

## Human Communicative Responses to Different Modes of Gaze Management by the Robot

**Zinina A. A.**

Kurchatov Institute / Moscow, Russia  
Russian State University for the  
Humanities / Moscow, Russia  
Moscow State Linguistic University /  
Moscow, Russia  
zinina\_aa@nrcki.ru

**Kotov A. A.**

Kurchatov Institute / Moscow, Russia  
Russian State University for the  
Humanities / Moscow, Russia  
Moscow State Linguistic University /  
Moscow, Russia  
kotov@harpia.ru

**Zaidelman L. Y.**

Kurchatov Institute / Moscow, Russia  
Russian State University for the  
Humanities / Moscow, Russia  
zaydelman\_ly@nrcki.ru

**Arinkin N. A.**

Kurchatov Institute / Moscow, Russia  
Russian State University for the  
Humanities / Moscow, Russia  
arinkin\_na@nrcki.ru

### Abstract

We investigated communicative reactions of people ( $N = 46$ ), while telling stories to two companion robots, who reacted differently to the human gaze (head turning). In response to a human gaze the "aversive" robot averted its gaze away from the user, while the "responsive" robot, lifted its head and showed a responsive gaze. We found that users with high level of emotional intelligence prefer the gaze responsive robot and better recognize the difference between the robots. Thus, these users constitute the core group for the technology. In this paper, we further examine behavioral patterns of people in the experiment situation: (a) shift of attention to the story; (b) shift of attention to the robot; (c) joint attention. We also distinguish the communicative reactions of people, mainly from the core group, to the aversive and responsive gazes of the robots: positive responses to gaze contact and negative responses to gaze aversion. We show that for some users the responsive gaze behavior of the robot may serve as positive feedback, increasing the number of human iconic gestures, while telling a story to the responsive robot, and decreasing the number of iconic gestures in a story to the aversive robot.

**Keywords:** Multimodal communication, human computer interaction, emotional computing, companion robots, gaze contact, eyes contact

**DOI:** 10.28995/2075-7182-2022-21-594-602

### 1. Introduction

Human gaze has multiple functions and is controlled by different systems, first of all, by attention. On the one hand, a human turns his eyes to an object that attracts his attention and tries to orient himself by switching his gaze between objects in the environment [17; 20]. At the same time, eye movements can support both voluntary (endogenous) attention, caused by internal stimuli, and involuntary (exogenous) attention, imposed by the external situation [19]. On the other hand, gaze is an expressive mean: with the help of gaze a subject has the possibility to intentionally or spontaneously indicate to the interlocutor the object of his attention, to direct the gaze of the interlocutor – to participate in the situation of joint attention, or to demonstrate meaningful mimic patterns: to turn eyes, to *shoot* with eyes, thus, informing the interlocutor a negative or positive evaluation. The semiotic functions of nonverbal actions, including the direction of gaze, were considered in the classic work by Adam Kendon [10]. The connection between attention and the communicative functions of the gaze is studied by O.A. Fedorova, who with the help of

modern methods (using eyeglasses-trackers) describes the strategies of distribution of visual attention of participants of natural communication: *general*, *context-dependent* and *individual* [5]. The competition of the cognitive systems for eye control has not only a perceptual but also an important communicative function [1; 18; 16]. Eye movements are an iconic sign indicating the communicative intentions or states of the subject. At the same time, this sign is extremely ambiguous: a person can interpret a side-gaze of an interlocutor (a) as *thoughtfulness* or *thinking* or (b) as *inattention* or attention shift to another object [11; 12]. The eyelid and eyebrow system, controlled by muscles or action units [2] can also directly serve the attention mechanisms (such as squinting or opening eyes when looking at an object), or express cognitive and emotional states (such as opening eyes and raising eyebrows to express a surprise).

Coordinated simultaneous gaze at an object is described as the effect of *joint attention* [6], where the attention as a cognitive function is shared between two participants in communication [4]. Joint attention is one of the basic mechanisms of child's development in the aspects of speech and social skills. As a cognitive function, it is much more than simultaneous gazing at an object – it involves shared intentional attitude toward the world [9] and includes the skills of (a) detecting the attention of another person, (b) attention management, (c) social coordination, and (d) theory of mind. Modeling of joint attention is an extremely promising area for effective human-machine interaction, and is widely investigated in real communications with robotic and virtual agents [3]. The use of artificial agents provides greater ecological validity as compared to classical protocols using a screen with pictures or schematic eyes – the gaze-cueing paradigm, for example, as in [4; 8].

We use the situation of interaction with the experimental robot companion F-2 to position a person in a certain pattern of communicative attention distribution by the robot interlocutor. The distribution of the person's attention, his communicative reactions, and a description of his experience after the experiment can serve to evaluate the person's response to different modes of gaze control by the robot.

## 2. Robotic responsive gaze experiment

We conducted an experiment in which we investigated whether different robot's responsive gaze could, on the one hand, change the attractiveness of the robot to the user and, on the other hand, cause certain behavior in users. We expected, that people with high emotional intelligence would be more sensitive to the robot with responsive gaze – they would prefer this robot and may demonstrate the responsive behavior in gestures and gaze, as compared to their interaction with gaze aversive robot. Within the experiment, subjects had to tell two robots 6 stories following a list of pictures by Herluf Bidstrup (Fig. 1). Each picture was represented as a stack of cards in random order. Before the experiment, the F-2 robots were introduced to the subjects. The main purpose of the study – to investigate the effect of the robot's responsive gaze – was not reported to the respondents. They were told that the developers were trying to train the robot to follow the story narrated by a human. After the experiment each participant filled out a questionnaire, describing his/her experience with the robots. We also used the *EmIn* – Emotional Intelligence Test [13; 14] to evaluate the level of emotional intelligence of the participants. The test consists of 46 questions; answers are scored and contribute to 4 scales, where the most important scale, following the hypothesis of the study, is *Understanding of Emotions of Others*.

### 2.1. Experimental conditions

The robots responded to human communication in two different ways. When a person looked away from the robot (e.g., looked aside or looked at the table to follow the cards), each of the robots looked down – at the table with the cards. When a person looked at the robot, the first robot (marked with a square) looked back: raised its head, opened its eyelids, and raised its eyebrows, while the second robot (marked with a triangle) demonstrated side gaze, looking left or right in random order (Table 1).

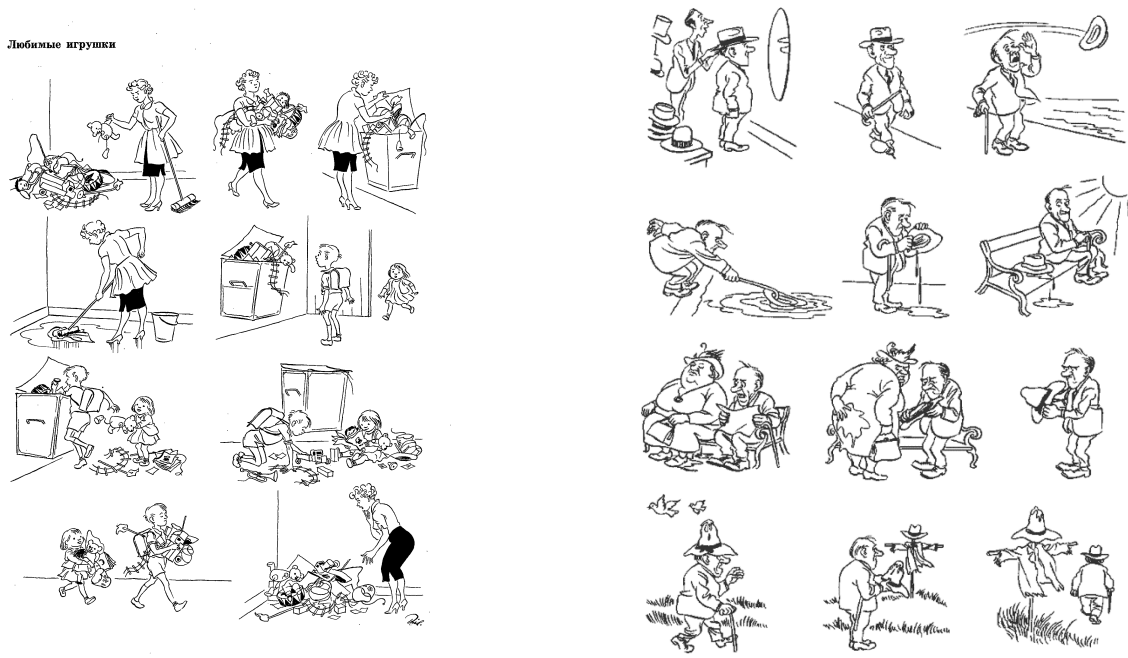


Figure 1: Example of stimulus material – a graphic story by Herluf Bidstrup. Each story was represented as a stack of separate cards in random order.

	Before or after the user's gaze	During the user's gaze
<p><b>Gaze-responsive robot</b> (marked with a square)</p>		
<p><b>Gaze-aversive robot</b> (marked with a triangle)</p>		<p>Left- or right-side gazes are selected randomly</p>

Table 1. Robots' reactions to human gaze: gaze responsive and gaze aversive behavior

## 2.2. Participants

A total of 46 subjects participated in the experiment (mean age of 27 years, 33 females). Three main attention zones were suggested in the experiment settings: participants could place cards on the table or show them to the left or right robot. Subjects could communicate with the left and right robot from the same position at the table, so the robots constantly maintained their respective behavior (Fig. 2).

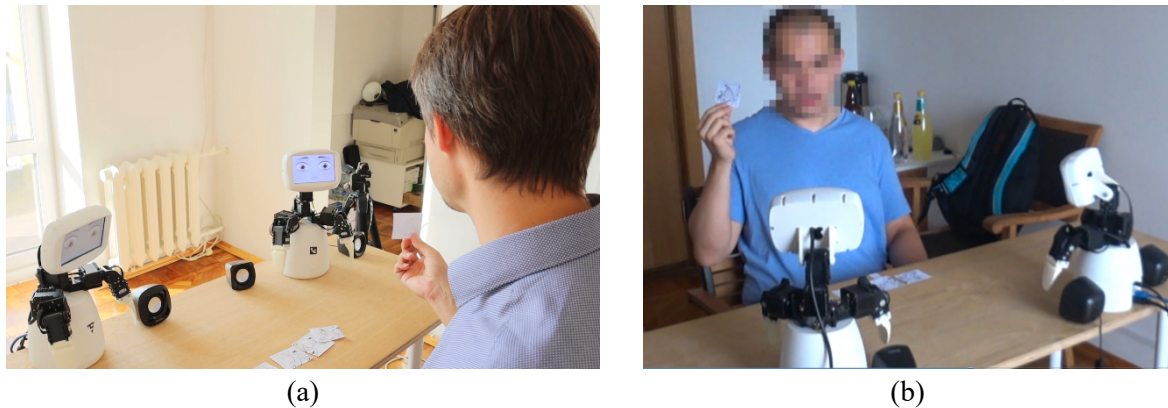


Figure 2: Experimental situation: (a) from the perspective of a participant and (b) front view

## 2.3. Robot movement and gaze recognition system

The robot control system controlled the robot's movement during the entire experiment. During greetings and speech utterances, the robot's head and arms were controlled by gestures coordinated with its speech production. During listening to a story, the robot's arms were controlled by the inactivity imitation component: random automanipulations or imitations of breathing have been performed. The vector of the user's attention was automatically identified by the orientation of the user's face and was recognized by a specially developed computer vision component based on OpenCV. This system had been preferred to an eye-tracker as a possible "built-in" solution for emotional companion robots, allowing us to avoid any calibration procedure, and thus, maintaining more natural communication. When the system had been recognizing a change in the user face vector, it had been translating messages like 'person looks\_at egocentric(robot)' or 'person looks away'. The first message caused the "responsive" robot to look back at the user and the "aversive" robot – to look away. When the user's gaze was moved away from a robot, this robot showed a downward gaze on the table.

## 3. Results

We found that subjects' emotional intelligence influences their preference of the robot, as well as their ability to recognize the difference in their gaze behaviors. Participants with high scores on the *Understanding of Emotions of Others* scale are better at recognizing the difference of gaze patterns demonstrated by the two robots ( $p < 0.05$ , Mann-Whitney U-test). In the group that noticed the difference, the level of emotional intelligence was 47.8 (SD 7.8), in the group that did not notice the difference, the level of emotional intelligence was 38.7 (SD 10.4). The number of people who have correctly identified the difference between two robots was: 8% (1 of 12 persons) among people, who have preferred the gaze avoiding robot, 23% (5 of 21 persons) among people, who have evaluated the robots equally, and 69% (9 of 13 persons) among people, who have preferred the gaze responsive robot. It means that people, recognizing the difference between gaze patterns of the robots significantly often ( $p < 0.01$ , Spearman correlation) prefer the robot with gaze responsive behavior. The results in more

details are described in [21]. In our further studies of people communicative behavior, we include all the participants, but consider as the core group people with high emotional intelligence, able to distinguish robot's gaze patterns. As we expect, these people can become the core user group for future companion robots.

### 3.1. Human communicative reactions to robot behavior

The vast majority of the subjects (63.0%,  $n = 29$ ) reported that they noticed a difference between the robots, however, only a third (32.6%,  $n = 15$ ) were able to describe this difference accurately – as a difference in the direction of the robots' gaze. Some subjects did notice the difference in the robots' movements, but interpreted the robots' actions differently. For example, one of the subjects noted that the gaze avoiding robot was *following the card with his gaze* – this could happen when the subject looked at the robot and showed the picture (moved it aside of his body), and the robot reacted to the gaze and moved its gaze sideways, as if following the picture. Another subject interpreted the robot's head-down action as an appropriate reaction of the robot to the sad events in a story. While these actions constitute some exceptions, we can identify some more general patterns human gaze distribution when interacting with robots with different types of gaze behavior.

(1) **Shift of attention to the story (side-gaze).** Most subjects (67.3%,  $n = 31$ ) could not accurately describe the difference between the robots. Many subjects concentrated on the subtask *construct a story from the cards*, they told stories without showing any gestures and only occasionally looking at the robots (Fig. 3). In turn, the automatic system could not register a sufficient number of incoming gazes and the number of robot's reactions and differences between the robots decreased accordingly: both robots looked down in front of them. This communicative behavior could be caused by several reasons: (a) the subjects were drawing parallels to the typical *construct a story from the cards* school task and believed that their creativity would be evaluated, (b) the subjects were involved in the events of the story and showed an up/side gaze, characteristic to thoughtfulness [11; 12], or (c) the subjects did not establish communicative contact with the robot due to its lack of communicative behavior.

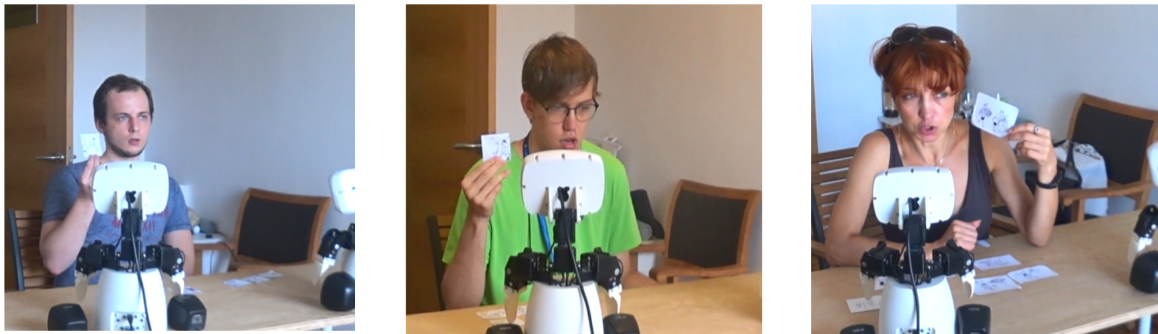


Figure 3: Participants look forward/upward, aside from the robot while telling the story

(2) **Shifting attention to the robot – following the direction of the robot's attention.** During the planning of the experiment, the responsive robot was designed as the more natural version, and the avoiding robot was designed with a kind of communicative deficit. At the same time, 26.0% of respondents ( $n = 12$ ) indicated that they liked the avoiding robot more, and in particular, that they liked *telling the story* to the avoiding robot (15.2%,  $n = 7$ ), because, for example, *the reactions of the avoiding robot were more relevant*. The deficit in the control of the avoiding robot's gaze had been provoking diverse communicative reactions by subjects. When the subjects were showing the card to the robot, and the robot was responding with an avoiding gaze, many subjects moved the card following the robot's gaze, as if trying to adapt to the direction of the robot's attention. This interaction with the supposed direction of the robot's attention was characteristic both for subjects who have noticed and described

the difference between the robots (Fig. 4) and for subjects who did not describe the difference (Fig. 5). This allows us to assume that the interpretation of the direction of the interlocutor's attention may be implicit: it triggers the subject's behavioral response, related to handling of other's attention, but it is not noticed by the subject himself and not represented in the report.

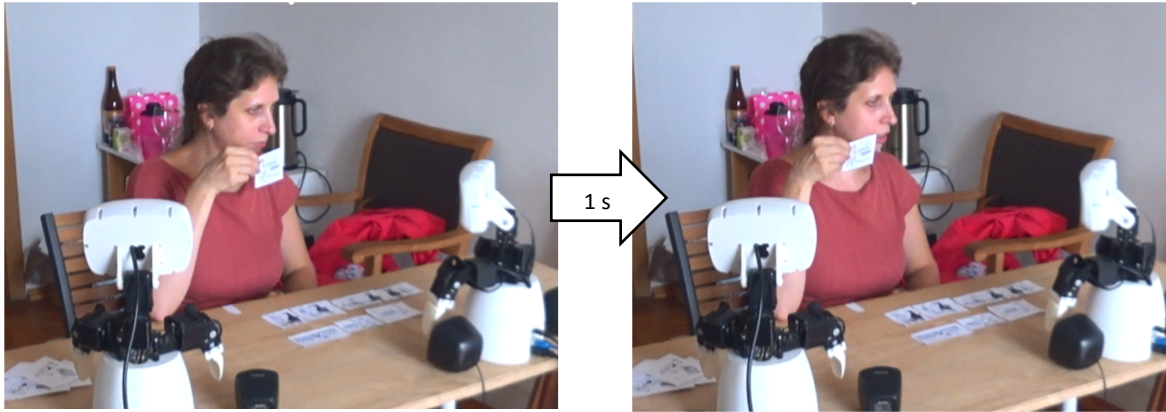


Figure 4: The participant places the card in front of the robot head and, after the robot replies with the avoiding gaze, moves the card in the direction of its gaze. In this case, the participant accurately described the difference between the robots in her report.



Figure 5: The participant places the card in front of the robot head and, after the robot replies with the avoiding gaze, moves the card in the direction of its gaze. In this case, the participant was unable to describe the difference between the robots in her report.

**(c) Joint Attention.** Many subjects behaved as if trying to attract the attention of the robot through the mechanism of joint attention. As they were showing the card toward the robot, they simultaneously moved their bodies forward and looked at the card, as if trying to gain additional attention from the robot to the picture (Fig. 6).

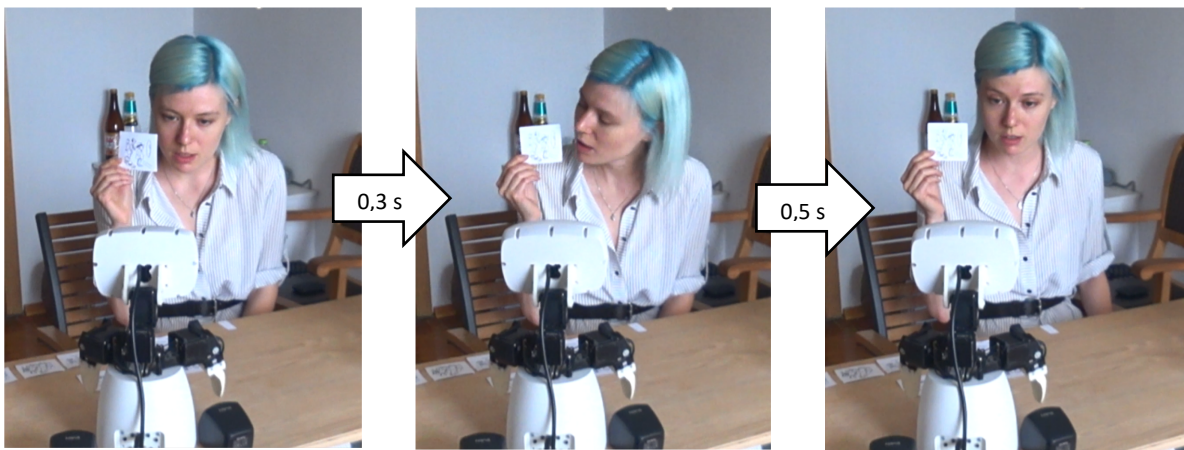


Figure 6: The respondent looks at the card facing the robot as if trying to get the robot's extra attention to the card

This mechanism, however, can be confused with another pattern, when subjects try to control the story and frequently peep into the card to follow the events. In several subjects this behavioral pattern is demonstrated at the first several stories, and is reduced to the end of the experiment, when they adapt to the experimental procedure.

### 3.2. Analysis of the behavior of the core group

As the core group for the communication behavior analysis, we selected subjects who noticed a difference in robots' behavior (32.6%,  $n = 15$ ), and who have also been characterized by a higher level of emotional intellect (*Understanding of Emotions of Others* score of EmIn test). These subjects showed different types of communicative responses when interacting with the robots.

**(a) No communicative response.** These subjects well controlled their communicative behavior, narrated stories while steady holding the card and showed no communicative gestures. In this case no difference between the subjects was observed during the interaction with different robots. This pattern is characteristic for subjects, who have demonstrated numerous auto-manipulations during the preparation of the story (when the cards are on the table but no communication with the robot is established yet), but not on the stage of storytelling.

**(b) Positive response to the robot's actions.** The subjects in this group accompanied their stories with expressive (mostly, iconic) gestures. The number of such gestures decreased over the course of three stories to the avoiding robot, and increased over the course of three stories for the responsive robot (Fig. 7, 8).

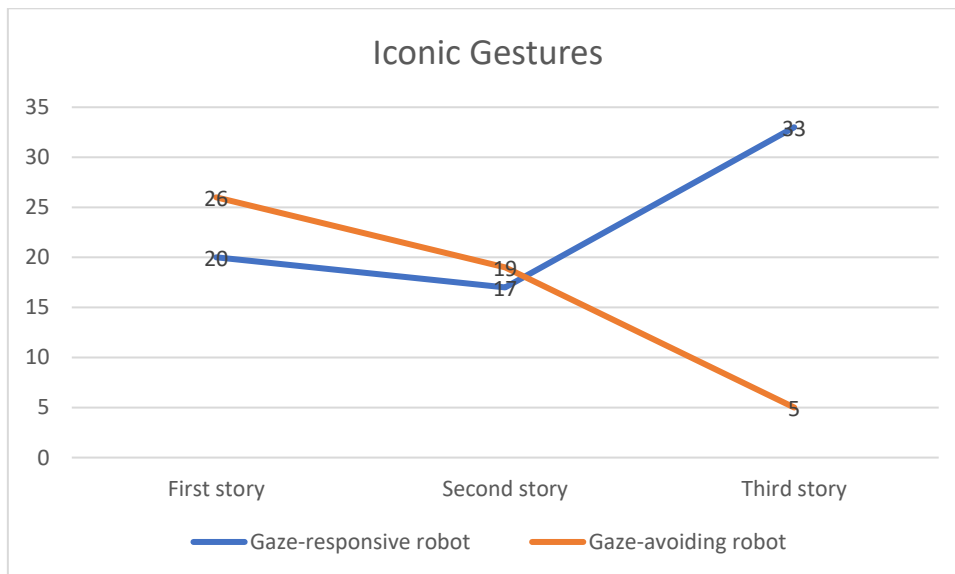


Figure 7: The number of iconic gestures in the stories – Participant №40 (female, 31)

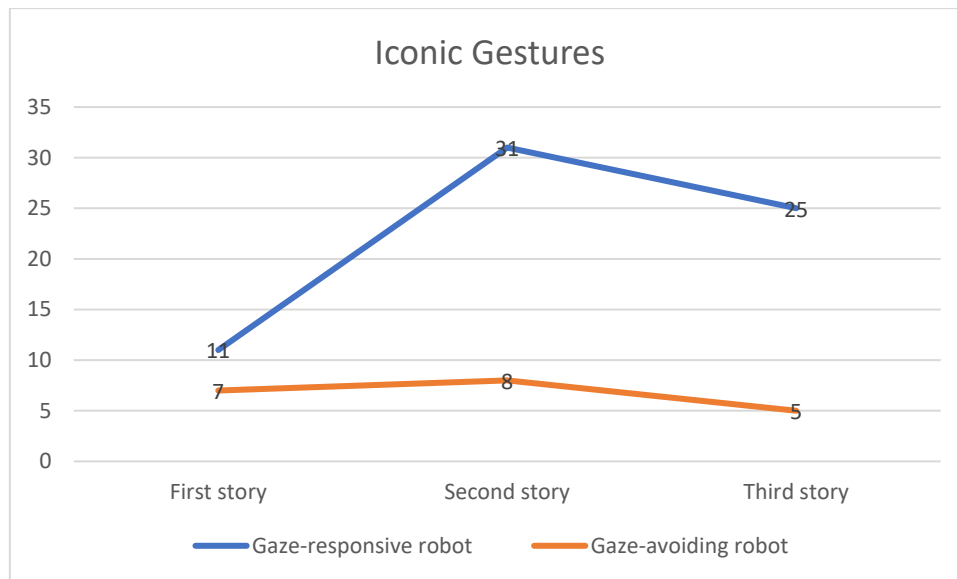


Figure. 8: The number of iconic gestures in the stories – Participant №24 (female, 30)

This behavior allows us to claim that some subjects receive communicative confirmation of attention from the responsive robot and use iconic gestures when telling him a story, while the lack of attention feedback from the avoiding robot causes them to reduce the number of gestures when telling him a story.

**(c) Negative response to robot' deficit.** Subjects tell the story neutrally to the responsive robot, but react negatively to the lack of attention from the avoiding robot. For example, subject №41 (female, 33) tries to get the attention of the avoiding robot before the beginning of each task: (1) she makes circular movements with the card and says *Are you listening to me?*, (2) while showing the card, she asks *Are you ready to listen to me? Are you looking at me?* and (3) while moving her gaze to the avoiding robot: *And you, are you ready to listen to my story?* Subject №35 (male, 35) tries to address the experimenter's assistant: *I don't think he's looking at me!*, and then – *Should I hold the card on the robot's right side then or on the left too? He [avoiding robot] keeps looking to the right.* Participant №38 (female, 22) tries to get feedback from the robot after the story – she holds up the card and says *Look!*, expecting a responsive gaze (or other feedback) from the robot. These reactions indicate expressed frustration during the participants' interaction with the robot with communicative deficit.

#### 4. Conclusion

Interactive gaze management for emotional companion robots constitutes a promising feature. While only a minority of people (32.6%,  $n = 15$ ) recognize the difference in robot's gaze patterns, these people also prefer the gaze responsive robot and are characterized by higher level of emotional intelligence. They may constitute the core user group for future emotional companion robots. These people may adjust their behavior in communication with the two robots: negatively reacting to the deficit of the gaze avoiding robot and/or positively reacting to the gaze responsive robot, in particular, by increasing the number of iconic gestures in the communication. People may try to adjust to the presumed attention of the robot by moving the card (object of the talk) after the robot's gaze – this may apply even to the people, who do not report the difference between the robot gaze patterns. People may also try to manipulate the attention of the robot through the mechanisms of joint attention – while showing a card and simultaneously looking at it. These results show that people (mostly with high level of emotional intelligence) successfully recognize natural gaze behavior patterns of companion robots and may demonstrate natural responsive cues – in gestures and joint attention. Thus, the robot may be perceived as a natural companion partner in the aspect of compound gaze behavior.



## Acknowledgements

The research was in part supported by the grant of the Russian Science Foundation No. 19-18-00547, <https://rscf.ru/project/19-18-00547/>

## References

- [1] Beyan C., Murino V., Venture G., Wykowska A. Computational Approaches for Human-Human and Human-Robot Social Interactions // *Frontiers in Robotics and AI*. — 2020. — Vol. 7.
- [2] Ekman P., Friesen W. Facial Action Coding System: A Technique for the Measurement of Facial Movement. — Palo Alto: Consulting Psychologists, 1978 — Vol. 3(2), pp. 5.
- [3] Chevalier P., Kompatsiri K., Ciardo F., Wykowska A. Examining joint attention with the use of humanoid robots-A new approach to study fundamental mechanisms of social cognition // *Psychonomic Bulletin & Review*. — 2020. — Vol. 27(2) — P. 217-236.
- [4] Falikman M.V. (2006) General psychology. In seven volumes: a student's course book for higher education institutes — Volume 4. Attention [Obshchaya psihologiya. V 7 t.: uchebnik dlya stud. vyssh. ucheb, zavedenij]. — M.: Academy.
- [5] Fedorova O.V. (2017), Distribution of visual attention of interlocutors in natural communication: 50 years later [Raspredelenie zritel'nogo vnimaniya sobesednikov v estestvennoj kommunikacii: 50 let spustya], *Cognitive Science in Moscow: New Research. Proceedings of the conference*, pp. 370-375.
- [6] Frischen A., Bayliss A. P., Tipper S. P. Gaze cueing of attention // *Psychological Bulletin*. — 2007. — Vol. 133. — P. 694-724.
- [7] Iriskhanova O.K., Cienki A. The semiotics of gestures in cognitive linguistics: Contribution and challenges // *Voprosy Kognitivnoy Lingvistiki*. — 2018. — Vol. 4. — P. 25-36.
- [8] Jording M., Hartz A., Bente G., Schulte-Rüther M., Vogeley K. The “social gaze space”: A taxonomy for gaze-based communication in triadic interactions // *Frontiers in Psychology*. — 2018. — Vol. 9.
- [9] Kaplan F., Hafner V. V. The challenges of joint attention // *Interaction Studies*. — 2006. — Vol. 7(2). — P. 135-169.
- [10] Kendon A. Some functions of gaze direction in social interaction // *Acta Psychologica*. — 1967. — V. 26. — P. 22-63.
- [11] Kotov A. A., Zinina A. A. (2015), Functional analysis of nonverbal communicative behavior [Funkcional'nyj analiz neverbal'nogogo komunikativnogo povedeniya], *Computational Linguistics and Intellectual Technologies [Komp'yuternaya Lingvistika i Intellektual'nye Tekhnologii]*, M., V. 1(14), pp. 299-310.
- [12] Kotov A., Zinina A., Filatov A. Semantic Parser for Sentiment Analysis and the Emotional Computer Agents // *Proceedings of the AINL-ISMW FRUCT*. — 2015. — P. 167-170.
- [13] Lyusin, D.V. (2006) A New Technique for Measuring Emotional Intelligence: The EmIn Questionnaire [Novaya metodika dlya izmereniya emocional'nogo intellekta: oprosnik EmIn], *Psychol. Diagn.*, Vol. 4, pp. 3-22.
- [14] Lyusin D.V. (2009) EMIN Emotional Intelligence Questionnaire: New psychometric data [Oprosnik na emocional'nyj intellekt EmIn: novye psihometricheskie dannye], *Social and Emotional Intelligence: From Models to Measurements*, M., pp. 264-278.
- [15] Müller C. Gesture and sign: Cataclysmic break or dynamic relations? // *Frontiers in Psychology*. — 2018. — Vol. 9. — P. 1651.
- [16] Pagnotta M., Laland K.N., Coco M.I. Attentional coordination in demonstrator-observer dyads facilitates learning and predicts performance in a novel manual task // *Cognition*. — 2020. — P. 201.
- [17] Pannasch S., Schulz J., Velichkovsky B.M. On the control of visual fixation durations in free viewing of complex images // *Attention, Perception, and Psychophysics*. — 2011. — V. 73(4). — P. 1120-1132.
- [18] Velichkovsky B.M. Communicating attention: Gaze position transfer in cooperative problem solving // *Pragmatics & Cognition*. — 1995. — V. 3(2). — P. 199-223.
- [19] Velichkovsky B. M. (2006) *Cognitive Science: Fundamentals of the Psychology of Cognition [Kognitivnaya nauka: Osnovy psihologii poznaniya]*, M.: Akademia, Vol. 1.
- [20] Velichkovsky B. M., Korosteleva A. N., Pannasch S., Helmert J. R., Orlov V. A., Sharaev M. G., Velichkovsky B.B., Ushakov V. L. Two visual systems and their eye movements: A fixation-based event-related experiment with ultrafast fMRI reconciles competing views // *STM*. — 2019. — V. 11(4) — P. 7-16.
- [21] Velichkovsky B. M., Kotov A., Arinkin N., Zaidelman L., Zinina A., Kivva K. From Social Gaze to Indirect Speech Constructions: How to Induce Impression that Your Companion Robot is a Conscious Creature // *Appl. Sci*. — 2021. — V. 11(21). — P. 10255.