Social Reflex Hypothesis on Blinking Interaction

Yuichiro Yoshikawa (yoshikawa@jeap.org)
Asada Synargistic Intelligence Project, Erato, JST
2-1 Yamadaoka, Suita, Osaka, 565-0871

Kazuhiko Shinozawa (shino@atr.jp)
2-2-2 Hikaridai, Keihana Science city, Kyoto, Japan

Hiroshi Ishiguro (ishiguro@ams.eng.osaka-u.ac.jp)
Asada Synargistic Intelligence Project, Erato, JST
Dept. of Adaptive Machine Systems, Graduate Shool of Engineering, Osaka Univ.
2-1 Yamadaoka, Suita, Osaka, 565-0871 Japan

Abstract

An interactive artificial agent is supposed to be a feasible tool to reveal how humans recognize and respond to another's response and should be developed through the investigation of the human cognitive mechanism. Previous studies have suggested that humans are sensitive to the responses of the interactant and might relate them to the interactant's communicativeness. Following on from the findings of the previous studies, this paper presents the *social reflex hypothesis* on nonverbal responsive interaction, which refers to the social effect and origins of one's unconscious response to the nonverbal behaviors of the interactant. To investigate the hypothesis, we conducted an experiment evaluating participants' impression on an on-screen agent that could blink in response to the participant's blinking. We found there was a non-linear relationship between the response latencies and the participants' *feeling of being looked at*, which accorded with the hypothesis. The implications of the result and the further work on other aspects of the hypothesis were discussed.

Keywords: Social reflex hypothesis; response latency; feeling of being looked at; interactive agent.

Introduction

An interactive artificial agent, including a communication robot and on-screen agents have been widely focused on as potential devices for an intuitive interface for humans (Kanda, Ishiguro, Imai, & Ono, 2004) and a therapeutic tool for communication disorders (Robins, Dickerson, Stribling, & Dautenhahn, 2004). To design an agent that responds to humans, it is necessary to know how humans recognize and respond to an agent's response. However, it is difficult to control the experimental conditions for revealing human cognitive mechanisms to the interactant since we can not completely control confederate humans. For example, it is difficult to inhibit and promote one's unconscious responses, which might form important aspects of how we recognize the other. Therefore, interactive artificial agents have been focused on as controllable "humans" in the experiment to complementarily approach the cognitive mechanism of a communication partner.

Experiments with artificial agents or objects without animate appearances have been extensively conducted to reveal how infants regard someone or something as being with important attributes for a communicative existence. From the experiment to compare an infant's re-enactment of the demonstration by a human and a mechanical manipulator, an animate appearance has been suggested to be a necessary aspect for infants to attribute intentions to the other (Melzoff, 1995). On the other hand, experiments using non-animate agents without animate appearance but with some behavioral aspects have been conducted to reveal the effect of these aspects. Through the experiments, behavioral aspects such as rationality (Gergely & Csibra, 2003), self-propelledness (Luo & Baillargeon, 2005), and interactiveness or contingency (Shimizu & Johnson, 2004) have been suggested to also be necessary for infants to attribute goals or intentions to the agents. Experiments using an agent with more animate appearance, such as a humanoid robot, have revealed that infants expect communicability of an interactive robot (Arita, Hiraki, Kanda, & Ishiguro, 2005). Furthermore, such studies using an agent with both an animate appearance and controllable, behavioral aspects have been expected to reveal their synergistic effects (Johnson, Slaughter, & Carey, 1998; Kamewari, Kato, Kanda, Ishiguro, & Hiraki, 2005).

However, in previous experiments, the participants usually observe the agents only from the third person's viewpoint. In other words, there have still been few studies directly analyzing how we recognize others through actively interacting with them. This might be caused by difficulties in building a sufficiently interactive agent. However, as appeared in a study using an android (Ishiguro & Minato, 2005), the recent development of technologies allows us to provide an interactive agent with an anthropomorphic appearance and powerful capabilities of sensing subtle human behaviors, with which it is expected to be sufficiently interactive to induce a natural response from humans. In other words, now we have artificial, controllable "humans" with which we can design experiments on the issue of the nature of the human cognitive mechanism necessary for an interactant's response.

Some previous studies with interactive robots have demonstrated that humans are sensitive to the interactant's response. Watanabe et al. have suggested that a responsive robot's nodding to a participant's voice could lead her to engage in conversation with it (Watanabe, Danbara, & Okubo, 2003). Not only the responses to voice but also to nonverbal signals such as mimicry of the partner's movement have been shown to make participants regard the mimicking robot highly (Bailenson & Yee, 2005) and as life-like (Yamaoka, Kanda, Ishiguro, & Hagita, 2006). However, the timing of responding was fixed in the previous work. In other words, it has not been deeply explored how a human's impression of the interactant's response could vary by the timing of responding. In this paper, therefore, we address the issue on the relationship between the response latency of the interactant and the participant's impression, especially the *feeling of being looked at*. In the rest of this paper, we first propose the *social reflex hypothesis* on responsive, nonverbal interaction. Then, to support our hypothesis, we describe an experiment with an anthropomorphic, on-screen agent that can respond to participants with eye blinking, and discuss the implication of the experimental results.

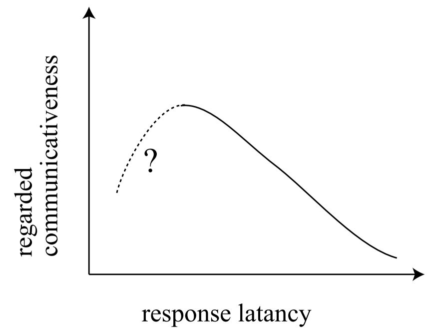


Figure 1: Social reflex hypothesis: relation between the *feeling of being looked at* by the partner and the partner's response latency

Social reflex hypothesis

Yoshikawa et al. (Yoshikawa, Shinozawa, Ishiguro, Hagita, & Miyamoto, 2006) have demonstrated that the responsive gaze shift of a robot could give participants a stronger *feeling of being looked at* by the robot even though the gaze shift is performed in either mimicking or non-mimicking manner. Furthermore, they have reported that the degree of such a feeling correlated with the amount of response latency, i.e., more rapid response could give a stronger *feeling of being looked at*. The *feeling of being looked at* can be supposed to be a feasible measure of the *feeling of being attended to* or the degree of a participant's regard of the interactant's communicativeness, and therefore will also be examined in this paper. From the findings in the previous work, it could be conjectured that humans' feeling of being looked at depends not on the ways and channels of responding but mainly on the response latency of the interactants. They have not examined the relation between the *feeling of being looked at* and the response latency around the range of almost zero latency. However, the relation might be drastically changed around there since humans seldom experience zero latency responses by their interactants.

In this paper, we propose the *social reflex hypothesis* as follows (see Fig.1). In interpersonal communication, humans unconsciously exhibit arbitrary nonverbal responses to the partner's nonverbal behaviors with a certain socially correct latency. Meanwhile, they unconsciously anticipate that their communication partner should respond within a certain range of latency. Hereby, humans might regard their interactant as communicative when the interactant responds to them within such a range even if the interactant is not animate. On the other hand, such socially correct range of latency is considered to be adapted through the feedback response from the interactants, or be optimized through the co-evolutionary process among the members of a society and then innately given to us. In either case of acquisition, the range would not be around zero latency since persons can not always respond to others with such small latencies. Moreover, since the only person who can respond to one's behavior with such small latencies is considered to be oneself, one would feel weird as if one confuses the interactant with oneself when one is facing the interactant that responds with such small latencies.

If the above hypothesis is true, we can obtain directly applicable requirements to design a communication robot. To investigate the social reflex hypothesis, we conducted an experiment with an interactive on-screen agent that could exhibit responsive blinking behavior to the participant's blinking. We made the participants evaluate their *feeling of being looked at* by the on-screen agent as a potential indicator of their unconscious regard of the interactant's communicativeness. Either of three conditions, namely *immediate*, *rapid*, and *slow* conditions in which different response latencies had been specified, were assigned to participants. In the *immediate* condition, it responded with about 0.0 second latency while in *rapid* and *slow* conditions it responded with about 0.5 and 2.0 seconds latency, respectively. We predicted that the ratio of feeling being looked at would increase along with the decreasing of the response latency but would suddenly decrease if the response latency approached to almost zero seconds (see Figure.1).

Experiment

Setup

An on-screen agent was shown on a 19-inch monitor display in this experiment (see Figure 2). It consisted of only a head that had the appearance of a gender-neutral adult without hairs or eyebrows. It could move its eyes in both pan and tilt axes as well as its eyelids. Although it could also move its mouth and neck, these degrees of freedom were fixed in this experiment.

Blinking by the participants was detected by a gaze detection device (Voxer, Nac Image Technology Inc.) and then fed to the blinking controller for the on-screen agent. Once the face of a participant is registered, the gaze detection device keeps tracking her face. Meanwhile, it emits infrared light on the participant's eyes and records the reflection to calculate where she is looking at. The reflected infrared image is used to calculate the participant's fixating point on a monitor display in front of her. The gaze detection device can also detect

Figure 2: Experimental setup

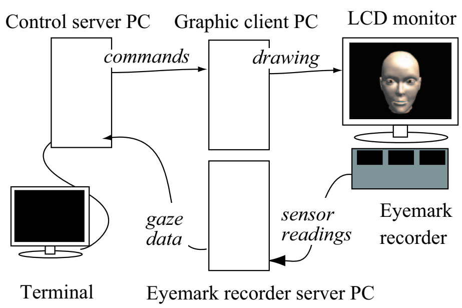


Figure 3: The information flow in experiment

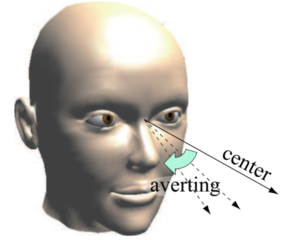
the sizes of the participant's pupils that tells us whether the person blinks or not since the sizes of the pupils are detected as zero during the blinking.

Figure 3 illustrates the information flow in the system of the experiment. The control server PC receives the data including participant's fixating point as well as the sizes of the pupils, and then judges whether the participant is blinking or not from the value of the pupil size. Based on the occurrence of the person's blinking, the on-screen agent determines whether to blink. The control signal of blinking is sent to the client program to draw the on-screen agent through a TCP connection. The data of the participant's fixating point and the size of the pupil were logged together with the parameters of the on-screen agent's DOFs.

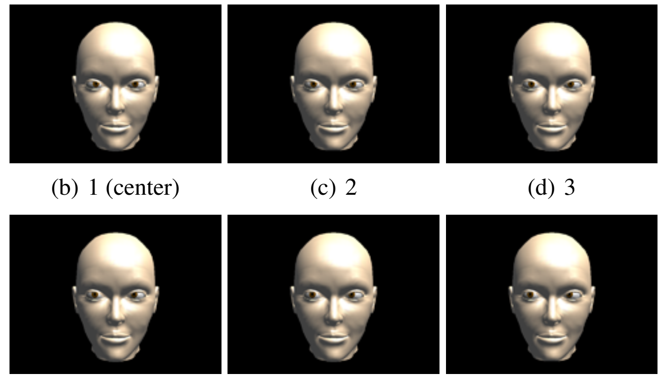
Procedure

The participants were told that they would attend to an experiment to reveal how a person perceives the other's gaze. They were asked to sit across from a monitor display and to look at an on-screen agent that sometimes looked at the participant while sometimes it did not. After the agent disappeared, they had to judge whether the agent had looked at them and say "yes" if they felt they had been looked at by the agent or "no" otherwise.

After this instruction, the experimenter first let a participant sit across from the monitor and the gaze detection device, and then calibrated them. The experimenter did not explain that the device was used for the blinking control of the on-screen agent, but only that it was used for analysis of one's gaze behavior when one looked at another.



(a) fixating angles



(b) 1 (center) (c) 2 (d) 3

(e) 4 (f) 5 (g) 6

Figure 4: The direction of the on-screen agent's eyes: (a) a schematic illustration of the angle to focus on or avert from the participant's face, and (b)∼(g) appearances of the on-screen agent in the order from (b) one which focuses on the participant's face to (g) one which fixates at most distant point from the participant's face

Each trial was started from the position of looking at the blank screen. When the participant chose to begin the next trial, she pushed a button in her hand, and an on-screen agent would be shown on the display until 20 seconds passed. The trials iteratively proceeded 48 times. The participant had two minutes break after each twelve trials.

Conditions

We compared three conditions of responsive blinking in which the on-screen agent responsively blinked to the participant's blinking with different response latency. In the *immediate* condition, it blinked just after the blinking of the participant was detected. Note that there were still 66 milliseconds time lag from the occurrence of the participant's blinking to the occurrence of the agent's one due to the delay in the sampling. In *rapid* and *slow* conditions, it waited 500 milliseconds and 2000 milliseconds, respectively, after the participant's blinking was detected, and then blinked. Note that plus 66 milliseconds time lag due to sampling delay also existed in these conditions. In addition, the on-screen agent blinked twice at the beginning of each session, namely at 0.3 and 0.55 seconds, in order to provoke the participants' blinking.

The on-screen agent sometimes looked at the participant but sometimes did not. These kinds of variations were per-

Table 1: The number of participants for each condition

Method	Male	Female	Total
Immediate	2	7	9
Rapid	4	5	9
Slow	1	6	7
Total	7	18	25

formed by drawing it as if shifting its fixating point in the horizontal plane (see Fig.4(a)). We picked up six positions for the fixating targets by the on-screen agent from the center to the left side of the participant in the horizontal line. The appearance of the on-screen agent are shown in Fig.4(b) $\sim$ (g).

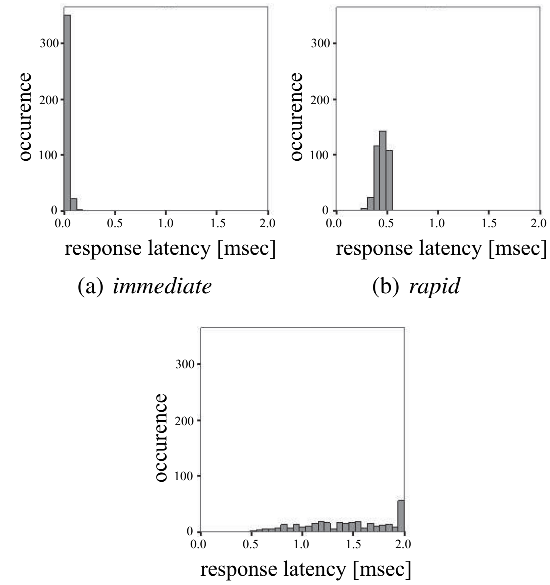
Result

Subjects attended an experiment in only one of the three conditions described in the previous section. The number of the participants for each condition is listed in Table 1. The age of the participants ranged from 19 to 49. The average age was 31.1 while the standard deviation was 8.2.

The statistics of blinking in the interaction

We observed 48 sessions for each participant, i.e., in total 1200 sessions from 25 participants. However, it was necessary to eliminate some data before the analysis, which were obtained in the sessions where the system might have failed to detect blinking. We regarded that the participant blinked every time when the readings of pupil sizes were zero. Since they were zero not only when the participant closed her eyes but also when the gaze detection device failed to track the participant's head, the system sometimes wrongly detected blinking. As exceptional data possibly including such false positives, we simply eliminated data in the sessions where the detected number of blinking or the total duration of closing participants' eyes was large. The exceptional thresholds of the number of blinking and the duration of closing eyes were set to 20 times and 10 seconds, respectively , each of which values was regarded as apparently abnormal. We also eliminated the cases where the on-screen agent did not have a chance to respond since the participant blinked almost at the end of the session or not at all. Through the automatic process to exclude such abnormal cases, we obtained data of 1091 out of 1200 sessions.

We calculated the average numbers of blinking of the participants and the on-screen agents during the sessions in three conditions. ANOVA and post-hoc tests revealed that there were significant differences between every pair both in the cases of participants (ANOVA: $p < 0.01, F(2,1088) = 48.4$; Dunnet's C: $p < 0.05$ between every pair) and agents (ANOVA: $p < 0.01, F(2,1088) = 71.7$; Hochberg's GT2: $p < 0.001$ between every pairs). Note that Dunnet's C was applied since we cannot assume that the variances of the number of blinking of the participant were equivalent among groups (Levene: $p = 0.026$).



(a) *immediate* (b) *rapid*

(c) *slow*

Figure 5: The histgram of response latency of the on-screen agents in (a) *immediate*, (b) *rapid*, and (c) *slow* conditions

The average response latency of the on-screen agent's blinking and its histgram for each condition are illustrated in Fig.5. The response latency is defined as the time interval between when the on-screen agent blinked and when the participant blinked last before the agent's blinking. Note that 74 cases with the response latency over five seconds were eliminated from the analysis as were the cases where the participant was regarded to ignore the last blinking of the on-screen agent and voluntarily blink. Although we can see that the on-screen agent almost succeeded in responding with the specified response latency in *immediate* and *rapid* condition, namely 0.0 and 0.5 seconds, the response latency in *slow* condition varied not only around the specified time, that is 2.0 seconds. This could happen when the participant blinked before 2.0 seconds past from her last blinking. However, we can see that the on-screen agent in the *slow* condition, at least, succeeded in responding slower than in other conditions since the response latency in this condition varied more than 0.5 seconds.

The average response latency of the participants' blinking and its histgram for each condition are illustrated in Fig.6. ANOVA concerning response conditions and fixating positions, and post-hoc tests revealed that there were significant differences between *immediate* and *slow* conditions (ANOVA: $p < 0.01, F(2,1014) = 7.90$; Dunnet's C test: $p < 0.05$). This result might imply that a person adapts her response latency to the interactants.

Feeling of being looked at

Fig. 7 shows the ratio of participants' answers to the question whether they feel being looked at by the on-screen agent.

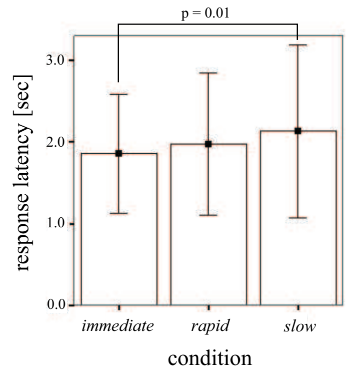


Figure 6: The average response latency of the participants

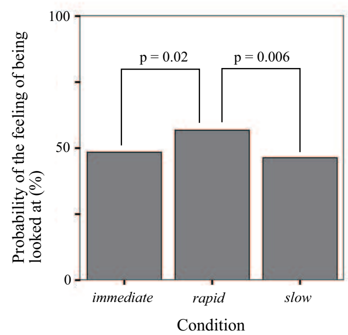


Figure 7: The probability of the feeling of being looked at in each condition

The distribution of the ratio among the fixating position of the on-screen agent is visualized in Fig.8. Since the answers and the conditions of response latency appeared to be dependent ($\chi^2 = 8.75, p = 0.013$), we applied Ryan's method and found differences of the probabilities between two pairs, namely *immediate* and *rapid* ($p = 0.023$) and *rapid* and *slow* ($p = 0.006$). Note that 1091 data were included in this analysis, which were obtained from the sessions where the on-screen agent could exhibit responsive blinking at least once. This result seems to support our hypothesis of responsiveness-based impression on the interactant where less amount of response latency cause stronger *feeling of being looked at* but the feeling is suddenly weakened if the response latency approximates to zero.

Discussion

Humans usually respond to their interactant within a certain range of response latency, or fail to keep turn-taking. We would feel that our interactant is not willing to communicate with us if the interactant does not sufficiently rapidly respond to us. Conversely, we would regard that the interactant is attending to us if the interactant rapidly responds to us. Such a feeling is regarded to be measured by asking about the participant's *feeling of being looked at* in this experiment.

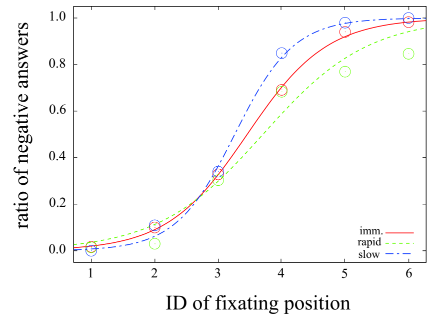


Figure 8: The ratio of negative answers on the question of whether the participant felt she had been looked at by the on-screen agent in relation to the fixating position of the on-screen agent

Through the analysis of the experiment, we confirmed that the participant's *feeling of being looked at* could be controlled by changing the latency of responsive blinking (see Fig.7). As well as apparent behaviors such as gaze shifting examined in the previous work (Yoshikawa et al., 2006), even subtle behaviors such as blinking seem to be sufficient for affecting a participant's impression when they are used in responsive manners. This result might imply the potential applicability of the social reflex hypothesis for arbitrary channels of responding. In other words, a channel-independent feature, namely timing of response, might be able to be a basis of the human cognitive mechanism for her interactant.

Furthermore, the comparison of the participant's *feeling of being looked at* in different conditions in Fig.7 seems to support the hypothesis of the detail relation between the feeling and the response latency (see Fig.1). In accord with the prediction, we can see that the *feeling of being looked at* increases along with decreasing of the response latency, as shown by the difference between *slow* and *rapid* conditions, but suddenly decreases if the response latency approaches to almost zero seconds, as shown by the difference between *rapid* and *immediate* ones.

Such a decrease around zero latency is considered to be caused by humans' anticipation of the interactant's response. We would expect that the interactant who pays attention to us should sufficiently rapidly respond to us as tested in the *rapid* condition. Meanwhile, we would not expect that the interactant responds to us too rapidly, namely with zero latency as tested in the *immediate* condition since we are not usually responded to by our interactant with such a small latency. In other words, someone who can respond to us with zero latency might be only ourselves. Therefore, the participants in the *immediate* condition might not feel like they are being looked at by the on-screen agent because it appeared for the participants not as interactant but as a weird existence that mixed the interactant and themselves.

As a preliminary investigation whether such self-other distinction is interrupted in the *immediate* condition, we asked the last 11 participants how old the on-screen agent looked.

This question was supposed to be a potential probe to measure how equivalently they regarded it with themselves. We found a weak tendency to positive partial correlation between the response latency and the transformed difference of the participant's age and the estimated one holding the participant's age fixed ($r = 0.61, p = 0.063$). Note that we evaluated the absolute difference of the tens place digits of the participant's age and the age of the on-screen agent estimated by the participant since such a difference was regarded to effectively reflect the participant's impression of age difference. This preliminary result seems to potentially support the occurrence of the interrupting self-other distinction caused by too rapid response since the participants felt less difference between themselves and the on-screen agent with less response latency. However, we need further experiments to obtain statistically significant results.

The social reflex hypothesis would provide us with the indicators to improve a communication robot in the aspects of an intuitive interface and therapeutic tool. For such an engineering purpose, we should conduct further investigations to examine to what extent we can apply the hypothesis. For example, examining the effects of the response with other channels and of the response with cross- or multi-channels would be important future issues. Furthermore, we need to address the issue of how we or a communication robot can find a suitable response latency since it might depend not only on channels but also on cultures and maybe even on persons. This issue should be investigated along with modeling studies of human cognitive developmental process such as on sensitivity to interpersonal timing (Striano, Henning, & Stahl, 2006) and on the concept of self (Rochat, 2001).

Conclusion

In this paper, we proposed the *social reflex hypothesis* on nonverbal responsive interaction in order to provide a model for the cognitive mechanism of a communication partner, which would be utilized to develop a communication robot. In the hypothesis, we supposed that unconscious nonverbal behaviors in response to a person could effect her impression of the interactant's communicativeness. We examined responsive blinking by an on-screen agent as a simple nonverbal channel and analyzed its effects on the *feeling of being looked at* that was held to be a feasible indicator of one's regard of communicativeness. In accord with the hypothesis, we found that rapid blinking responses could give stronger *feeling of being looked at* than slow or immediate ones. The decreased *feeling of being looked at* for immediate response was argued as the effect of interruption of the self-other distinction. We might conclude that further studies not only to validate and extend the hypothesis but also to apply it for designing a communication robot would be mutually beneficial for both areas of research.

Acknowledgment

The authors would like to thank Dr. Rodney Berry for his kind help to construct the experimental setup.

References

Arita, A., Hiraki, K., Kanda, T., & Ishiguro, H. (2005). Can we talk to robots? ten-month-old infants expected interactive humanoid robots to be talked to by persons. *Cognition*, *95*, B49-B57.

Bailenson, J. N., & Yee, N. (2005). Digital chameleons: Automatic assimilation of nonverbal gestures in immersive virtual environments. *Psycho. Sci.*, *16*, 814-819.

Gergely, G., & Csibra, G. (2003). Teleological reasoning in infancy: the naïve theory of rational action. *Trends in Cognitive Science*, *7*(7), 287-292.

Ishiguro, H., & Minato, T. (2005). Development of androids for studying on human-robot interaction. In *Proc. of 36th international symposium on robotics.*

Johnson, S., Slaughter, V., & Carey, S. (1998). Whose gaze will infants follow? the elicitation of gaze following in 12-month-olds. *Developmental Science*, *1*(2), 233-238.

Kamewari, K., Kato, M., Kanda, T., Ishiguro, H., & Hiraki, K. (2005). Six-and-a-half-month-old children positively attribute goals to human action and to humanoid-robot motion. *Cognitive Development*, *20*, 303-320.

Kanda, T., Ishiguro, H., Imai, M., & Ono, T. (2004). Development and evaluation of interactive humanoid robots. *Proc. of IEEE*, *92*(11), 1839-1850.

Luo, Y., & Baillargeon, R. (2005). Can a self-propelled box have a goal? psychological reasoning in 5-month-old infants. *Psyhological Science*, *16*(8), 601-608.

Melzoff, A. (1995). Understanding the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, *31*(5), 838-850.

Robins, B., Dickerson, P., Stribling, P., & Dautenhahn, K. (2004). Robot-mediated joint attention in children with autism. *Interaction Studies*, *5*(2), 161-198.

Rochat, P. (2001). Origins of self-concept. In *Blackwell handbook of infant development* (p. 191-212). Blackwell.

Shimizu, Y., & Johnson, S. (2004). Infants' attribution of a goal to a morphologically unfamiliar agent. *Developmental Science*, *7*(4), 425-430.

Striano, T., Henning, A., & Stahl, D. (2006). Sensitivity to interpersonal timing at 3 and 6 months of age. *Interaction Studies*, *7*(2), 251-271.

Watanabe, T., Danbara, R., & Okubo, M. (2003). Effects of a speech-driven embodied interactive actor "interactor" on talker's speech characteristics. In *Proc. of the 12th IEEE international workshop on robot-human interactive communication* (p. 211-216).

Yamaoka, F., Kanda, T., Ishiguro, H., & Hagita, N. (2006). How contingent should a communication robot be. In *Proc. of ACM 1st annual conf. on human-robot interaction* (p. 313-320).

Yoshikawa, Y., Shinozawa, K., Ishiguro, H., Hagita, N., & Miyamoto, T. (2006). Responsive robot gaze to interaction partner. In *Proc. of Robotics: Science and Systems ii.*