Longitudinal Child-Robot Interaction at Preschool

Hideki Kozima and **Cocoro Nakagawa**
National Institute of Information and Communications Technology
Hikaridai 3-5, Seika, Soraku
Kyoto 619-0289, Japan
{xkozima,cocoron}@nict.go.jp

Abstract

This paper reports a year-long observation of 27 typically-developing preschoolers (three-year-olds) interacting with an interactive robot, *Keepon*, which is a simple creature-like robot capable of expressing attention (by gaze/posture) and simple emotions (by body movement). *Keepon* was placed in their playroom and tele-controlled by a human operator (wizard) in a remote room. Throughout 25 three-hour-long sessions, the children showed not only individual actions, such as approach to, exploration of, and interaction with *Keepon*, but also collective social actions, where the children spontaneously and actively situate *Keepon* in their circle of, for example, playing house. This field study suggests that *Keepon*'s infantile appearance and capabilities would induce from the children (1) various prosocial behavior as if they took care of or protected *Keepon* and (2) projection of their social expectation, such as a meaning of body movement and a role in pretense play, to *Keepon*. The interaction data has been shared among the teachers and the parents for improving and motivating their child care practices.

Introduction

In the field of pedagogical services for children, teachers and parents have growing interest in and need for observing and analyzing children's peer interactions in their everyday situations, where their communicative competence and performance are naturally exhibited. However, children's peer interaction would certainly be one of the most difficult human activities to investigate, since it cannot easily be transcribed in a symbolic or quantitative form. Video recording and analysis help us a lot, but the presence of a video camera (and, of course, the person who operates it) would usually spoil the natural interactions among children. As we step back the camera away from the interactants, we would then loose the rich flow of emotional and attentional exchange. How can we observe and describe the rich flow in the intercorporal and intersubjective interactions?

We describe here our trial of utilizing an interactive robot for observing children's peer interaction in the playroom at a preschool (Figure 1), which suggests a novel way to observe human communicative behavior in everyday situations. The robot functions not only as an observation device (e.g., video camera) but also as an interaction partner for children; here we may equate "the robot" both with "the observer" and with "the interactant". Next section introduces

Figure 1: *Keepon* in the peer interaction among children in their playroom at a preschool. (Courtesy of *Kyoto Shimbun*)

the interactive robot, which enables us to tele-participate in the interaction among children and record the interaction from its own perspective. The following two sections describe child-robot interactions in a laboratory setting and a year-long longitudinal observation at the preschool, where the robot interacted with 27 three-year-olds in their playroom. Final section discusses phenomenological meaning of the robot-mediated participating observation in the children's everyday situations.

Interactive Robots for Children

There has been a growing interest in designing interactive robots that human children can naturally and intuitively interact with (Dautenhahn 1999; Breazeal 2000; Billard 2002; Michaud 2002). This research trend is motivated by the assumption that the underlying mechanism for children's embodied interaction and its development is the fundamental substratum for human social interaction in general. We describe here some psychological findings of early development in communication, our design principle of interactive robots for children, and the robot, *Keepon*, as an implementation of the design principle.

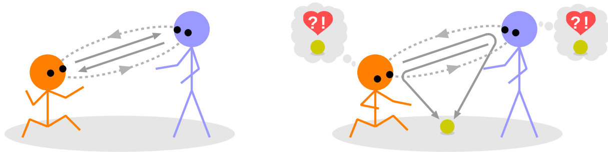

Figure 2: Exchange of emotions through eye-contact (left), and sharing perceptual information of and emotional attitude towards the target (right).

Figure 3: *Infanoid*, the child-like humanoid robot, engaging in eye-contact (left) and joint attention (right) with a human interactant. (The video monitors behind show *Infanoid*'s peripheral and foveal views.)

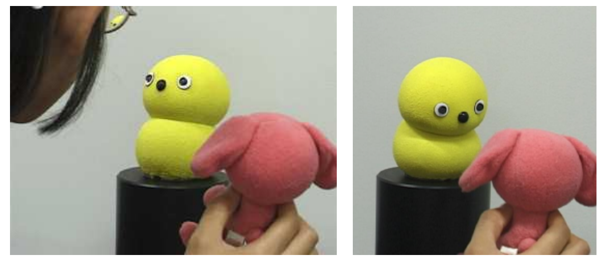

Figure 4: *Keepon*, the creature-like robot, performing eye-contact and joint attention with a the human interactant. (Note that even with this simplicity, *Keepon* retains the capabilities for exchanging attention and emotions.)

Eye-Contact and Joint Attention

Eye-contact and joint attention are fundamental activities that maintain child-caregiver interactions (Trevarthen 2001; Tomasello 1999). A child and a caregiver spatially and temporally correlate their attention and emotions, in which they refer to each other's subjective feelings of pleasure, surprise, or frustration (Dautenhahn 1997; Zlatev 2001; Kozima 2003). We believe all communication emerges from this mutual reference.

Eye-contact is the joint action of two individuals looking into each other's face, especially the eyes (Figure 2, left). It serves not only to monitor each other's state of attention (e.g., gaze direction) and of emotion (e.g., facial expressions), but also to temporally synchronize interactions and to establish mutual acknowledgment (Kozima 2004), such as "My partner is aware of me" and "My partner is aware that I am aware of her".

Joint attention is the joint action of two individuals looking at the same target by means of gaze and/or pointing (Butterworth 1991), as illustrated in Figure 2 (right). At the first stage of its development, the caregiver actively follows and guides the child's attention so that the child can easily capture the target; then the child gradually becomes able to follow and guide the partner's attention; finally the child and caregiver coordinate such initiations and responses, forming a dance of attention. Moreover, at the first stage, children are only capable of coordinating attention with others in relation to a visible object or event; later in its development, invisible psychological entities, such as emotions and concepts, will be incorporated into the target of attention. Joint attention not only provides the interactants with shared perceptual information, but also with a spatial focus for their interaction, thus creating mutual acknowledgment (Kozima 2004), such as "I and my partner are aware of the same target" and "Both of our actions (such as vocalization and facial expressions) are about the target".

Alternating between eye-contact and joint attention, the interactants can monitor each other's perceptual target (by joint attention) and emotional attitude towards the target (by monitoring emotion under eye-contact). We believe this mutual monitoring of mental states is the fundamental layer of any kinds of interpersonal communication.

Socially Interactive Robots

With inspiration from the psychological study of preverbal communication, we first developed an upper-torso humanoid robot, *Infanoid* (Kozima 2002), which is 480mm tall, the approximate size of a four-year-old human child (Figure 3). It has 29 actuators and a number of sensors that enables it to express attention (by direction of gaze, face, and pointing), facial expressions (by moving eyebrows and lips), and other hand and body gestures. Image processing of video streams enables the robot to detect a human face and direct its gaze to the face (eye-contact), and to estimate the direction of attention from the face and search the attention line for any salient object (joint attention), as shown in Figure 3. With a simple habituation mechanism, the robot alternates between eye-contact and joint attention with a human in the loop.

Through our preliminary experiments of child-robot interaction, we observed that most of the children enjoyed social interactions, where they read the robot's attention and emotions in order to coordinate their behavior. However, we also found that *Infanoid* conveys overwhelming information to some of the children, especially those under three years of age. This is probably because of its anthropomorphic but strongly mechanistic appearance. Each of the moving parts induces qualitatively different meanings in the children, who then need effortful integration of the separated meanings into a holistic recognition of a social agent.

Minimal Design

Having learned from the younger children's difficulty in understanding holistic information from *Infanoid*'s mechanistic appearance, we then built a simple creature-like robot,

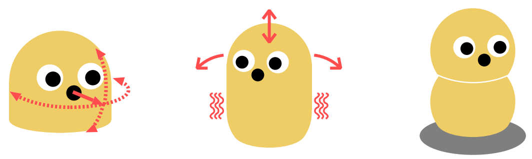

Figure 5: Designing *Keepon*'s appearance for the interactive functions for expressing attention (left), emotion (middle), and the final sketch (right).

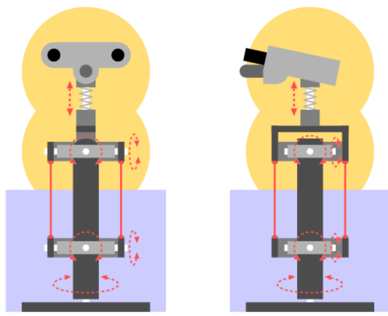

Figure 6: *Keepon*'s structure: Three motors drive the lower gimbal which is connected to the upper gimbal with four wires. Another motor pulls the head down for bobbing.

Keepon (Figure 4), which has a minimal design for exchanging attention and emotions with people, especially babies and toddlers, in the simplest and most comprehensive ways.

The simplest robots for our purpose would be a robot that can display its attention (what the robot is looking at), as exemplified in Figure 5 (left). Presence of active attention strongly suggests that the robot can perceive the world as we see and hear. The next simplest one would be a robot that displays not only attentive expression but also emotional expression. Emotive actions strongly suggest that the robot has a "mind" to evaluate what it perceives. However, if we add a number of additional degrees of freedom to the robot's emotional expression, this information flood would so overwhelm the children that they hardly grasp the gestaltic meaning. So, we decided to utilize the basic body movement for expressing some emotional states, such as pleasure (by rocking the body from side to side), excitement (by bobbing the body up and down), and fear (by vibrating the body) as shown in Figure 5 (middle). In the final design stage, we made a neck/waist line, as shown in Figure 5 (left). The neck/waist line provides a clear distinction between the head and the belly, which would give an anthropomorphic (but not overwhelming) impression to the children. The neck/waist line also provides life-like deformation of the body as the robot changes its posture.

***Keepon*, the Essence**

As an implementation of the design principle described above, *Keepon* has got a yellow snowman-like body, 120mm tall, made of soft silicone rubber. The upper part (the

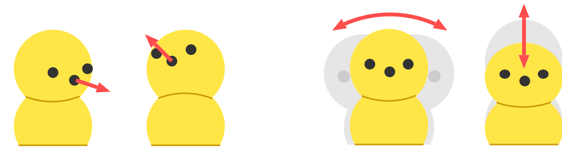

Figure 7: *Keepon*'s function: expressing attention by orienting its head (left) and expressing emotions by rocking and/or bobbing its body (right).

"head") has two eyes, both of which are color CCD cameras with wide-angle lenses (120 deg, horizontally), and a nose, which is actually a microphone. As illustrated in Figure 6, the lower part (the "belly") contains small gimbals and four wires with which the body is manipulated like a marionette using four DC motors and circuit boards in the base. Since the body is made of silicone rubber and its interior is relatively hollow, *Keepon*'s head and belly deform whenever it changes posture or someone touches it.

The simple body has four degrees of freedom: nodding (tilting) ± 40 deg, shaking (panning) ± 180 deg, rocking (side-leaning) ± 25 deg, and bobbing (shrinking) with a 15-mm stroke. These four degrees of freedom produce two qualitatively different types of actions:

- **Attentive action:** *Keepon* orients towards a certain perceptual target in the environment by directing the head (i.e., gaze) up/down and left/right. It appears to perceive the target. This action includes eye-contact and joint attention. (Figure 7, left)
- **Emotive action:** *Keepon* rocks from left to right and/or bobs its body up and down keeping its attention fixed on a certain target. It gives the impression of expressing emotions, such as pleasure and excitement, about the target of its attention. (Figure 7, right)

Note that *Keepon* can express "what" it perceives and "how" it evaluates the target with these two actions. These communicative functions of *Keepon*'s actions can easily be understood by human interactants, even babies and toddlers.

Interaction in the Lab

Before going into the preschool, we have carried out preliminary experiments of child-robot interaction using *Keepon*, in order to verify the effect of our minimal design for attentive and emotive exchange with younger children.

Method

Although *Keepon* can be operated in "automatic" mode, where the robot alternates between eye-contact and joint attention in the same way *Infanoid* does, we intentionally used "manual" mode, where a human operator (or a "wizard", usually at a remote PC) controls the robot's attention (orientations), emotions (bodily expressions), and vocalizations. The operator watches video from the on-board cameras and listens to sound from the on-board microphone; the operator also watches video from the off-board camera for the side view of the interaction, which helps the operator to figure

Figure 8: Interaction between *Keepon* and typically developing children: a 2-year-old girl showing a toy to the robot (left) and soothing the robot (right).

out what is going on, especially when the on-board cameras are occasionally covered by a hand or a face.

To perform interactive actions on the robots, the operator uses a mouse to select points of interest on the camera monitor and uses key-strokes that are associated with different emotive actions. The robot thus (1) alternates its gaze between a child's face, the caregiver's face, and sometimes a nearby toy, and (2) produces a positive emotional response (e.g., bobbing its body several times with a "pop, pop, pop" sound) in response to any meaningful action (e.g., eye-contact, touch, or vocalization) performed by the child. We thus manually controlled *Keepon* because *Keepon* should wait for a child's spontaneous actions, and when the child directs an action, *Keepon* should respond with appropriate timing and manner.

Unfolding Interaction with Keepon

We have observed 25 typically developing infants in three different age groups, i.e., 0-year-olds, 1-year-olds, and over-2-years-olds, interacting with *Keepon* (Figure 8). Each child, together with the caregiver, was seated in front of *Keepon* on the floor; there was no particular instruction for them. Interaction continued until the infants became tired or bored; on average, each infant's dealings lasted about 10 to 15 minutes. We found from these observations that infants in each age group showed different styles of interaction.

- **0-year-olds:** Interaction was dominated by tactile exploration using the hands and/or mouth. The infants did not pay much attention to the attentive expressions of the robot, but they exhibited positive responses, such as laughing or bobbing their bodies, to its emotive expressions.
- **1-year-olds:** The infants demonstrated not only tactile exploration, but also awareness of the robot's attentive state, sometimes following its attention. Some of the infants mimicked its emotive expressions by rocking and bobbing their own bodies.
- **Over-2-year-olds:** The infants first carefully observed the robots behavior and how caregivers interacted with it. Soon the infants started social exploration by showing it toys, soothing (by stroking its head), or verbal interactions (such as asking questions).

The differences between the interactions of each age group reflects differences in their ontological understanding of *Keepon*. The 0-year-olds recognized the robot as a "moving thing" that induced tactile exploration. The 1-year-olds first encountered a "moving thing", but after observing the robot's responses to various environmental disturbances, they recognized that it was an "autonomous system" that possesses attention and emotion as an initiator of its own expressive actions. The over-2-year-olds, through the recognition of a "thing" and a "system", they found that the robot's responses, in terms of attention and emotion, had a spatiotemporal relationship with what they had done to it; finally they recognized it as a "social agent" with which they could play by exchanging and coordinating their attention and emotions.

Figure 9: A self-contained version of *Keepon*, equipped with a battery and wireless connections.

Interaction at the Preschool

Having tested the safety issues and learned the typical ways children show in the interactions with *Keepon*, we then proceeded to the preschool, where we longitudinall observed natural interaction with children in everyday situations.

Method

We observed how a group of 27 children in a class of three-year-olds (average CA 4:0 throughout the year-long observation) interacted with *Keepon* in the playroom of their preschool (Figure 1). In each session, at around 8:30 a.m., one of the teachers brought *Keepon* to the playroom and put it on the floor with other toys. In the first 90 minutes, the children arrived in the preschool one after another, gradually formed clusters, and played freely with each other and with *Keepon*. In the next 90 minutes, under the guidance of three teachers, the children engaged in various group activities, such as singing songs, playing musical instruments, and doing paper crafts. *Keepon* was moved as necessary by the teachers so that it did not interfere with the activities; sometimes it sat beside the teacher who was telling a story, or sat on the piano watching the children who were singing or dancing.

A wireless version of *Keepon* (Figure 9) was placed in the playroom just like one of the toys on the floor. We tele-controlled *Keepon*'s attentional expression, emotional expression, and vocalizations (of simple sound) from a remote room by watching the video from *Keepon*'s on-board camera and an off-board camera on the ceiling of the playroom.

We recorded live interactions with the children from *Keepon*'s perspective (Figure 10). In other words, we

Figure 10: A child seen from the subjective viewpoint of *Keepon* as the first person of the interaction. (From the preliminary observations in our laboratory.)

recorded all the information from the subjective viewpoint of *Keepon* as the first person of the interaction. Strictly speaking, this subjectivity belongs to the operator; however, the interaction is mediated by the simple actions that *Keepon* performs, and every action *Keepon* performs can be reproduced from the log data. Therefore, we may say that *Keepon* is both subjective (i.e., interacting directly with children) and objective media (through which anyone can re-experience the interactions), enabling human social communications to be studied.

Observation

In the longitudinal observations (25 three-hour sessions), the children showed various spontaneous interactions with *Keepon*, individually and in a group, whose style changed over time. Here are some anecdotes about what *Keepon* experienced in the playroom. (Here , "Sn" stands for "the n-th session".)

- In S1, the children showed shyness and embarrassment to *Keepon*, not knowing well what it was and how they should do with it. From S2, they gradually started various interventions to *Keepon* — from beating to feeding.
- In S5, a girl NK (hereafter NK/f) put a cap on *Keepon*. When the cap was gone, a boy YT (hereafter YT/m) put his own cap on *Keepon*. In S7, when it was lost again, TK/m and NK/f soothed *Keepon*, saying, "Did you loose your cap?" and "Endure being without your cap."
- In S6, KT/f played with *Keepon* in the outdoor playground; a boy in the 4-year-old class came to *Keepon* and told KT/f, "This is a camera. This is a machine," but KT/f insisted, "No, this is Keepon's eyes!"
- In S8, pointing to an insect cage, SR/f guided *Keepon*'s attention to it. In S9, when NR/m beat *Keepon*'s head several times, HN/f stopped him by saying, "It hurts! It hurts!" During reading time in S11, NK/f and TM/m came up and showed their picture books to *Keepon*.
- In S13, FS/m and TA/m, strongly beat *Keepon*'s head a couple of times, as if demonstrating their braveness to each other. YT/f and IR/f, observing this, approached *Keepon* and checked if it had been damaged, then YT/f said to *Keepon* and IR/f, "Boys are all alike. They all hit *Keepon*," stroking its head gently.
- In S16, after a blank of a couple of sessions, NK/f came to *Keepon* and said, "We haven't see each other for a while," as if soothing *Keepon*'s loneliness.
- In S17, YT/f taught *Keepon* some words — showing it the cap, she said, "Say, *Bo-shi*," then switched to *Keepon*'s knitted cap and said, "This is a *Nitto Bo-shi*, that you wear in winter." (*Keepon* could only respond to her by bobbing its body with the "pop, pop, pop" sound.)
- Also in S17, after two girls hugged *Keepon* tightly, other girls found a scar in its head. NK/f pretended giving medicine to *Keepon* with a spoon, saying, "Good boy, you'll be all right."
- In S19, after playing with *Keepon* for a while, IZ/m asked other children nearby, "Please take care of *Keepon*." IZ/m managed to get an OK from KT/f, then left from *Keepon*.
- In S22, after all the children practiced a song with the teachers, several of them ran to *Keepon* and asked one by one, "Was it good?", to which *Keepon* responded by nodding and bobbing for praise.
- In S23, NZ/m noticed *Keepon* had a flu mask and asked *Keepon*, "Caught a cold?" NK/f then put a woolen scarf around *Keepon*'s neck, then NR/m and YS/f asked NK/f, "Is he ill?" and "Got a cold?"
- In 25, NK/f gave a toy sled to *Keepon*. *Keepon* showed a preference to another toy NK/f was holding. After some negotiation, NK/f brought another sled and persuaded *Keepon*, "Now you have the same hing as mine."

Especially during free play time (the first 90 minutes), the children showed a wide range of spontaneous actions, not only dyadic one between a particular child and *Keepon*, but also n-adic one, in which *Keepon* functioned as a pivot of interpersonal play with peers and sometimes with teachers. Since the children were generally typically developing, they often spoke to *Keepon*, as if they believed that it had a "mind". They interpreted *Keepon*'s responses, although they were merely simple gestures and sounds, as having communicative meanings within the interpersonal context. We have never observed this with autistic children at the same age, who rarely interacted with peers (Kozima 2006). Compared with the experimental setting, where children became bored after 15-minute interactions, it is interesting that children in the preschool never lost interest even after 20 sessions.

Discussions and Conclusion

We reported in this paper our longitudinal observations of 27 preschool children's peer interactions in their playroom. The peer interaction was facilitated and observed by an interactive robot, *Keepon*, which was tele-controlled by a remote operator. Qualitative analysis of the children's dyadic interaction with *Keepon* and n-adic interaction mediated by *Keepon* suggests the following:

- The children were able to approach *Keepon* with a sense of curiosity and security. This was probably because the children intuitively understood the gamut of *Keepon*'s possible actions (e.g., action repertoire and range of motion) and perception (e.g., to look and to hear).

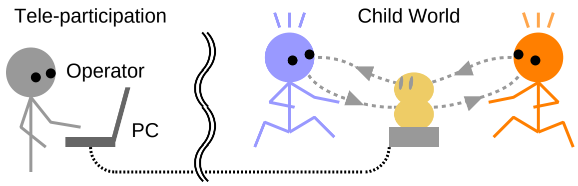

Figure 11: Tele-participation in *Child World*: the operator subjectively experiences the *in situ* interactions, while anyone could also re-experience it from the video data.

- Through the approach and exploration phases (the first several sessions), the children gradually attribute "mental states" (e.g., wanting a cap, being lonely, having pain, likes and dislikes) to *Keepon*.
- Some of the prosocial actions the children exhibited to *Keepon* were probably copied from what they had been done by their caregivers (e.g., feeding food or medicine, soothing and praising). Being small, helpless, and immobile, *Keepon* would be an ideal target for the children to imitate such caretaking behaviors.
- Each child exhibited a different style of interaction that changed over time, which would tell us a "story" about his or her personality and ability. This rich individual profile would not be thoroughly obtained by the snapshot result of a developmental test.

The stories of the children accumulated as video data have been utilized by teachers at the preschool for improving their pedagogical services. We also provided the video data to parents with the hope that it may positively influence their own child-care.

What we have done in the field is "participating observation", where *Keepon* functioned not only as a camera but also as the agent who actually interacted with the children. The human operator tele-controlled *Keepon* and recorded the interactions from the perspective of *Keepon* as the first person of the interactions (Figure 11). In other words, the operator transferred his or her viewpoint to the position of *Keepon*, where he or she could interact with the children by means of the robot's simple, small appearance and comprehensive actions. Therefore, the video data contains subjective experience that *Keepon* (and so the operator) had in the interaction, which can then be re-experienced and re-interpreted by anyone including the children's caregivers. To summarize, *Keepon* provides the operator with both subjective experience and interpretation of interaction *and* objective observation open to anyone to re-experience and re-interpret the interaction.

Acknowledgments

The authors are grateful to Ms. Ikuko Hasegawa (Hachioji Preschool, Omihachiman, Shiga, Japan) for her infinite support for this field study.

References

A. Billard, Play, dreams and imitation in Robota, in K. Dautenhahn, et al. (eds.), *Socially Intelligent Agent*, pp. 165–173, Dordrecht, The Netherlands: Kluwer Academic Publishers, 2002.

C. Breazeal and B. Scassellati, Infant-like social interactions between a robot and a human caretaker, *Adaptive Behavior*, Vol. 8, pp. 49–74, 2000.

G. Butterworth and N. Jarrett, What minds have in common is space: Spatial mechanisms serving joint visual attention in infancy, *British Journal of Developmental Psychology*, Vol. 9, pp. 55–72, 1991.

K. Dautenhahn, I could be you: The phenomenological dimension of social understanding, *Cybernetics and Systems Journal*, Vol. 28, pp. 417–453, 1997.

K. Dautenhahn, Robots as social actors: Aurora and the case of autism, *Proceedings of the International Cognitive Technology Conference*, pp. 359–374, 1999.

H. Kozima, Infanoid: A babybot that explores the social environment, K. Dautenhahn, et al. (eds.), *Socially Intelligent Agent*, pp. 157–164, Dordrecht: Kluwer Academic Publishers, 2002.

H. Kozima and A. Ito, From joint attention to language acquisition, J. Leather and J. van Dam (eds.), *Ecology of Language Acquisition*, pp. 65–81, Dordrecht: Kluwer Academic Publishers, 2003.

H. Kozima, C. Nakagawa, and H. Yano, Can a robot empathize with people?, *International Journal of Artificial Life and Robotics*, Vol. 8, pp. 83–88, 2004.

H. Kozima and C. Nakagawa, Social robots for children: Practice in communication-care, *IEEE International Workshop on Advanced Motion Control*, pp. 768–773, 2006.

F. Michaud and C.Théberge-Turmel, Mobile robotic toys and autism, in K. Dautenhahn, et al. (eds.), *Socially Intelligent Agent*, pp. 125–132, Dordrecht, The Netherlands: Kluwer Academic Publishers, 2002.

M. Tomasello, *The Cultural Origins of Human Cognition*, Cambridge, MA, USA: Harvard University Press, 1999.

C. Trevarthen, Intrinsic motives for companionship in understanding: Their origin, development, and significance for infant mental health, *Infant Mental Health Journal*, Vol. 22, pp. 95–131, 2001.

J. Zlatev, The epigenesis of meaning in human being, and possibly in robots, *Minds and Machines*, Vol. 11, pp. 155–195, 2001.