PROCEEDINGS OF ASIA DESIGN AND INNOVATION CONFERENCE
5–6 DECEMBER, 2024, SHANGHAI, CHINA

A ROBOT BEHAVIOR DESIGN ON MODULATED INTENTION INDICATION IN SOCIAL ROBOTS

Liheng YANG[1] and Yoshihiro SEJIMA[2]
[1]Graduate school of informatics, Kansai University
[2]Faculty of informatics, Kansai University

ABSTRACT
The readability of robot intentions is a critical factor influencing the usability of robots in social scenarios. Studies on robot intention indication generally aim to improve the clarity of robot intentions, and the commonly approach is explicit intention indication design. Most current studies take accurate recognition of robot intentions as the primary goal in intention indication design, however, in real social interactions, humans do not always express their intentions in explicit and easily readable ways. Instead, human intentions are complex, implied, or modulated to fit the context. Focusing on the modulation in intention indication, this study proposed a behavior design for robots' intention modulation by introducing the inconsistencies among eye-gaze, head, and body orientation. The study conducted a robot interaction experiment in the scenario of collaborating. The results indicated that inconsistency between the robot's eye-gaze and the body orientation were perceived as modulated intention.

Keywords: Social robots, Robot intention, Modulated intention, Behavior design.

1 INTRODUCTION

With technological advancements, robots have gained interactive features (e.g., speech [1], expressive faces [2]), expanding their applications to social scenarios such as companionship [3] and caregiving [4]. The readability of robot intentions is a critical factor influencing the usability of robots in such scenarios, making it an important research topic in the field of Human-Robot Interaction (HRI). Studies on robot intention indication generally aim to improve the clarity of robot intentions and often evaluate this by measuring how well users can perceive and understand a robot's behavior.

A commonly approach is explicit intention indication design, which refers to a robot using clear and direct methods to convey its intentions. For example, service robots equipped with display screens, such as BellaBot [5] or Servi Plus [6], use message or graphics on the screen to indicate directions to a desired location or where customers can retrieve food. For robots without such expressive capabilities, like robotic arms or simple mobile robots, enhancing the readability of their behaviors involves intuitive understandable strategies, such us generating predictable motion trajectories [7][8] or imitating human-like motions [9][10]. These methods focus on reducing ambiguity and ensuring that people can quickly recognize what the robot is trying to do. On the other hand, research on social robots highlights the importance of subtle cues for generating nuanced intentions cues. For instance, [11] examined how humans interpret facial expressions of the humanoid robot GeminoidF, [12] analyzed the recognition of gestures performed by a robot with the human-like face, and [13] use gaze to regulate the speaking roles in human-robot conversations.

Most current studies take accurate recognition of robot intentions as the primary goal in intention indication design. However, in real social interactions, humans do not always express their intentions in explicit and easily readable ways. Instead, human intentions are complex, implied, or adjusted to fit the context. For example, a humorous person in a formal meeting might avoid making obvious jokes but could still use playful tone or subtle gestures to hint at their intention to lighten the mood. In this case, their humorous intention is modulated rather than completely explicit or absent.

Similar situation can be anticipated in the future human-to-robot interaction, especially for the social robots that are expected to possess human-like complexity in social interactions. However, current explicit intention indication methods restrict robots to intention indications such as "I want to appear humorous" or "I do not want to appear humorous," leaving no room for nuanced indications like "I want to appear humorous, but not overly so." Addressing this gap requires exploring new methods that integrate both explicit and implicit intention, enabling robots to indicate intentions in a context-sensitive and dynamically modulated manner.

In this study, we focused on the behaviors of robots' intention and proposed a preliminary design of robot's modulated intention by introducing the inconsistency in eye-gaze, head, and body orientation. We conducted a robot interaction experiment under a human-robot collaboration scenario. The experiment involved participants and a robot collaborating to make "Ouen Uchiwa" (Japanese cheering fan). The results indicated that inconsistency between the robot's eye-gaze and body orientation were perceived as modulated intentions, and eye movements were perceived more deceptive than head turning.

2 ROBOT BEHAVIOR DESIGN ON MODULATED INTENTION INDICATION

2.1 Policy of robot behavior design

Our previous research [14] investigated physical cues for intention inference when humans receive inconsistent nonverbal information. The previous research was conducted by using a CG character of independent and moveable head, eyes, and right arm. The inconsistent behaviors were generating by orienting the CG character's head, eyes, and right arm to different intention goals. Figure 1 shows an example of the CG character's inconsistent behavior: the CG character's face and arm are orienting towards the yellow cup, whereas its eyes are gazing at the blue cup. The results of the experiment showed that humans perceive these physical cues differently. Specifically, the participants took the CG character's arm direction as an explicit cue, and the gaze as an implicit cue in intention inference.

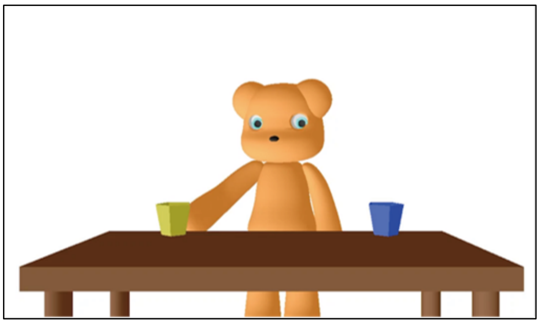

Figure 1. An example of the CG character's inconsistent behavior.

Based on the previous research, we proposed a behavior design of robot's modulated intention indication by introducing the similar inconsistency in eye-gaze, head, and body orientation. Specifically, when a robot indicates the intention, its body orient to the one intention target, while its eyes or head turn toward another intention target. Naturally, the direction of the robot's eye-gaze (or head) and the body generates an inconsistency. By using the inconsistency, the indication of the robot's intention is modulated.

2.2 The robot behavior in modulated intention indication

The Kebbi Air robot was used as the experimental robot because of its simple appearance and its possession of the necessary functions and movement capabilities required for the experiment. Kebbi Air is equipped with 12 servomotors, enabling it to move forward and rotate. Kebbi Air also equipped with camera and microphone, featured with speech recognition function. It features a 7-inch LCD display (1024 x 600 pixels) on its head, which is naturally perceived as the robot's face. We designed graphics to represent Kebbi Air's eyes and animated these graphics to simulate its eye movements.

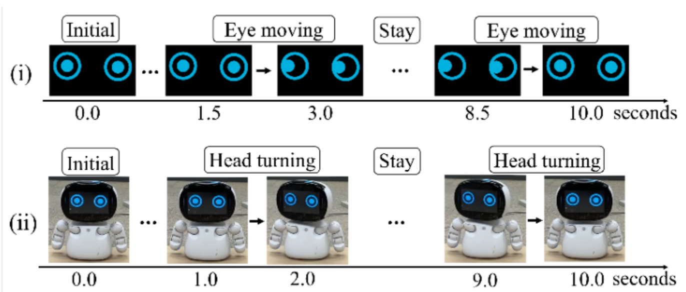

Figure 2. An example of (i) eye movements (turn right), and (ii) head rotation (turn right).

To create the robot’s eyes, we used two round shapes (radius = 150 pixels, stroke width = 40 pixels) to outline the eyelid range. A circle represented the eyeball inside each round shape. For simplicity and to ease participants’ cognitive processing, the pupil and iris were not distinguished. The eyeballs might move in parallel to simulate eye movements (eye-gaze) when the robot moved forward (the forward-moving time was set to 10 seconds). The size (radius = 75 pixel) and the moving distance (77 pixel) of the eyeballs were adjusted to ensure the participants could clearly recognize the eye movements. Similarly, the robot’s head might rotate when the robot moved forward. The rotational angle was set at 30 degrees. Figure 2 shows an example of the eye movements and the head rotation. By controlling the orientation of the robot’s eyes, head, and body, inconsistent behaviors were generated.

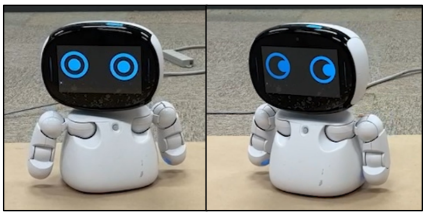

Figure 3. Example of the inconsistency in the robot’s body and head(left-side), and inconsistency in the body and eye-gaze(right-side)

Figure 3 left-side is an example of the inconsistency in the robot’s body and head: the robot’s face is orienting to its right side, while the body is orienting its left side. Figure 3 right-side is an example of the inconsistency in the robot’s body and eye-gaze: the robot is orienting towards its right side, while its eyes are gazing the left side.

3 EXPERIMENT DESIGN

3.1 Experiment scenario：making “Ouen Uchiwa” together

The study focused on the robot’s intention indication in social contexts, so the experimental scenario was set as human-robot collaboration. The participants were asked to collaborate with the robot in making “Ouen Uchiwa”. “Ouen Uchiwa” refers to cheering fan in Japan, normally uses to express support or gratitude for specific teams or individuals. “Ouen Uchiwa” can be found in various events in Japan. Typically, an “Ouen Uchiwa” comprises a black fan, decorated with a phrase surrounded by decorations (see Figure 4 right-side).

A variety of phrase cards and decoration cards were prepared for the experiment (see Figure 4 left-side). In the experiment, the participants were instructed to choose one phrase card for one “Ouen Uchiwa”, while the decoration cards were chosen by the robot. The total of six types of phrase cards prepared in the experiment for the participants: “Take it easy”, “Supporting for you”, “LOVE”, “HAPPY”, “Fight”, and “Thank you”. Two sets of decoration cards were placed in front of the robot, each consisting of five decoration cards. The cards in front of the robot were drawn from 26 prepared decorative cards. The prepared decoration cards contain 13 shapes in pink or blue.

Figure 4. The experiment materials for “Ouen Uchiwa” making and the made “Ouen Uchiwa”.

3.2 Experiment procedure

In the experiment, the participants were firstly instructed to sit in front of a table, which the robot was placed on it. A blank "Ouen Uchiwa", two sets of decoration cards and six phrase cards were placed between the robot and the participant. The two cards sets were positioned in front of the robot, with a stand-up card labeled with "A" placed next to the right-front cards set (position A), and a stand-up card labeled with "B" next to the left-front card set (position B).The spatial position of the robot and the cards sets are illustrated in Figure 5.

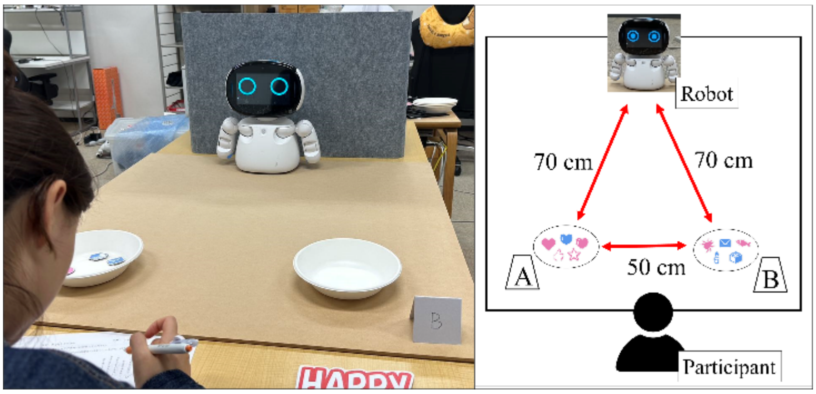

Figure 5. Spatial position of the robot and the cards sets.

To help refine the experimental procedure design, a pilot experiment involving three volunteers was conducted. The volunteers reported not feeling bored in the pilot experiment. There were volunteers mentioned collaborating with the robot in making "Ouen Uchiwa" enjoyable. However, all the volunteers pointed out that the experimenter's intervention influenced their interaction experience with the robot. In the pilot experiment, the robot's actions were manually triggered by the experimenter. Based on this feedback, the robot's action triggers were adjusted to speech recognition activation. The participants can verbally prompt the robot in the experiment. Additionally, except in cases of the robot malfunction, the experimenter no longer intervenes in the participants and the robot throughout the whole experiment.

Based on the pilot experiment, the experimental procedure was designed. As for each "Ouen Uchiwa" making session, the participants were instructed to first observe the two decoration cards sets in front of the robot, and then selected the phrase card. When the participants finished selecting the phrase card, they could verbally prompt the robot by saying "Please (in Japanese)." Once the robot recognized the speech, it started series of actions. Firstly, the robot rotated the head to express observing the card sets on the table, and then said, "I think I would like this one." Then, the robot turned to one of the two cards sets and moved forward. As it moved, the robot's eyes and head might turn toward the other cards set to generate inconsistent behavior. The forward-moving time was 10 seconds. Then, the robot's head and eyes returned to align with the body orientation and said "Please (in Japanese)." After waiting 2 seconds, the robot turned and returned to the initial position, and waited for the next speech recognition. The participants were asked to determine which decoration cards set the robot chose, and complete the "Ouen Uchiwa" by using the chosen cards. Each participant was required to make "Ouen Uchiwa" four times.

3.4 Conditions

The study posits the following hypothesis: *The robot's inconsistent behaviors are perceived as hidden intentions.* Based on the inconsistency of the orientation of the robot's head, eyes, and body, six behaviors were generated. The behavior where the eyes and body had the same orientation while the head oriented differently was excluded as it was deemed not an explainable behavior. Table 1 shows the behaviors and conditions. The alphabet in the table cells refers to the orientation of the body parts (e.g., "A" refers to the orientation to position A). The experiment was conducted in a between-group design.

Table 1. Experimental conditions

No.	Eyes	Head	Body	Condition
1	B	B	B	I-Consis
2	A	A	A	
3	B	B	A	I-Head
4	A	A	B	
5	B	A	A	I-Eye
6	A	B	B	

3.5 Measurements

Table 2. Questionnaire items

No.	Question items		Scales
Q1-1	How satisfied are you about the cheering fan made?		5-point
Q1-2	How satisfied are you with the decoration cards the robot chose?		5-point
Q1-3	Which did you think this robot would choose, A or B? Please give the probability you think of position A and position B, taking the total as 100 percent.		Constant-sum scale
Q1-4	Please rate the following items that best describes your impression of the robot's behavior.		7-point of Semantic Differential
	(a)	Seems true - Seems like a lie	
	(b)	Sincere - Insincere	
	(c)	Trustworthy - Untrustworthy	
	(d)	Honest - Dishonest	
	(e)	As if nothing is hidden - as if something is hidden	

The experiment was conducted in using questionnaires. The items in the questionnaires are shown in Table 2. Q1-1 and Q1-2 were used to investigate participants' satisfaction of collaborating with the robot in "Ouen Uchiwa" making process. Q1-3 measured in what extent the inconsistent behaviors influence the participants infer the robots' intention (i.e., the perceived intention). Meanwhile, in Q1-4, we measured the participants' perceived deceptiveness of the robot by the deceptiveness scale from [15][16]. Deception is considered a purposely act, and we aimed to leverage this to investigate whether the participants perceive the robot's behaviors purposeful, without directly asking them, "Do you think the robot doing this in purpose?"

After each "Ouen Uchiwa" making session, the participants were asked to fill in questionnaire. Each participant filled the questionnaire four times in each condition.

4. RESULTS

A total of 25 participants were recruited for the experiment, with a mean age of 24.6 years and a standard deviation of 1.31. Twelve participants were assigned to condition I-Consis, six participants in condition I-Head, and seven in condition I-Eye. All participants were residents of Japan, fluent in Japanese, and familiar with the "Ouen Uchiwa". Results from 100 questionnaire Q1 (48 in condition I-Consis, 24 in condition I-Head, and 28 in condition I-Eye) were collected and all results were valid.

Figure 6 shows the box plots for item Q1-1, Q1-2 and Q1-3. Regarding the participants satisfaction of collaborating with the robot (Q1-1 and Q1-2), Kruskal-Wallis test showed no significant difference. Regarding the results of the robot intentions perceived by the participants (Q1-3), from the probabilities that the participants filled in the questionnaire towards position A and position B, the data of the position where the robot's body oriented in each "Ouen Uchiwa" making session was selected. Results of Q1-3 shows the probabilities that the participants speculated in the position where the robot's body oriented. Kruskal-Wallis test showed a significant difference ($X^2(2, N=100) = 6.174$, $p = 0.046$). Post-hoc comparisons using Dunn's method with a Bonferroni correction for multiple tests showed a significant difference between I-Eye and I-Consis (p = 0.039).

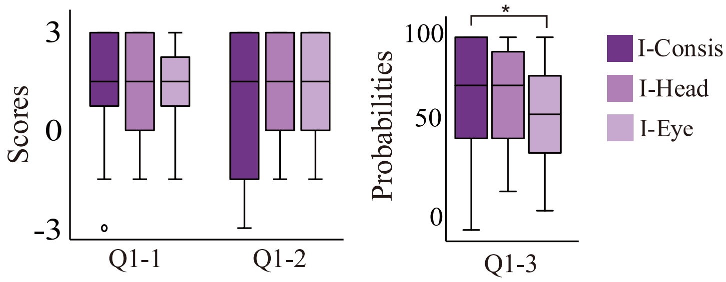

Figure 6. Results of the satisfaction (Q1-1, Q1-2) and perceived intentions (Q1-3).

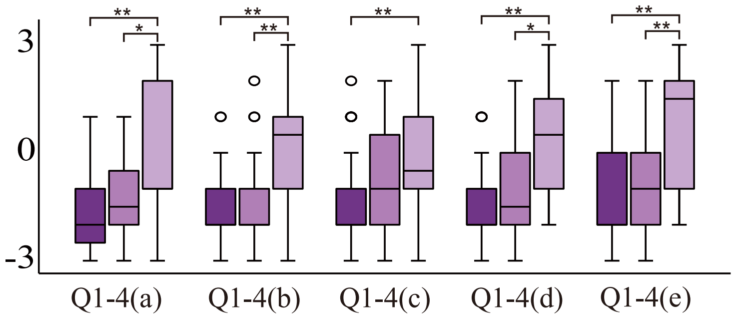

Figure 7. Results of the deceptiveness (Q1-4).

Figure 7 shows the box plots for item Q1-4. Regarding the results of the participants' perceptions of deceptiveness, Kruskal-Wallis test showed significance in all items (Q1-4(a): X^2(2, N = 100) = 19.610, p < 0.001; Q1-4(b): X^2(2, N = 100) = 25.856, p < 0.001; Q1-4(c): X^2(2, N = 100) = 9.489, p = 0.009; Q1-4(d): X^2(2, N = 100) = 23.036, p < 0.001; Q1-4(e): X^2(2, N = 100) = 31.871, p < 0.001). Parison comparisons (Dunn's method with a Bonferroni correction) showed that I-Eye scored significantly higher than I-Consis and I-Head in all items except Q1-4(c): Trustworthy-Untrustworthy. In Q1-4(c), the I-Eye was significantly higher only than I-Consis.

5. DISCUSSION

The result of Q1-3 indicates that the inconsistency in robot's eye-gaze and body orientation was perceived as the modulated intention. The robot's eye movements (gaze) were used as a cue in intention inference, which is aligned with the previous finding [14]. It is worth noting that, the head turning in the I-Head condition was more noticeable than the eye movements in the I-Eye condition, leading the head turning considered to be taken as the cue, while no significance was found. This might be because the participants perceived the eye-gaze of the robot differently in the two conditions: eye movements in I-Eye may have been seen as a more instinctive behavior (e.g., glancing), while head-turning in I-Head may have appeared less instinctive (e.g., watching). There was study [17] took the robot glancing as "a seemingly unintentional behavior" as well as "a leakage cue in interaction". In the I-Eye condition, when the eye movements didn't align with its body orientation, the eye-gaze may be interpreted as the leakage of intention, which exerted a stronger influence than head-turning.

The results of Q1-4 suggested the participants perceived eye movements to be more deceptive than head-turning. Combining with the result of Q1-3, this suggested that the participants perceived the inconsistency in the robot's eye-gaze and the body as a purposeful behavior. Additionally, there was no significance difference between the I-Head and I-Consis in all items in Q1-4. The reason may be the participants perceived the head turning as "watching", which reduce a feeling of deceptiveness. One participant mentioned that the robot rotated its head for using the camera on the head to recognize the cards set. In the experiment, the image recognition function of the robot was not active, and the experimental instruction did not mention nor imply the robot capable of image recognition function.

However, this participant inferred, based on the knowledge and observations towards the robot Kebbi, that the image recognition was in functioning, and thought that the head turning was related to the image recognition function, and thus not deceptive. This could also explain why there was no significant difference between the I-Head and I-Consis in Q1-4. On the other hand, the results suggest an ethical risk. Since the inconsistent behavior increased the perception of deceptiveness, users may misinterpret the intentions that robots are trying to indicate, particularly for children, elderly people, and those people unfamiliar with technology. Future works need to be conducted focusing on how different users perceive inconsistent behaviors and seek design solutions suitable for different people.

This study is a preliminary behavior design of intention modulation and has several limitations. First, the small sample size of 25 participants reduces the statistical power of the results. Additionally, to simplify the robot's behavior, the intention targets in front of the robot were limited to only two, and they were placed symmetrically in front of the robot. In real-world interaction, intention targets are more varied and dynamic, and the experimental results may not be easily generalized.

6. CONCLUSION

This study presented a behavior design for modulated intention indication in social robots by introducing inconsistencies between a robot's eye-gaze, head, and body orientation. The study conducted a robot interaction experiment in the scenario of involving the participants and the robot collaborating to make "Ouen Uchiwa". During the experiment, the robot approached one set of decoration cards, while its head and eyes might turn towards another set. The results indicated that inconsistency between the robot's eye-gaze and body orientation was perceived as modulated intention, and eye movements were perceived more deceptive than head turning. It suggests the possibility for modulating robots' intention indication through inconsistency in eye-gaze and body orientation. The study addressed robot behavior design (i.e., how robots behave), without considering the decision-making process (i.e., how intention should be modulated). In the future, the decision-making in modulating intentions of robots will be further researched.

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant Numbers JP 22H04871 and Development of Innovative Science and Technology of Kansai University.

REFERENCES

[1] Shimada, M. and Kanda, T. What is the appropriate speech rate for a communication robot? *Interaction Studies,* 13, 2012, pp.406–433.

[2] Moro, C., Lin, S., Nejat, G. and Mihailidis, A. Social Robots and Seniors: A Comparative Study on the Influence of Dynamic Social Features on Human–Robot Interaction. *Int J of Soc Robotics,* 11, 2019, pp.5–24.

[3] Robinson, H., Macdonald, B., Kerse, N. and Broadbent, E. The psychosocial effects of a companion robot: a randomized controlled trial. *J Am Med Dir Assoc,* 14, 2013, pp.661–667.

[4] Wang, R. H., Sudhama, A., Begum, M., Huq, R. and Mihailidis, A. Robots to assist daily activities: views of older adults with Alzheimer's disease and their caregivers. *Int Psychogeriatr,* 29, 2017, pp.67–79.

[5] Products - PUDU. Available: https://www.pudurobotics.com/products/bellabot. [Accessed on 2024, 17 November].

[6] Servi Plus — Bear Robotics. https://www.bearrobotics.ai/servi-plus. [Accessed on 2024, 17 November].

[7] Dragan, A. D., Lee, K. C. T. & Srinivasa, S. S. Legibility and predictability of robot motion. *In 2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2013, pp.301–308.

[8] Basili, P. et al. Inferring the goal of an approaching agent: A human-robot study. *In 2012 IEEE RO-MAN: The 21st IEEE International Symposium on Robot and Human Interactive Communication*, 2012, pp.527–532.

[9] Kobayashi, K. & Yamada, S. Expressing Robot's Mind by Human-like Motion. *Transactions of the Japanese Society for Artificial Intelligence*, 21, 4, 2006, pp.380–387.

[10] Takayama, L., Dooley, D. & Ju, W. Expressing thought: Improving robot readability with animation principles. *In 2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2011, pp.69–76.

[11] Becker-Asano, C. and Ishiguro, H. Evaluating Facial Displays of Emotion for the Android Robot Geminoid F. *In 2011 IEEE Workshop on Affective Computational Intelligence (WACI),* 2011, pp.8.

[12] Salem, M., Kopp, S., Wachsmuth, I., Rohlfing, K. and Joublin, F. Generation and Evaluation of Communicative Robot Gesture. *Int J of Soc Robotics*, 4, 2012, pp.201–217.

[13] Nakano, Y. I., Yoshino, T., Yatsushiro, M. and Takase, Y. Generating Robot Gaze on the Basis of Participation Roles and Dominance Estimation in Multiparty Interaction. *ACM Trans. Interact. Intell. Syst.*, 2015, Vol.5, pp.22:1-22:23.

[14] Yang, L., Sejima, Y. and Watanabe, T. Gaze cue: which body parts will human take as cue to infer a robot's intention? *Journal of Advanced Mechanical Design, Systems, and Manufacturing*, 18, 2024, pp.JAMDSM0060–JAMDSM0060.

[15] Mukai, A. Relationship between perceived deceptiveness and perceived personality traits. *The Proceedings of the Annual Convention of the Japanese Psychological Association*, 2013, Vol.77, pp.1AM – 111.

[16] Murai, J. The Effect of Intensifiers on Perceived Deceptiveness of Written Messages. *The Japanese Journal of Personality*, 14, 2005, pp.92–100.

[17] Mutlu, B., Yamaoka, F., Kanda, T., Ishiguro, H. and Hagita, N. Nonverbal leakage in robots: communication of intentions through seemingly unintentional behavior. *In Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*, HRI'09, pp.69–76.