Directing Attention through Gaze Hints Improves Task Solving in Human-Humanoid Interaction

Eunice Mwangi[1], Emilia I. Barakova[1], Marta Díaz[2], Andreu Català Mallofré[2], Matthias Rauterberg[1]

[1]Designed Intelligence Group
Department of Industrial Design
Eindhoven University of Technology, Netherlands

[2]Technical Research Centre for Dependency Care and
Autonomous Living (CETpD)
Technical University of Catalonia, Spain

ABSTRACT

In this paper, we report an experimental study designed to examine how participants interpret and perceive social hints from gaze exhibited by either a robot or a human tutor when carrying out a matching task. The underlying notion is that knowing where an agent is looking provides cues that can direct attention, to an object of interest during a task. In this regard, we asked human participants to play a card matching game in the presence of either a human or a robotic tutor. In one case, the tutor gave hints to help the participant find the matching cards by gazing toward the correct match, in the other case, the tutor only looked at the participants and did not give them any help. The performance was measured based on the time and the number of tries taken to complete the game. Results show that gaze hints (helping tutor) made the matching task significantly easier (fewer tries) with the robot tutor. Furthermore, we found out that the robot's gaze hints were recognised significantly more often than the human tutor gaze hints, and consequently, the participants performed significantly better. The reported study provides new findings towards the use of non-verbal gaze hints in human-robot interaction, and lays out new design implications, especially for robots used for educative purposes.

Keywords - Gaze-based interactions; gaze perception; game-based human-robot interaction; embodied cues; attentional cues; directed attention; facial orientation.

I. INTRODUCTION

At present, robots are showing increasing potential to be effectively incorporated into various social settings, for example, in educational and therapeutic facilities for children and in nursing homes for elderly among others [5, 9, 23]. Accordingly, it is increasingly important to design intuitive social behaviors for robots. In human interactions, people rely on non-verbal cues such as gaze, gestures, body language, and facial expressions to communicate meaning. Feldman and Rimé, showed that non-verbal cues provide 93% of the meaning exchanged in interaction [17]. These findings suggest that implementing these cues in social human-robot interaction may increase its naturalness and effectiveness.

In an attempt to reach natural interaction, much research has been done into social human-robot interaction on mapping non-verbal human behavior to robots [10, 19]. Among non-verbal behaviors, the gaze is considered as primary source of information [4, 20, 22]. Gaze plays a significant role in social interaction, particularly in directing attention. Moreover, gaze behavior facilitates a range of social functions during human - human interactions such as communicating their emotions, intentions, and what they are attending to [18, 20, 39]. Additionally, it was shown that at a very young age, children can follow the gaze of their parents or their caregivers [11, 12]. On the contrary, children with development disorders such as children with ASD show typical difficulties in producing and reading non verbal communicative behaviours both gestures and gaze-based [33]. Perhaps, one of the most significant roles of gaze is its capability to direct attention to objects of interest in the environment facilitating the formation of joint visual attention [16, 18]. Joint attention is argued to be the basis for early language learning and normative development in children [7, 11, 12].

The use of eye gaze behaviour in human-robot interaction is gaining the much-needed attention as well. Prior research has demonstrated how gaze can be used to build better interactions with robots [13, 19, 26, 27, 28, 29 30, 37]. "Leakage cues" –that are unintentional and unconscious both in their production and perception have been studied by Mutlu et al. [28] who showed that participants could read and interpret such cues from robot gaze. They evaluated the perception of gaze cues that were implemented on ASIMO, Robovie, and Geminoid robots. While the above work focuses on unintentional and unconscious use of gaze cues in a competitive setting, in the current study we focus on deliberate attention directing eye gaze cues in a collaborative setting with a tutor. Palinko et al. [31] studied the impact of the use of eye- or head-based gaze estimation in a human-robot interaction experiment with the iCub robot. The robots used in the studies mentioned above can move their eyes independently of the head. Since cheaper and better accessible robots are the more feasible choice for use in education and elderly care, in the current study, we want to investigate whether robots such as NAO can also convey such deliberate meaning through gaze cues. We further examine whether humans can read such cues and accept help from the robot, and, in turn, if these cues influence the decision-making of the interacting human.

For this purpose, we designed an experimental study to evaluate the effects of gaze hints in the context of the educational gameplay. We asked participants to play a matching card game in the presence of a human or a robotic tutor. The aim of the study was to determine if gaze hints from the tutor can direct attention and, in turn, influence the choices of human partners. Figure 1 describes the interaction flow for the designed study: -
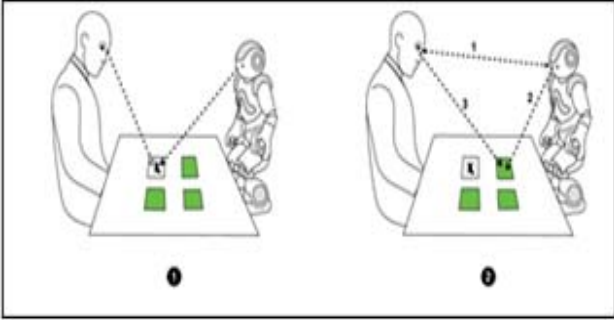

Figure 1: The interaction flow: P-Participant (left) turns over a card; T-Tutor (right) looks at the selected card; Tutor gazes at participant to draw attention; T–Tutor moves on to look at a matching.

We expected that the participants would notice the gaze of the tutor while the tutor was gazing at different cards on the table and that they would follow the tutor's lead to the matching card. Thus, we hypothesized that gaze cues (facial orientation and gaze direction) from the tutor would help in drawing the participant's attention to the matching card, and subsequently influence their choices. To create a more accurate measure for the interactive behaviours of the human participants, we incorporated use of eye tracking to make a precise recording of the gaze behaviour of the interacting human.

In this paper (Section II), we describe prior work on gaze in social human-robot interactions. We further describe the methodology and the design of the human-human and human-robot experimental set-ups in section III, and the results of the study in Section IV. Finally, in Section V, we discuss the findings and limitations of our work and give directions for future work.

II. RELATED WORK

Given the critical role of gaze in human communication, research into designing social gaze behaviors for robots has been extensive [2, 13, 31, 34]. Andrist *et al.* [3] combined three functionalities including face-tracking, head detection, and gaze aversions to create social gaze behaviors for conversational robots. In an evaluation study, the participants indicated they perceived the designed gaze as more intentional. Admoni *et al.* [1] addressed the impact of frequency and duration of gaze on the perception of attention during the human-robot interaction concluding that shorter, more frequent fixations are better for signifying attention than longer and less frequent fixations.

In a storytelling setting, Mutlu et al. showed that participants recalled the story better when the robot looked longer at them [27]. Yoshikawa et al. [41] explored both responsive and non-responsive gaze cues and found that the responsive gazes gave a

strong effect of "feeling of being looked at" during interaction. Moon et al. [26] studied the effects of gaze behaviors in a handover task. They found that gaze cues can improve the hand-over timing and the subjective experience in a hand-over tasks. Weise et al. showed that individuals with ASD are more likely to follow gaze of robots, relative to humans [43]. Boucher et al. [8] studied gaze effects on the speed of communication in both human–human and human-robot interaction collaborative works. Their results demonstrate that human participants can use gaze cues of a human or a robot partner to improve their performance in physical interaction task.

Several studies have considered the ability of people to read and follow robot gaze. For example, using a guessing game, Mutlu et al. [28] show that participants can read and interpret leakage cues from robot gaze even faster when the robot is more human-like. Their designed gaze behaviors were evaluated on ASIMO, Robovie, and Geminoid robotic platforms that can move their eyes independently of the head direction. This poses the question to what degree simpler and more available robots can also perform gaze cuing effectively.

In this line of research, Cuijpers et al. [14] used NAO robot, which has no moveable eyes and measured the region of eye contact with the robot, they concluded that perception of gaze direction with NAO robot is similar to a human looker. Mwangi et al. [29] examined the ability of people to correctly guess the head direction towards different target positions (cards) on a table using the NAO robot and showed that participants perceive the head (gaze) direction of NAO robot more accurately for close objects and also that the cards positions left and right of the player are recognised with different accuracy. These related works suggest that robots without movable eyes as Nao robot can be used as well for providing gaze cues.

Prior work has also focused on the role of gaze in joint attention. Pfeiffer-Lessmann *et al.* [34], examined the timing of gaze patterns in interactions between humans and a virtual human to build a joint operational model for artificial agents. Yu *et al.* [42] studied the timing patterns of gaze when interacting with either a robot or a human in a word learning task. Their eye-tracking result revealed that people pay more attention to the face region of the robot than that of the human during a word learning task.

In the present comparison study, we aim to determine whether gaze hints from a tutor (either human or robot) can direct player attention and therefore influence the choices of human partners in a game-play. This brings us to several sub-questions: i) are the gaze cues provided noticed, ii) is it understood as helping behavior and iii) does the help provided by the tutor influence performance. Based on the literature of non-verbal communication and gaze behavior, our assumption is that knowing where another agent is looking provides hints that can help to cue attention and influence decisions and thoughts, and therefore improve the performance.

III. MATERIALS AND METHODS

A. The Experimental setup

To test our hypothesis, we formulated an experimental task, in which participants interacted with a tutor –either a robot or a

human- in a variant of the Memory card game. The game is played with fourteen cards (7 matching pairs), with images of black dogs, varying slightly in shape, to increase the level of difficulty. On the table, there were 14 cards arranged in a rectangular layout. The layout had six columns and three rows for a total of 18 cards, or 9 pairs. We placed the cards in the first two rows, and two cards in the middle positions of the third row. The distance between the cards on the X-axis was 6 cm; on the Y-axis, it was 10 cm. This arrangement was informed by our prior experiment [29], where we examined whether the participants can accurately perceive gaze direction and what was the resolution needed for the head angles of the robot to direct at different card locations on the board (see Figure 4).

At the beginning, the cards were laid face down on the board, and then the player selected a card and tried to find a matching card. If the cards turned face up were similar (a pair of matching cards), then the player continued to match the cards; otherwise, the participant turned the cards face down and made a new try/move. The goal was to find all pairs in the smallest number of moves/tries, and shortest time, possible. An attempt (try) consists of choosing two cards; the game ended when the participant found all the matching pairs. Although the game is better and more enjoyable when two players play against each other, it can also be played by a single player. In our research experiment, the participant played the game alone, in the presence of either a robot or a human tutor.

Figure 2: The Set-up: Human–robot set-up (Top); Human-human set-up (Bottom).

1) Human - Human Set-up

Figure 2 (bottom) depicts the human-human set-up. The tutor and the participant sat across the table, approximately 160 cm apart. The human tutor was trained to follow a pre-defined protocol of steps that detailed the rules of how to introduce the game and the sequence of how to shift her gaze during the game. Figure 3 captures the sequence of tutors' gaze (human and robot tutor) from the eye tracking videos in the help condition. The tutor first looks at the chosen card; then looks to the face of the participant, and then looks to the matching card. This sequence

of gaze shifting was consistent for both tutors in the help condition.

The same person acted as the tutor for all the participants, and during all sessions. In front of the tutor was a printed photo of the card locations on the board layout (see Figure 4, left). After each session with a participant, the cards were re-arranged. In this set-up, we logged the gaze of both the tutor and the participant. The gaze data of the tutor were registered using the eye-tribe tracker, while the participant wore SMI eye-tracking glasses to capture their eye gaze behavior. The experiment was recorded using I-view ETG software, and also with a video camera. The reason for recording the gaze of the tutor was to examine if the tutor gaze was consistent in all sessions, and if varying behavior on the part of the tutor influenced the participant gaze data in any way.

2) Human - Robot Interaction Setup

In the robot condition a humanoid robot NAO developed by Aldebaran Robotics [33], a personal computer, a web-cam, and the same memory game (see Figure 2) was used. NAO is a 57cm tall robot with a moveable head and facial features that bear a resemblance to those of a child. As a result of its minimalistic design and perception capabilities, NAO robot has been adopted widely for research focused on therapeutic training, or for general educational/pedagogical purposes. To develop the game, we used Java Programming Language. Each card was labelled with a unique card code, and placed in a fixed position on the board layout marked with a head pitch and yaw angle on the computer layout. The algorithm was applied such that, after scanning the code of the selected card, the robot head angles shifted to the card position of the chosen card, then to the face of the participant (assumed at NAO initial position), and then to the location of the matching card. The design of the help state gaze for the robot follows the concept of attention-directing gaze movement in human communication. Another aspect we considered is the timing of gaze behaviour; we invited fellow students to the lab and played the robot gaze motions to different cards on the table for them with two different timings. We used the timing of head movements that they regarded as more natural.

The figure below shows the image sequence of the designed tutor gaze: -
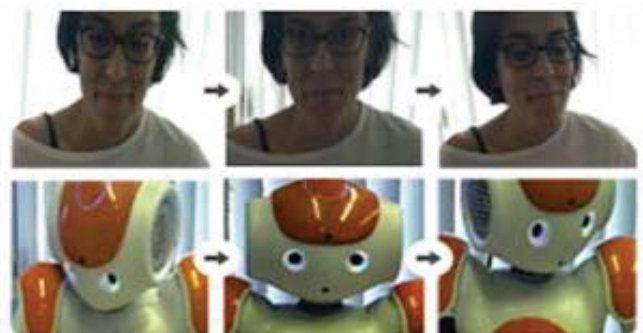

Figure 3: Tutor gaze captured from eye tracking videos: Human tutor (Top); Robot tutor (Bottom); Tutor looks at the chosen card; looks to the participant, and then looks to matching card.

B. Experimental Design and Conditions

The study followed a two - by - two (Tutor_Type: Human or Robot) and (Help_Type: Help vs. No-Help) mixed factorial design experiment. The Tutor_Type variable was manipulated as a between-participants and the Help-Type as a within-participants. Participants interacted either with a robot tutor or a human tutor in two conditions.

Both experimental conditions, Help and No_Help, were executed by the tutor for evaluation (see Table 1). In the help condition, the tutor provided gaze hints to help the participant find the matching cards. While introducing the game, the tutor informed the participant that she would help them. However, the tutor did not explicitly reveal the modality she would use to help. The tutor remained silent the entire session; when the participant turned the card upwards, the tutor gazed to see the flipped card and continued to look at the matching card. In the No_Help Condition, the tutor remained silent and only stared at the participant.

Participants were evenly assigned to play in the presence of either the robot or the human tutor. Each participant interacted with the tutor in both conditions, and the order of conditions was counterbalanced across trials. In both set-ups, the robot and the human tutors performed similar actions.

C. Hypothesis

This work presents an experimental study aimed at investigating the effect of gaze cues provided by a human or a robot tutor during a memory game with cards. In this regard, we formulated two hypotheses as outlined below regarding how help (presence of gaze hints) and the type of tutor (human/robot) and affect participants' performance of the task, and also how participants' gaze of the participants would differ between interactions with the human and humanoid tutor.

Hypothesis One (H1): Participants will perform better in the Help condition than in the No_Help condition

- H1.1: Participants will complete the task in less time in the Help condition than in the No_Help condition

- H1.2: Participants will complete the task with fewer tries in the Help condition than in the No_Help condition.

Hypothesis Two (H2): The type of tutor (human/robot) will influence the participants' task performance both in terms of time and numbers of tries.

Hypothesis Three (H3): The type and style of tutor will influence participants' gaze behaviour with the tutor during the play

- H1: Participants will look more into the tutor's face in the Help condition than in the No_Help condition.

- H2: Participants will look more into the robot tutor's face than in the human tutor's

D. Measurements

To evaluate the above mentioned hypothesis we employed the following measures:-

Task Performance: - We identified two primary objective measures that are notably used to measure performance in memory game: (1) Duration: - the time it takes the participants to find all pairs of matching cards on the table; and (2) Number of tries:- the total number of attempts required to find all matching cards. A "try" consists of choosing two cards. All sessions were video-recorded to facilitate the analysis according to of these measures.

Gaze Behaviour: - While playing the game, the participants wore SMI eye tracking glasses, to capture the gaze direction. We recorded the these data, in order to compare the gaze patterns when the participants interacted with the human and the robot. Eye tracking provides data with high temporal resolution and can, therefore, reveal precise patterns of gazing behavior.

Perceptions: - We used a questionnaire to evaluate participants perceptions of tutor behaviours, particularly perceived likeability, perceived presence of the tutor, and also the feedback about the task. We included open-ended questions in the questionnaire asking subjects to list the helping cues that they observed or searched for in the tutor's behaviour. At the end, we conducted semi-structured interviews to assess whether participants perceived any differences between the two conditions of the game. During the sessions, we documented all the observations made.

E. Procedure

Before the participants entered the room, the experimenter placed the cards in their correct locations on the board layout as shown in Figure 4.

Figure 4: Card arrangement: Set-up One (Help) Left; Set-up Two; (No_Help) Right.

Upon receiving informed consent, the experimenter verbally provided the participant with details regarding the task and directions on how to play the game, including the use of SMI Eye-Tracking Glasses. At this point, the experimenter made sure not to disclose the objective of the study. The SMI eye-tracking glasses were fitted on the participant. The experimenter then performed a calibration procedure for the eye-tracking system.

The participant followed the tutor's instructions to complete the game. Once the game ended, the participants filled out a post-experiment questionnaire After each session, the study took a pause of around four minutes; the participant was requested to wait outside the room for the experimenter to reorganize the game for the second session. This procedure was then repeated for the other condition. After interacting with the tutor in both conditions (Help and No_Help Condition), the experimenter collected demographic details and interviewed the participants to get more information on any differences they may have observed in the tutor's behaviors between the two conditions.

F. Participants Profile

The 20 participants who took part in the study (Gender: 11 males and 9 females, Age: 19-33) were students recruited from the university campus. The participants were from different cultural backgrounds, namely, China and different European countries. The study took 30 minutes, and participants were given a coupon worth 10 Euro at the end of the experiment for their participation.

IV. RESULTS

The paragraphs below report on the results, from our performance, gaze and subjective analysis. We first describe results from the performance measures, in order to provide background for the eye-tracking and subjective measures.

A. Performance measures

For objective analysis, we conducted a mixed-model ANOVA in SPSS, with the repeated measure Help_Type (Help vs. No_Help) as the within - subject factor and the Tutor_Type (Robot or Human) as the between -subject factor. We analyzed the results of 20 participants (10 for the Robot tutor condition, and 10 for the Human tutor condition, for a total of 20 trials in the Help condition and 20 trials in the No-Help conditions across both tutor conditions. We analysed the effect of Help and Tutor type on the following two performance measures:-

Duration: We obtained the duration from video recordings, this being the period between the participant starting to play the game and completing it. **Number of tries:** We counted the number of tries that participants used from our video recordings.

Tables 1, and 2, and Figure 5 provide results from the defined objective measures:-

Table 1: Performance Measures; Mean Duration (s) and Number of tries

Durations (s)

| Tutor_Type | Help | | No_Help | | |
	Mean	Std. Deviation	Mean	Std. Deviation	N
Human	118.00	38.306	118.70	45.019	10
Robot	124.40	27.774	145.10	83.523	10
Average	121.20	32.730	131.90	66.692	20

Number of tries

| Tutor_Type | Help | | No_Help | | |
	Mean	Std. Deviation	Mean	Std. Deviation	N
Human	16.20	5.116	19.00	5.121	10
Robot	11.30	3.057	17.00	5.925	10
Average	13.75	4.811	18.00	5.487	20

Effect of Tutor_type on performance measures: - Tests of between-subject effects based on the F tests of the averaged variables show a significant multivariate effect of Tutor_Type on performance measures (p=0.048). There was no significant main effect of Tutor_Type on duration (F (1, 18) =0.913, p=0.352). However, we found a significant main effect of Tutor_Type on the number of tries (F (1, 18) =5.253, p=0.034).

Effect of Help_Type on performance measures: - Tests of within-subject effects based on the F tests of the averaged variables shows a significant multivariate effect of Help_Type (p=0.003). Thre was a significant main effect of Help_Type on the number of tries (F (1, 18) =7.009, p=0.016). However, there was no significant main effects on duration between the Help and No_Help conditions. (F (1, 18) =0.428, p=0.521).

We found no significant Help_Type by Tutor_Type interaction (p = 0.654). Similarly,there was no Significant Help_Type by Tutor_Type interaction effects on the duration F (1, 18) =0.374, p=0.549 and number of tries (F (1, 18) =816, p=0.378).

Table 2: Effect of Help at different levels of the Tutor_Type

Duration (s)

Tutor_Type	Help_Type		MD	Std. Error	Sig.
Human	Help	No_Help	-.700	23.133	**0.976**
Robot	Help	No_Help	-20.7	23.133	**0.383**

Number of Tries

Tutor_Type	Help_Type		MD	Std. Error	Sig.a
Human	Help	No_Help	-2.8	2.27	**0.233**
Robot	Help	No_Help	-5.7	2.27	**0.022***

Effect of Help at different levels of the Tutor_Type:-

Duration: For the Human condition, there was no significant difference in duration of the Help and No_Help condition (p=0. 976). Similarly, the mean difference in duration between the Help and No_Help condition was not significant for the Robot condition (p=0. 383).

Number of Tries: For the Human condition, there was no significant difference in the number of tries between the Help and No_Help (p =0. 233). However, for the Robot condition, the mean difference in the number of tries between the Help and No_Help condition is (MD = 5.7), was significant (p=0.022); indicating participants performed significantly better, with fewer tries, with help from the robot tutor than without help.

Counterbalanced Orders: Performance Measures

Participant interacted with the tutor in both conditions(Help and No_Help) in a different order: Group 1 (Order: Help; No_Help) and Group 2 (order No_Help; Help). To examine order effects on performance measures, we created a data matrix with six columns. SubjectID, Group: Group 1 (Help; No_Help) and Group 2 (No_Help; duration and number of tries , two for each condition. From the analysis, we found a significant Group * Help interaction effect on duration; thus, there was an overall order effect. The differences were larger for Group 2 (No_Help; Help) than for Group 1 (Help; No_Help). However we found no significant Group * Help interaction effects on the Number of tries.

Pairwise comparisons (adjustment for multiple comparisons: Bonferroni) between the Robot and the Human condition, reveal that participants used significantly fewer tries with help from the Robot tutor than with the help from Human tutor (p=0.018, two-tailed). However, there were no significant differences in the number of tries for the two groups (Human and Robot) when help was not present (p=0. 430, two-tailed), assuming equal variances. Comparing both groups on duration, we found no significant differences in both Help (p=0. 674, two-tailed) and No_Help (p=0. 391, two-tailed) conditions.

20% of the participants in the Human condition reported noticing help gaze hints while 60% in the Robot condition said they recognised gaze hints. For the Human condition, there was no significant difference between those who reported identifying the gaze hints and those who did not report identifying the gaze hints across both measures(duration and number of tries). For the Robot condition, there was a significant correlation between noticing gaze hints and the number of tries. Participants who reported seeing the gaze hints in the Robot condition performed ignificantly better, measured by the number of tries than those who did not report identifying the gaze cues. However, there was no significant difference in duration for those who identified gaze hints in the robot condition and those who did not report identifying the gaze hints. We provide probable explanations for these findings in the discussion section. Figure 6 shows how individuals who did not recognise that the tutor was helping performed if compared with participants who made use of the hints. The figure shows the number of attempts and duration in the "Help" condition with text representing subjects' awareness of the hints (YES or NO) for both Human and Robot Tutors.
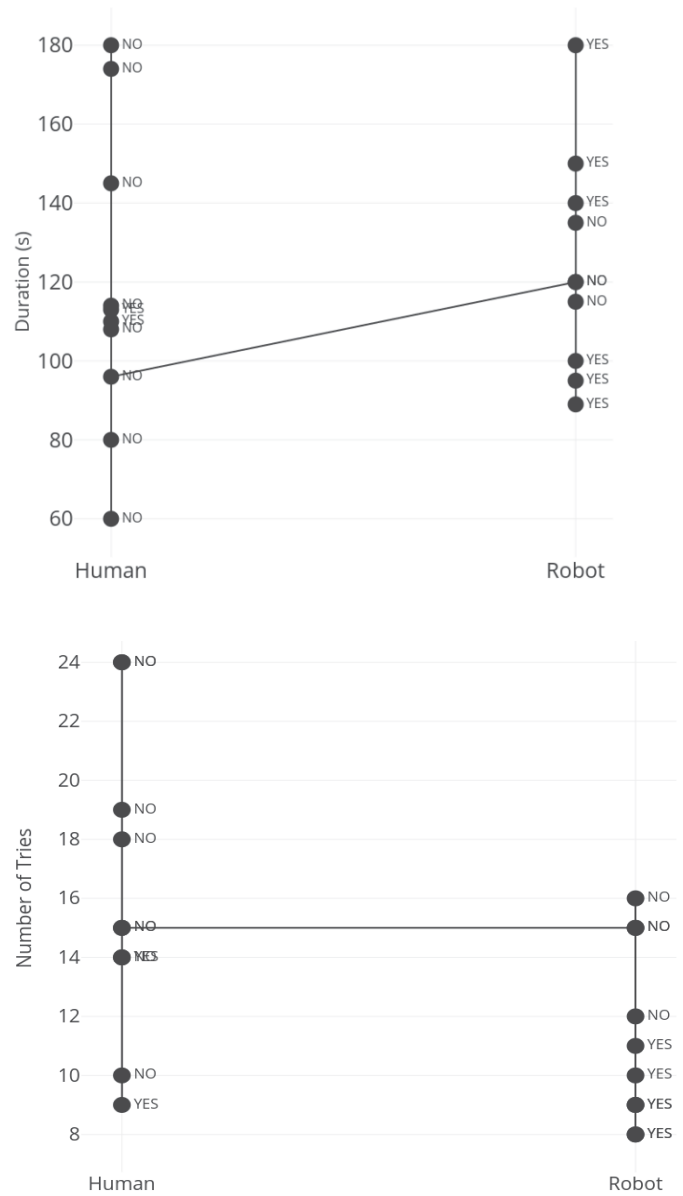

Figure 6: Noticing gaze in Help_condition and performance measures (Top; Duration; Bottom;Number of Tries)
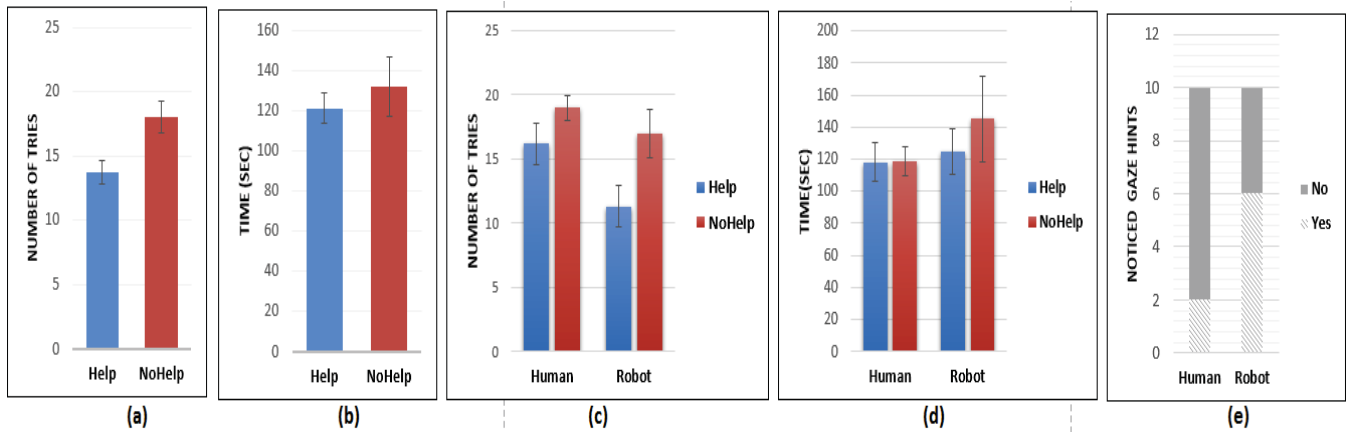
Figure 5: (a) Number of tries with and without help. (b) The time (seconds) it took participants to find matching cards with and without help. (c) The number of tries with and without the help for the two tutors. (d) The time (seconds) to match all the cards with and without the help from the two tutors.

B. Eye – Gaze Measures

In this section, we analyse the eye gaze patterns when the participants interacted with the human and the robot tutor. To analyse the recorded gaze data from the video, we used Begaze software to create custom trials of the video recordings, which included only the segment from the participant starting to play the game until completion. From the trial images, we cropped the face of the robot and that of the human tutor as our area of reference (AOI_Face) as shown in figure 7 below: -
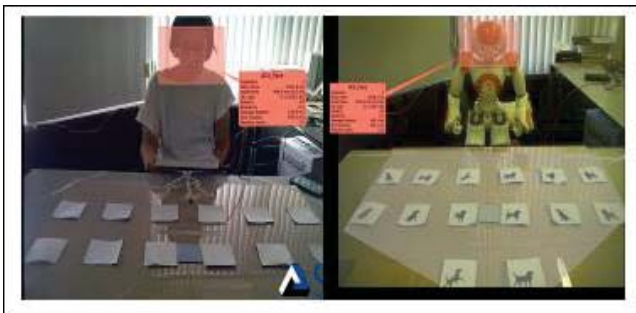

Figure 7: Area of interest (AOI) region: Face of the robot and the human tutor.

Next, we exported the metrics related to the trials and area of interest to SPSS software. To analyse the eye gaze measures, we conducted a mixed-model ANOVA in SPSS, with the repeated measure Help_Type (Help vs. No_Help) as the within - subject factor and the Tutor_Type (Robot or Human) as the between - subject factor. We analyzed the results of 20 participants (10 for the Robot tutor condition, and 10 for the Human tutor condition, for a total of 20 trials in the Help condition and 20 trials in the No-Help conditions across both tutor conditions. We analysed the effect of Help and Tutor type on the following eye gaze measures:

Fixation count: this refers to the number of fixations within an area of interest(AOI_face).

Fixation time ; Refers to the total fixation time on the face area (AOI_Face) and is often associated with attention and visual processing, in milliseconds

Dwell time: The time spent looking at the area of interest, calculated by summing up the time the gaze coordinates were within the face area (AOI_Face).

Table 3: Eye - tracking measures
Fixation count (AOI_Face)

	Human	Robot	Average
Help	4.6	19.8	12.2
No_Help	2.9	5.6	4.2
Average	3.7	12.7	8.2

Fixation time (Ms.) (AOI_Face)

	Human	Robot	Average
Help	1097.9	4205.7	2651.8
No_Help	558.2	1001.5	779.9
Average	828.1	2603.6	1715.9

Dwell Time(Ms.) (AOI_Face)

	Human	Robot	Average
Help	1760.53	5749.22	3754.88
No_Help	285.76	4237.06	2262.41
Average	1028.15	2408.14	3008.65

Results of the eye - gaze analysis show that participants generated different patterns depending on the tutor they were interacting with. There was no significan main effect of Help on Fixation count: $p=0.081$; Fixation time: $p=0.070$. However we found that there was a significan main effect of the tutor type on the fixation measures: Fixation count: $p=0.020*$ Fixation time: $p=0.050*$.

Comparing the effect of Help on eye - tracking measures at different levels of the Tutor, we found significant differences in both Help and No_Help conditions for the robot condition for

both Fixation measures; Fixation_count *(p=0.028)* and Fixation time *(p=0.028)*. However, there were no significant differences on all measures for the human condition during Help and the No_Help conditions. Results show that the total number of eye fixations on the robot face was significantly higher than the number of those on the human.

C. Participants' perceptions

For subjective analysis, we conducted a mixed model analysis of variance in SPSS [], with the repeated measure Help_Type (Help vs. No_Help) as the within - subject factor and the Tutor_Type (Robot or Human) as the between -subject factor. We analyzed the results of 20 participants (10 for the Robot tutor condition, and 10 for the Human tutor condition, for a total of 20 trials in the Help condition and 20 trials in the No-Help conditions across both tutor conditions. To facilitate the analysis of the Likert scale data, we coded the data as follows: 1: Completely Disagree; 2: Disagree; 3: Neutral; 4: Agree; and 5: Completely Agree. We also included the 'not applicable option, and these data were treated as missing values during the analysis. Figure 7 and Figure 8 summarize the results from our subjective measures, particularly the perceived likeability and presence of the tutor.

Likeability: To evaluate the likeability of the tutor, participants were asked to rate the following three statements about the tutor: Tutor was pleasant; Tutor was kind; Tutor was likeable; and Tutor was friendly in a five point likert scale.

There was a significant multvariate effet of help on likeability measures (p=0.046). We found a significant main effects of Help on the Pleasant measure $(F (1, 16) =7. 828, p=0. 015)$. However, there was no significant effects of Help on the participants rating of tutor's Kindness $F (1, 16) =1. 361, p=0. 260)$, Friendliness $(F (1, 16) =2. 028, p=0. 174)$, and Likeable measures $(F (1, 16) =1. 361, p=0. 260)$. There was no significant main effects of Tutor_ type on any of the Likeability variables used. We found no significant multivariate interaction (Tutor_Type * Help_Type) effects on presence measures (p=0.144). We found significant interaction effects interaction on pleasant measure $(F (1, 18) =4.893, p=0.040)$; However there were no sfnifican interaction effects on the other three measures.
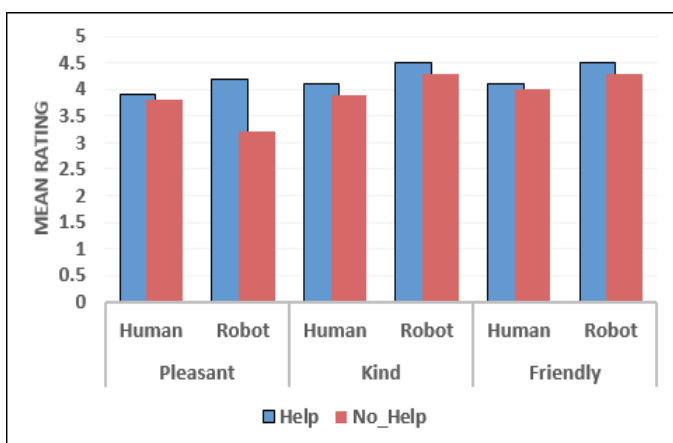

Figure 8: Comparing the effect of Help on Likeability measures at different levels of the Tutor_Type.

Pairwise comparisons based on the estimated marginal means between the two tutors on the likeability measures show no significant difference regarding how both tutors were rated for all likeability measures used.

Comparing the effect of Help_Type on Likeability measures at different levels of the Tutor_Type, we found that the participants rated the robot tutor as more socially pleasant *(p=0.005)* and more likeable (p=0.047) when the robot tutor gave gaze hints than when they did not. However, there was no significant difference in ratings for Kindness and Friendliness. On the other hand, there was no significant difference in how participants rated the human tutor in both Help and No_Help conditions across all Likeability measures used.

Presence: A test of within-subject effect shows a significant multivariate effect of Help on presence measures (p=0.003). There was significant main effects of Help_Type on Tutor_Presence $(F (1, 18) =15.059, p=0. 001)$; Tutor caught the participants attention $F (1, 18) =9.529, p=0. 006$; and Tutor was attentive $F (1, 18) =6.600, p=0.019))$. A test of between-subject effect shows a significant multivariate effect of Tutor_Type on presence measures (p=0. 001). There was a significant main effect of Tutor_Type on Tutor caught the participant's attention measure $F (1, 18) =5.921, p=0. 026)$; However the main effects of Tutor_Type on Tutor_Presence $(F (1, 18) =3.009, p=0. 100)$; and Tutor was attentive $F (1, 18) =1.638, p=217)$ were not significant.

We found no significant multivariate interaction (Tutor_Type * Help_Type) effects on presence measures (p=0.090). We found significant interaction effects interaction on Tutor's perceived presence$(F (1, 18) =4.366, p=0.051)$; and how the tutor's behaviour caught the attention of the participant $((F (1, 18) =4.235, p=0.054)$; but the interactions effects on perceived attentiveness $(F (1, 18) =0.492, p=0.491)$ were not significant. Pairwise comparisons on the effect of Help on Presence measures, at different levels of the Tutor_Type, show that participants rated the robot tutor as more socially present (p=0.001), and more attentive (p=0.33) in the Help condition than in the No_Help condition. Moreover, they indicated that the robot tutor behaviour caught their attention more during help than when there was no help (p = 0.002). However, we found no significant difference on how the participant rated the human tutor across all presence measures in both Help and No_Help conditions.

Pairwise comparisons based on the estimated marginal means between the two tutors on the presence measures show that the robot tutor behaviour caught the attention of the participants significantly more compared to the human tutor (p=0.026). However, there was no significant mean difference in how both tutors were rated for the other two measures.
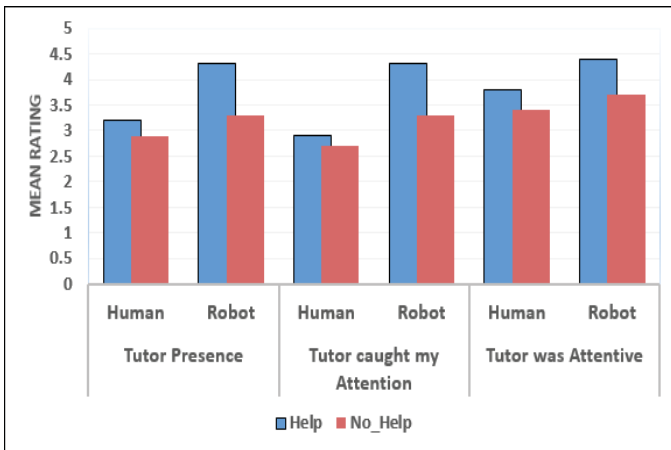
Figure 9: Comparing the effect of Help on presence measures at different levels of the Tutor_Type.

D. Post-experiment interview

In the end, we conducted a post-experiment interview. We asked participants whether they noticed the help from the tutor and, if they did, whether that influenced their choice of cards. We also asked what cues they expected the tutor would use to help them in the game. Lastly, we examined how they perceived the robot tutor head movements: whether they were natural, too fast, or too slow, and if they considered the behavior of the robot to be automatic.

Eight (8) participants in the human condition reported not noticing the help, i.e., the gaze hints from the tutor. Most of them indicated they were focused on the game and did not look at the tutor. At least eighteen (18) of all the twenty (20) participants in both human and robot conditions said they expected verbal/vocal/audio help from the tutor, as illustrated in the excerpts below:

> P003: [Human condition] *"I expected verbal hints from the tutor, I didn't pay attention to the tutor."*
>
> P006: [Human condition] *"I noticed the gaze hint but it took a while to get to the eyes. I expected the speech, for example, when I get closer to the matching card the tutor says something like warmer!!, or hey, you have seen this card before. If there is no help, you are forced to remember, but now I relied on help."*
>
> P009: [Human condition] *"I was focused on the game so I did not look at the tutor, maybe in the future you can put lights with different colors under the card or give some sound instructions. Audio feedback could be important".*
>
> P017: [Robot condition] *"I noticed and used the help as direction, but I prefer more speech."*

Four (4) people in the robot group did not notice the gaze hints. However, all of them reported seeing the head movements, but felt like the robot was following their moves rather than directing their attention to the matching cards as illustrated in the following excerpts:

> P016: [Robot condition] *"I noticed it (the robot) was looking around but felt like it was following my moves, not showing me the positions. I felt as if we were both exploring."*

> P018: [Robot condition] *"It seemed like he sees what I see! I did not get help! I was more focused on the task, though, I thought he was watching what I was doing."*
>
> P011: [Robot condition] *"I was focusing on the cards and I was not looking at the robot. I thought the help would be vocal, the robot looked like it was moving with me, following me."*
>
> P012: [Robot condition] *"I did not feel like it was helping, just thought he was looking at the cards I was turning. I did not get the hint. Maybe it should have been verbal clues for example, say left, right, or centre."*

We recorded mixed responses from participants as regards how they perceived the direction of gaze for the robot tutor.

> P018: [Robot condition] *"I got that the robot was helping, it was looking at the right card. Afterwards, I felt the movements of the head were slower."*
>
> P020: [Robot condition] *"I got the tutor help - it looked at my card, then me, and then to the matching card. I just got one problem - to read the gaze direction from the angle. The head movements were pretty well-paced, not too slow, not too fast."*
>
> P14: [Robot condition] *"I got that the tutor was helping, was looking at the right card, but it was not easy to tell which one it was looking at. The robot felt natural and the speed was nice. I felt like the robot was fully automatic."*
>
> P16: [Robot condition] *"Yes, I noticed the help, pretty natural, difficult to see the direction, I was not sure which card when the cards were closer."*

V. DISCUSSION & CONCLUSION

The present paper presents an experimental study designed to examine whether gaze cues from a human or a robot tutor can direct attention, and influence the choices of human partners in a card matching game. Specifically, users are asked to pick matching card pairs and have varying levels of assistance in locating the matching card. The conditions of assistance are Help and No_Help from a human tutor, and a robot tutor. The help involves the tutor looking at the matching card once the player has made a card selection. The paper details a study design on comparing task performance measured by time and number of tries of Help vs No_Help, and robot vs human Findings from the study support the formulated hypotheses:

In our first hypothesis (Hypothesis One), we projected that participants would perform better in the Help condition than in the No_Help condition. Number of tries supported this assumption. Participants used significantly fewer tries to find all the matching cards with help from the tutor than without help (F $(1, 18) =7.009$, p=0.016). However, we found no significant time difference between the two conditions. In our second hypothesis, (Hypothesis Two): The type of tutor (human/robot) will influence the participants' task performance both regarding time and numbers of tries. According to participants awareness of the tutor's hints, a significantly higher number of participants reported identifying the help cues and using the gaze information to pick the matching cards during the Robot-Help condition than in the Human-Help condition. Consequently, participants performed significantly better, measured by the number of tries,

with help from the robot tutor than from the human tutor. We found that participants identified all the pairs of matching cards with significantly fewer tries with help from the robot tutor (p=0.022) than without help. However, there was no significant difference in the number of tries, with or without help, in the human tutor condition (p=0.233).

Further analysis shows a significant correlation between noticing gaze hints and the number of tries. Participants who reported seeing the gaze hints in the robot condition performed significantly better, measured by the number of tries than those who did not report identifying the gaze cues. However, there was no significant difference in duration for those who identified gaze hints in the robot condition and those who did not report identifying the gaze hints. For the human condition, there was no significant difference between those who reported identifying the gaze hints and those who did not report identifying the gaze hints across both measures. However, the few participants who noticed the help in the human condition noticed it too late in the game to take advantage of it.

The significant difference between the number of participants who noticed the gaze hints of the human or the robot tutor might relate to diverse factors. First, the novelty effect of the robot on participants attention, supported by the fact that participants with the robot tutor spend more time to fulfil the task even when the number of tries was less. From the eye gaze, data results indicate that participants spent more time looking at the robot hence more time to complete the task. Another explanation could be that the sounds of the robot's motors displaying the head movements –to perform gaze behaviour- possibly attracted participants attention, making the robot's gaze behaviour more salient than human's. Besides, the robot's gaze behaviour is achieved with large head movements, whereas the human tutor's gazes are much subtle, based more on the eyes movements than in the head motion. According to the particular salience in the experimental setting, we can consider that robot's gaze behaviour is an overt cue while the humans are a covert cue, which affects its communicative effectiveness in the context of assisting the player. Though covert cues can influence interaction even without being aware of them, in this particular context and regarding the game flow, the hints providing information about the matching card position should be noticed by participants to be effective.

With this respect, our findings differ from the findings of Mutlu et al. [28] or in collaborative scenarios (Palinko et al. [31] who showed that gaze is a powerful communicative signal also without explicit awareness from the observer. In the current study, the lack of awareness was shown to cancel the informative content of gazing, which we believe is because of the lack of subtlety of the NAO eyes which does not have the impactful embodiment of human-like eyes. For the Human condition ,there is the concept of intimacy regulation, where humans control their gazes to regulate the level of intimacy with their interaction partners [4]. This may mean that in the human condition participants restricted the gaze behaviour according to social rules, which was not present in the robot case. Additionally the human condition, subjective evaluation indicated that a majority expected verbal help from the tutor, as the more natural modality of communication in the face to face situation, even more provided the human tutor addressed the participant verbally in the introduction of the activity.

Results do not show a significant effect on game duration, even during the robot condition. There are several possible explanations for this fact. Firstly, as soon as the participant noticed that the robot tutor was helping them with gaze hints, they waited until the robot showed them the matching card, even when they had an idea of where the matching card was. Secondly, as revealed in the interview responses, it took a while for some of the participants to read the gaze direction of the robot. The difficulty in reading gaze could be attributed to the limitations of the robot used as the experimental platform for this study, as it lacks articulated eyes. The third probable reason is the duration of head motions during attention shifts from the flipped card to the face of the participant, and then to the matching card. Again, as shown in the qualitative responses, some of the participants indicated that the head movements of the robot tutor were slow. In this sense, the robot's help could be considered detrimental regarding task performance only when measured by time: waiting for robots' hints increased the execution time while improved the accuracy of selections.

We recorded diverse responses from participants as regards how they perceived the direction, timing and intent of head movement (gaze) of the robot tutor. A few of the participants reported identifying the robot head movements but felt like it was following their moves rather than directing their attention to the matching card, missing the informative content of tutor's gaze. This brings up the issue of the timing of gaze behaviour and directions of head movements, which are two interesting aspects of our future work. Further analysis of the gaze behaviour of the participants collected during the experiment will reveal finer, micro-level details of the interactive player-tutor looking behaviour, including complex gaze patterns (i.e. gaze following, joint attention) to gain an understanding of the differences between human-robot and human-human nonverbal communication. Interesting coordinated sequences to investigate are, responsive gaze behaviours to either the robot or the human tutor head turn or gaze shift. Besides, the analysis of sequences of coordinated gaze behaviour could shed light on the participant's attributions of tutor's intent in the flow of the game. The attributed meaning of tutor behaviour could be inferred from the moment when participants check tutor gaze behaviour –look at the face-, for instance before selecting a card –looking for cues- or just immediately after a card selection –looking for confirmation.

Finally, we observed that participants in the robot condition appeared unperturbed, however, in the human condition, participants seemed a little uneasy. This observation, and especially the reasons for it can be investigated in future studies. One possible explanation is the unnaturalness of the human tutor interactive behaviour, shifting from a casual verbal based communication during the introduction and the briefing phase into an absence of verbal communication during the game in both Help and No-help conditions.

The use of a card game aims to add to our long-lasting research line on social inclusion which includes training social skills to children with ASD [5, 33] and improving the involvement in activities of elderly with dementia [32] which

was shown to be very beneficial for these user groups. In such settings often games are used to encourage the children to better engage in training practices [5, 23, 33], or to stimulate higher engagement in activities and thus longer mental alertness in elderly [32]. Huskens et all [33], and Barakova et al. [5] have incorporated game based training in clinical studies, while Palinko et al. [31] gives a plausible framework for introducing robots with designed gaze behaviors in general education. Perugia et al. [32] used cognitive games and robots to engage elderly by mental stimulation on an emotional level.

Future work involves examining the temporal aspects of gaze and movements in human-human interactions to build more realistic interactive robot gaze behaviours. To further this research, using human studies is a promising approach, as the robots are intended for interaction with people. In the future, we hope to combine robot's head-directions behaviours, different gaze movements, and other social cues derived from Time and Flow Effort of Laban Movement Analysis [23].Additionally, to improve our design and communication capabilities, we will combine NAO with a gaze tracker, to detect gaze directions of the human, and to control robot responses.

ACKNOWLEDGEMENT

REFERENCES

[1] Admoni, H., Hayes, B., Feil-Seifer, D., Ullman, D., & Scassellati, B. (2013, March). Are you looking at me? perception of robot attention is mediated by gaze type and group size. In Proceedings of the 8th ACM/IEEE international conference on human-robot interaction (pp. 389-396).

[2] André, E., & Chai, J. Y. (2010, February). Workshop: eye gaze in intelligent human machine interaction. In Proceedings of the 15th international conference on Intelligent user interfaces (pp. 431-432).

[3] Andrist, S., Tan, X. Z., Gleicher, M., & Mutlu, B. (2014, March). Conversational gaze aversion for humanlike robots. In Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction (pp. 25-32).

[4] Argyle, M., Ingham, R., Alkema, F., & McCallin, M. (1973). The different functions of gaze. Semiotica, 7(1), 19-32.

[5] Barakova, E. I., Bajracharya, P., Willemsen, M., Lourens, T., & Huskens, B. (2015). Long-term LEGO therapy with humanoid robot for children with ASD. Expert Systems, 32(6), 698-709.

[6] Barakova EI, Chonnaparamutt W: Timing sensory integration. IEEE Robotics & Automation Magazine 2009, 16:51-58.

[7] Baron-Cohen, S. (1997). Mindblindness: An essay on autism and theory of mind. MIT press.

[8] Boucher, J. D., Pattacini, U., Lelong, A., Bailly, G., Elisei, F., Fagel, S., ... & Ventre-Dominey, J. (2012). I reach faster when I see you look: gaze effects in human–human and human–robot face-to-face cooperation. Frontiers in neurorobotics, 6, 3.

[9] Bemelmans, R., Gelderblom, G. J., Jonker, P., & de Witte, L. Socially Assistive Robots in Elderly Care: A Systematic Review into Effects and Effectiveness. Journal of the American Medical Directors Association, 13(2), 114-120

[10] Breazeal, C., Kidd, C. D., Thomaz, A. L., Hoffman, G., & Berlin, M. (2005, August). Effects of nonverbal communication on efficiency and robustness in human-robot teamwork. In 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (pp. 708-713).

[11] Brooks, R., & Meltzoff, A. N. (2002). The importance of eyes: how infants interpret adult looking behavior. Developmental psychology, 38(6), 958.

[12] Brooks, R., & Meltzoff, A. N. (2014). Gaze following: a mechanism for building social connections between infants and adults. In Mechanisms of social connection: from brain to group, 167-183.

[13] Broz, F., Lehmann, H., Nakano, Y., & Mutlu, B. (2012, March). Gaze in HRI: from modeling to communication. In Proceedings of the seventh annual ACM/IEEE international conference on Human-Robot Interaction (pp. 491-492).

[14] Cuijpers, R. H., & van der Pol, D. (2013, October). Region of eye contact of humanoid Nao robot is similar to that of a human. In International Conference on Social Robotics (pp. 280-289). Springer International Publishing.

[15] Delaunay, F., de Greeff, J., & Belpaeme, T. Keeping an eye on you: The influence of robotic faces on gaze direction reading.

[16] Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. Neuroscience & Biobehavioral Reviews, 24(6), 581-604.

[17] Feldman, R. S., & Rimé, B. (1991). Fundamentals of nonverbal behavior. Cambridge University Press.

[18] Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. Psychological bulletin, 133(4), 694.

[19] Imai, M., Kanda, T., Ono, T., Ishiguro, H., & Mase, K. (2002). Robot mediated round table: Analysis of the effect of robot's gaze. In Robot and Human Interactive Communication, 2002. Proceedings. 11th IEEE International Workshop on (pp. 411-416).

[20] Kendon, A. (1967). Some functions of gaze-direction in social interaction. Acta psychologica, 26, 22-63.

[21] Kirchner, N., Alempijevic, A., & Dissanayake, G. (2011, March). Nonverbal robot-group interaction using an imitated gaze cue. In Proceedings of the 6th international conference on Human-robot interaction (pp. 497-504).

[22] Kleinke, C. L. (1986). Gaze and eye contact: a research review. Psychological bulletin, 100(1), 78.

[23] Kozima, H., Nakagawa, C., & Yasuda, Y. (2005, August). Interactive robots for communication-care: A case-study in autism therapy. In ROMAN 2005. IEEE International Workshop on Robot and Human Interactive Communication, 2005. (pp. 341-346).

[24] Lourens T, van Berkel R, Barakova E: Communicating emotions and mental states to robots in a real time parallel framework using Laban movement analysis. Robotics and Autonomous Systems 2010, 58:1256-1265.

[25] Meltzoff, A. N., Brooks, R., Shon, A. P., & Rao, R. P. (2010). "Social" robots are psychological agents for infants: A test of gaze following. Neural Networks, 23(8), 966-972.

[26] Moon, Aj. et al., 2014. Meet me where i'm gazing. Proceedings of the 2014 ACM/IEEE international conference on Human-robot interaction - HRI '14, pp.334–341.

[27] Mutlu, B., Forlizzi, J., & Hodgins, J. (2006, December). A storytelling robot: Modeling and evaluation of human-like gaze behavior. In 2006 6th IEEE-RAS International Conference on Humanoid Robots (pp. 518-523).

[28] Mutlu, B., Yamaoka, F., Kanda, T., Ishiguro, H., & Hagita, N. (2009, March). Nonverbal leakage in robots:

communication of intentions through seemingly unintentional behavior. In Proceedings of the 4th ACM/IEEE international conference on Human robot interaction (pp. 69-76).

[29] Mwangi, E., Barakova, E., Zhang, R., Diaz, M., Catala, A., & Rauterberg, M. (2016, October). See Where I am Looking at: Perceiving Gaze Cues With a NAO Robot. In Proceedings of the Fourth International Conference on Human Agent Interaction (pp. 329-332).

[30] Palinko, O., Sciutti, A., Schillingmann, L., Rea, F., Nagai, Y., & Sandini, G. (2015, August). Gaze contingency in turn-taking for human robot interaction: Advantages and drawbacks. In Robot and Human Interactive Communication (RO-MAN), 2015 24th IEEE International Symposium on (pp. 369-374).

[31] Palinko, O., Rea, F., Sandini, G., & Sciutti, A. (2016, 9-14 Oct. 2016). Robot reading human gaze: Why eye tracking is better than head tracking for human-robot collaboration. Paper presented at the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS).

[32] Perugia, G., Díaz Boladeras, , M., Mallofré, A. C., Rauterberg, M., & Barakova, E. (2017, 17-20 July 2017). Modelling engagement in dementia through behaviour. Contribution for socially interactive robotics. Paper presented at the 2017 International Conference on Rehabilitation Robotics (ICORR).

[33] .Huskens B, Verschuur R, Gillesen J, Didden R, Barakova E (2013) Promoting question-asking in school-aged children with autism spectrum disorders: effectiveness of a robot intervention compared to a human–trainer intervention. Dev Neurorehabil 16:345–356

[34] Pfeiffer-Lessmann, N., Pfeiffer, T., & Wachsmuth, I. (2012). An operational model of joint attention-timing of gaze patterns in interactions between humans and a virtual human. In Proceedings of the 34th Annual Conference of the Cognitive Science Society (CogSci 2012), N. Miyake, D. Peebles, and RP Cooper, Eds. Austin, TX USA: Cognitive Science Society (pp. 851-856).

[35] Poggi, I., Pelachaud, C., & De Rosis, F. (2000). Eye communication in a conversational 3D synthetic agent. AI communications, 13(3), 169-181.

[36] Robotics, A. (2010). Nao robot reference manual. Version 1.10. 10.

[37] Ruhland, K., Andrist, S., Badler, J., Peters, C., Badler, N., Gleicher, M., ... & Mcdonnell, R. (2014, April). Look me in the eyes: A survey of eye and gaze animation for virtual agents and artificial systems. In Eurographics State-of-the-Art Report (pp. 69-91).

[38] Sidner, C. L., Lee, C., Morency, L. P., & Forlines, C. (2006, March). The effect of head-nod recognition in human-robot conversation. In Proceedings of the 1st ACM SIGCHI/SIGART conference on Human-robot interaction (pp. 290-296). ACM.

[39] Vertegaal, R., Slagter, R., Van der Veer, G., & Nijholt, A. (2001, March). Eye gaze patterns in conversations: there is more to conversational agents than meets the eyes. In Proceedings of the SIGCHI conference on Human factors in computing systems (pp. 301-308).

[40] Yamazaki, A., Yamazaki, K., Kuno, Y., Burdelski, M., Kawashima, M., & Kuzuoka, H. (2008, April). Precision timing in human-robot interaction: coordination of head movement and utterance. In Proceedings of the SIGCHI C onference on Human Factors in Computing Systems (pp. 131-140).

[41] Yoshikawa, Y., Shinozawa, K., Ishiguro, H., Hagita, N., & Miyamoto, T. (2006, August). Responsive Robot Gaze to Interaction Partner. In Robotics: Science and systems

[42] Yu, C., Schermerhorn, P., & Scheutz, M. (2012). Adaptive eye gaze patterns in interactions with human and artificial agents. ACM Transactions on Interactive Intelligent Systems (TiiS), 1(2), 13.

[43] Wiese E, Müller HJ, Wykowska A: Using a Gaze-Cueing Paradigm to Examine Social Cognitive Mechanisms of Individuals with Autism Observing Robot and Human Faces. In Social Robotics: 6th International Conference, ICSR 2014, (pp.370-379).