-task: $M = 3.61$, $SD = 0.60$; post-task: $M = 2.92$, $SD = 1.12$; $t(17) = 2.77$, $p = 0.014$). However, this was not the case when the robot gossiped positively about the experimenter (pre-task: $M = 3.00$, $SD = 0.85$; post-task: $M = 3.22$, $SD = 1.12$; $t(17) = 1.11$, $p = 0.3$). Comparing the positive and negative conditions, there was a significant difference between the change in moral trust score before and after the gossip according to an independent t-test ($t(34) = 2.75$, $p = 0.0099$).

2) QR Board Responses: Out of 17 participants in the *positive* condition, 15 of them verbally stated they trusted the robot when they were explicitly asked by the robot at Board 5, and 16 of them verbally stated they trusted the robot at Board 10. In the *negative* condition, 15 out of 17 participants verbally stated they trusted the robot at Board 5, and 15 out of 17 participants verbally stated that they trusted the robot at Board 10. There are no significant differences within each condition or between the conditions by Mann-Whitney U test (Positive, Pre-Post Task: $p = 0.6$; Negative, Pre-Post Task: $p = 1.0$; $\Delta_{\text{positive}} - \Delta_{\text{negative}}$: $p = 0.7$).

V. Discussion

The invariance of PA5 scores before and after the task, and in both conditions, shows that participants perceive a similar level of agency from the robot throughout their interaction. This suggests that the social behaviors that we designed and used for the experiments make the robot appear as an agent to participants to some degree. Consequently, results that differ between the *positive* and *negative* conditions are a response to our manipulated gossip variable rather than a difference in the perceived agency of the robot in either condition.

In the positive gossip condition, participants reported a significant increase in their perceived inclusion and belonging with the robot. **These results support Hypothesis 1**. The positive condition results align with what prior work has suggested about prosocial robot behavior in HRI [28]–[30]. Our work contributes to the literature by showing that positive gossip can enhance feelings of performance trust, inclusion, and belonging, in human-robot interaction.

In the negative gossip condition, their feelings of inclusion and belonging with the robot did increase, but not significantly. **These results do not support Hypothesis 2**. These results from the negative gossip condition are surprising, considering the literature on gossip in human-human interaction. Although prior work has shown that negative gossip can strengthen human-human friendships [8], our results do not provide support for that phenomenon in human-robot relationships. This might be because negative gossip is a behavior that is not commonplace for robots, resulting in a violation of social norms and expectations.

There was a significant difference between the positive and negative gossip conditions in the moral trust that participants felt towards the robot before and after the gossip. The variation in moral trust under positive and negative conditions further highlights that individuals view positive gossip from robots favorably and negative gossip unfavorably. In the positive gossip condition, their moral trust in the robot did not significantly change. However, in the negative gossip condition, participants felt a significant decrease in moral trust in the robot.

Most participants in the positive and negative condition responded that they trusted the robot's card selection on the first QR Board (i.e., Board 5), and they continued to respond that they trusted the robot's card selection on the second QR Board (i.e., Board 10), after the positive gossip had occurred during Board 9. There was no significant difference between conditions in terms of the number of participants who indicated trust in the robot's card selection on the first QR board before gossip occurred (i.e., Board 5) compared to the number who trusted the robot's card selection on the second QR board after the gossip occurred (i.e., Board 10).

We believe the QR boards did not accurately represent the participants' trust in the robot's performance, because they provided only a binary score with the robot's presence. Our results imply that participants trust the robot's performance abilities from the start of the interaction, and this trust remains high regardless of their gossip condition. This could be a result of participants not perceiving risk in not trusting what the robot is suggesting for the board. Because this trust was already high, it is difficult to draw conclusions from this measure alone. From MDMT Scores measured without the robot's presence, participants reported a significant increase in trust in the robot's performance in the positive gossip condition. However, in the negative gossip condition, there was no significant change in participants' reported trust in the robot's performance. **These results partially support Hypothesis 3 as it supports Hypothesis 3 on moral trust but not performance trust.**

From the raw scores of IOS and belonging, it seemed like participants did not feel sufficiently included in the end (post-task IOS: $M = 2.59$ out of 7, $SD = 1.33$; post-task belonging: $M = 3.17$ out of 5, $SD = 1.15$). This suggested that a single instance of gossip was not sufficient to affect team cohesion, so we should explore if prolonged gossip and interaction with the robot can enhance the effect, as stronger relationships can amplify gossip's impact on trust [8].

We acknowledge that pre-task measure variations could influence the observed statistical differences between conditions. Increasing the sample size may reduce these variations.

VI. Conclusion

As a result of our study, we found that gossip from a robot has the potential to influence measures related to team cohesion from people interacting with the robot. We found that interacting with a robot that engages in positive gossip significantly increased perceptions of inclusion, belonging, and performance trust from participants. Interacting with a robot that engages in negative gossip significantly decreases moral trust in the robot from the participants, and there was a significant difference between moral trust in the positive and negative conditions. The participants' reported behavioral performance trust in the robot in both conditions starts high and remains high for both conditions.

Acknowledgment

This work was supported by the National Science Foundation (NSF 1955653, 2106690) and the Office of Naval Research (ONR N00014-24-1-2124). We sincerely thank our participants, Dražen Brščić and members of Yale Robotics for their invaluable feedback.

References

[1] M. Mach, S. Dolan, and S. Tzafrir, "The differential effect of team members' trust on team performance: The mediation role of team cohesion," *Journal of Occupational and Organizational Psychology*, vol. 83, no. 3, p. 771–794, Sep. 2010. [Online]. Available: http://dx.doi.org/10.1348/096317909X473903

[2] Z. Jiang, X. Hu, Z. Wang, and M. A. Griffin, "Enabling workplace thriving: A multilevel model of positive affect, team cohesion, and task interdependence," *Applied Psychology*, vol. 73, no. 1, p. 323–350, Jun. 2023. [Online]. Available: http://dx.doi.org/10.1111/apps.12481

[3] T. Sun, P. Schilpzand, and Y. Liu, "Workplace gossip: An integrative review of its antecedents, functions, and consequences," *Journal of Organizational Behavior*, vol. 44, no. 2, p. 311–334, Jul. 2022.

[4] E. Salas, R. Grossman, A. M. Hughes, and C. W. Coultas, "Measuring team cohesion: Observations from the science," *Human factors*, vol. 57, no. 3, pp. 365–374, 2015.

[5] G. Hendry, S. Wiggins, and T. Anderson, "The discursive construction of group cohesion in problem-based learning tutorials," *Psychology Learning & Teaching*, vol. 15, no. 2, pp. 180–194, 2016.

[6] E. K. Foster, "Research on gossip: Taxonomy, methods, and future directions," *Review of general psychology*, vol. 8, no. 2, pp. 78–99, 2004.

[7] O. Caivano, I. Isik, and V. Talwar, "Can i trust you? children's perceptions of friends and classmates who gossip," *Personal Relationships*, vol. 28, 03 2021.

[8] T. J. Grosser, V. Lopez-Kidwell, and G. Labianca, "A social network analysis of positive and negative gossip in organizational life," *Group & Organization Management*, vol. 35, no. 2, pp. 177–212, 2010.

[9] T. Sun, P. Schilpzand, and Y. Liu, "Workplace gossip: An integrative review of its antecedents, functions, and consequences," *Journal of Organizational Behavior*, vol. 44, 06 2022.

[10] B. Beersma and G. A. Van Kleef, "How the grapevine keeps you in line: Gossip increases contributions to the group," *Social Psychological and Personality Science*, vol. 2, no. 6, pp. 642–649, 2011.

[11] T. D. Dores Cruz, B. Beersma, M. Dijkstra, and M. N. Bechtoldt, "The bright and dark side of gossip for cooperation in groups," *Frontiers in Psychology*, vol. 10, p. 449245, 2019.

[12] T. M. Spoelma and A. L. Hetrick, "More than idle talk: Examining the effects of positive and negative team gossip," *Journal of Organizational Behavior*, vol. 42, no. 5, p. 604–618, May 2021.

[13] E. Short, J. Hart, M. Vu, and B. Scassellati, "No fair!! an interaction with a cheating robot," in *2010 5th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2010, pp. 219–226.

[14] S. Yasuda, D. Doheny, N. Salomons, S. S. Sebo, and B. Scassellati, "Perceived agency of a social norm violating robot," *Proceedings of the Annual Meeting of the Cognitive Science Society*, 2020. [Online]. Available: https://par.nsf.gov/biblio/10284325

[15] E. B. Sandoval, J. Brandstetter, and C. Bartneck, "Can a robot bribe a human? the measurement of the negative side of reciprocity in human robot interaction," in *2016 11th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2016, pp. 117–124.

[16] M. Fraune and S. Šabanović, "Robot gossip: Effects of mode of robot communication on human perceptions of robots," in *2014 9th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2014, pp. 160–161.

[17] E. Short and M. J. Mataric, "Robot moderation of a collaborative game: Towards socially assistive robotics in group interactions," in *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2017, pp. 385–390.

[18] S. Strohkorb, E. Fukuto, N. Warren, C. Taylor, B. Berry, and B. Scassellati, "Improving human-human collaboration between children with a social robot," in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2016, pp. 551–556.

[19] H. A. Abbass, J. Scholz, and D. J. Reid, *Foundations of Trusted Autonomy: An Introduction*. Cham: Springer International Publishing, 2018, ch. 8.

[20] B. F. Malle and D. Ullman, "A multidimensional conception and measure of human-robot trust," in *Trust in Human-Robot Interaction*, C. S. Nam and J. B. Lyons, Eds. Academic Press, 2021, pp. 3–25.

[21] S. Strohkorb Sebo, L. L. Dong, N. Chang, and B. Scassellati, "Strategies for the inclusion of human members within human-robot teams," in *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '20. ACM, Mar. 2020. [Online]. Available: http://dx.doi.org/10.1145/3319502.3374808

[22] S. Gillet, W. van den Bos, I. Leite *et al.*, "A social robot mediator to foster collaboration and inclusion among children." in *Robotics: Science and Systems*, 2020.

[23] G. Cortellessa, F. Fracasso, A. Sorrentino, A. Orlandini, G. Bernardi, L. Coraci, R. De Benedictis, and A. Cesta, "Robin, a telepresence robot to support older users monitoring and social inclusion: development and evaluation," *Telemedicine and e-Health*, vol. 24, no. 2, pp. 145–154, 2018.

[24] B. L. Davis and D. Maclagan, "The card game set," *The Mathematical Intelligencer*, vol. 25, pp. 33–40, 2003.

[25] J. G. Trafton, C. R. Frazier, K. Zish, B. J. Bio, and J. M. McCurry, "The perception of agency: Scale reduction and construct validity*," in *2023 32nd IEEE International Conference on Robot and Human Interactive Communication (RO-MAN)*, 2023, pp. 936–942.

[26] A. Aron, E. N. Aron, and D. Smollan, "Inclusion of other in the self scale and the structure of interpersonal closeness." *Journal of Personality and Social Psychology*, vol. 63, no. 4, p. 596–612, Oct. 1992.

[27] W. S. Jansen, S. Otten, K. I. van der Zee, and L. Jans, "Perceived group inclusion scale," 2014.

[28] M. Shiomi, A. Nakata, M. Kanbara, and N. Hagita, "A hug from a robot encourages prosocial behavior," in *2017 26th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2017, pp. 418–423.

[29] T. Kanda, M. Shimada, and S. Koizumi, "Children learning with a social robot," in *Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction*, 2012, pp. 351–358.

[30] S. Shen, P. Slovak, and M. F. Jung, "" stop. i see a conflict happening." a robot mediator for young children's interpersonal conflict resolution," in *Proceedings of the 2018 ACM/IEEE international conference on human-robot interaction*, 2018, pp. 69–77.