Combining Embodied Models and Empirical Research for Understanding the Development of Shared Attention

Ian Fasel[1], Gedeon O. Deák[1], Jochen Triesch[1], and Javier Movellan[2]
[1]Department of Cognitive Science and [2]Institute for Neural Computation, UC-San Diego
9500 Gilman Drive, La Jolla, CA 92093-0515
{ianfasel, deak, triesch}@cogsci.ucsd.edu, movellan@inc.ucsd.edu

Abstract

The capacity for shared attention is a cornerstone of human social intelligence. We propose that the development of shared attention depends on a proper interaction of motivational biases and contingency learning mechanisms operating in an appropriately structured environment. Atypical contingency learning leads to deficits in shared attention as seen in children with autism. To test this theory, we describe a unique research effort that combines theoretically rigorous modeling techniques using both simulated and robotic learning systems with novel empirical investigations of social learning and development in infants and toddlers with and without developmental disabilities. We believe that studying embodied learning models, whose input data (from a real or virtual caregiver) is modeled after real infant-caregiver interactions, will lead to a better understanding of the development and dysfunction of shared attention.

1. Introduction

The social self emerges from modest beginnings. Compared to other mammals, human neonates are nearly helpless, nearly insensate, and slow to respond (when awake). By about three months, however, infants typically show burgeoning social responsiveness (Cole & Cole, 1996), and by 3-6 months infant-caregiver dyads typically engage in complex patterns of reciprocal interaction (Kaye, 1982). These patterns will eventually include imitation, turn-taking games, seeking to share a focus of attention, becoming upset by separation and showing anxiety toward strangers, and using parents' affect to interpret ambiguous events (Morales et al., 1998; Scaife & Bruner, 1976; Walden & Ogan, 1984). By the first birthday, normally developing infants show robust gaze-following (Deák et al., 2000), as well as a variety of other sophisticated pre-linguistic communication and shared attention skills, such as pointing and requesting behaviors.

These joint or shared attention skills, defined as re-orienting or re-allocating attention to a target *because* it is the object of another person's attention, play a critical role in autonomous mental development. Joint attention is the foundational mechanism by which independent agents coordinate their activity to develop more complex interactions (e.g., communication), and it supports inferences about other agents' current and future activity, both overt and covert. Investigations into the computational nature of shared attention may give insight into a variety of complex behaviors in humans, and may provide critical information about mental disorders such as autism.

How does shared attention develop? One of the most prominent developmental theories of shared attention, which has informed the pioneering effort to model social learning in robotic systems by (Breazeal & Scassellati 1998, 2000), is Baron-Cohen's theory of social-cognitive modules (Baron-Cohen, 1995). Baron-Cohen's model makes strong claims about social knowledge and its cognitive underpinnings. He posits several discrete mechanisms involved in shared attention: a primitive Eye Direction Detector (EDD), and later-evolving faculties including an Intentionality Detector (ID), a Shared Attention Mechanism (SAM), and two Theory of Mind modules (TOMs). These modules come online in a stereotypical time-course, and eventually result in a socially sophisticated individual.

Baron-Cohen's model is a useful description of key elements of shared attention, however it provides little detail about the mechanism of these modules, or how these modules come online. For understanding development, it may be more useful (and more parsimonious) to take a systems approach, with an eye towards understanding how complex behaviors might emerge from simple biases and learning mechanisms. For instance, rather than appeal to a missing or malfunctining innate SAM to explain the dysfunction of shared attention in autistic children, it may be possible to attribute these abnormal social behaviors to a (presumably more general) cognitive difficulty in inhibiting or filtering information (Kootz, Marinelli, & Cohen, 1982; Pennington & Ozonoff, 1996). Under this explanation, the abundance of information in social situations might be hard for autistic children to filter, and therefore aversive. This would explain behaviors like avoiding eye contact and ignoring others' communication bids. Shared attention deficits, and other social and cognitive deficits, might therefore result from an affect- and cognition-driven social avoidance that limits social input and social learning.

We therefore propose an alternative learning model for development of shared attention, as well as a unique methodology to explore it. Because shared attention is a mechanism for expressing and inferring internal, mental states through external, bodily relationships, we believe that an understanding of this development may best be obtained both through careful attention to social behavioral patterns, as well as through theory testing using embodied models such as robots. Below, we describe our model, and then present a description of an ongoing research effort to test it through a combination of virtual and embodied modeling techniques with novel empirical investigations of social learning and development in infants and toddlers with and without developmental disabilities.

2. A Developmental Model of Shared Attention

Developmental theory in the last decade has shifted from nativist and modular approaches as researchers have recognized similarities across cognitive processes underlying various skills. The dynamic systems approach (Thelen & Smith, 1994; Elman et al., 1996) parsimoniously attributes the emergence of complex cognitive skills (e.g., finding objects, imitation, word learning) to basic processes of attention and pattern learning in subsymbolic distributed networks (Deák, 2000; Diedrich et al, 2001; Jones, 1996). Our model of the development of shared attention takes such a dynamical systems approach. We hypothesize that early emerging, or deeply canalized preferences, with relatively simple learning mechanisms and information seeking routines, might result in emergent behaviors such as gaze following in normally developing infants. Moreover, specific parameter settings in the learning model might inhibit the development of shared attention in children with autism.

The Basic Set

We propose that a basic set of key ingredients are sufficient to develop shared attention. These are (1) a set of motivational biases to look at and shift attention between interesting things, (2) a learning mechanism which takes advantage of the temporal structure of predictable, contingent interactions, and (3) a structured environment providing strong correlation between where parents look and where interesting things are. Let us explore these in more detail.

Motivational Biases In our model, strong infant social orientation is an important early factor driving the development of shared attention. Normally developing infants prefer social stimuli, and respond selectively to parents, even smiling in anticipation of a feeding. They enjoy seeing faces in general (Dannemiller & Stevens, 1988), and caregivers in particular (Field et al., 1984). These preferences are not limited to faces; for example most infants prefer their mother's voice (DeCasper & Fifer, 1980). In contrast, many children with autism do not show normal social preferences (e.g., Dawson et al., 1998; Hobson, Ouston, & Lee, 1988).

Habituation A basic learning mechanism critical to cognitive development is habituation. Though its role in social learning is seldom discussed, we hypothesize that "triadic interactions," in which infant and caregiver attend to one another and to a third object/event (e.g., a spoon for feeding), rest on a cycle of habituation in which attention to an object of attention (e.g., spoon) decreases over time, resulting in shift of attention to another object (e.g., adult's face). Butcher et al. (2000) showed that between 8 and 12 weeks, infants begin to shift visual attention from a central stimulus when a peripheral stimulus is introduced. More specifically, by 6 months infants begin breaking mutual gaze with their mother to look at distal objects. Also, interest in inanimate objects increases from 3 to 6 months [1], so competition for infants' attention between one interesting stimulus (i.e. caregiver's face) and another (e.g. a colorful toy) should emerge, and possibly be responsible for gaze alteration.

Contingent Interactions It seems that predictable interactions are enjoyable to infants. Watson (1972) found that 2-month-old infants detect when an object is responsive to their behavior, and this elicits positive affect (i.e., smiling; cooing). By 3 months infants come to prefer to interact with stimuli that produce highly, but not perfectly, predictable responses (Watson & Ramey, 1987).

We believe that this "tuning" of preference for high predictability plays a critical role in later communication and social learning. For example, Movellan and Watson (1987) tested how infants interpret objects as having a "line of regard." Ten-month-old infants interacted with a non-human-like robot. For an experimental group, the robot moved a mechanical "face" in response to the child and to stimulus events. For a control group, the robot made the same behaviors, but independent of the infant's actions. Infants in the experimental group showed more vocalize-wait sequences, social behaviors, and expressions of delight and interest. Control infants quickly lost interest in the robot. The most relevant result was that experimental infants looked more often in the direction specified by the robot's "head" orientation (Fig. 1; see also Johnson et al., 1998). Apparently, the discovery that another agent's gaze is a cue worthy of monitoring relies on the infant's ability to detect the contingency structure in interactions with that agent.

Learning Mechanism Given these motivational biases to shift attention between social stimuli and interesting objects in a partially predictable environment, we believe that the infant brain is especially tuned to seek out and learn contingencies in human interactions to maximize internal rewards due to these biases.

Work in neural computation shows that neural learning algorithms are modulated by global parameters (see below) whose optimal values must become "tuned" to the structure of the environment. One learning algorithm that makes particular sense in this context is temporal difference or TD-learning (Sutton & Barto, 1988). This algorithm has generated a great deal of excitement in the in the computational neuroscience community as a model of learning in the brain (Houk. et al. 1995; Dayan et. al, 2000), due in part to a series of critical experiments by Schultz et al. (1997) in which it was found that the activity of dopaminergic neurons in the basal ganglia changed during learning in accordance to the TD-learning model.

Parameters that are optimal for learning in highly predictable environments may be ineffective in less predictable environments. It has been proposed that in the brain global parameters are controlled by neuromodulator systems that project diffusely from the brainstem to the cerebral cortex, basal ganglia, and cerebellum (e.g., Katz, 1999). More specifically, Doya (2000) proposed that the acetylcholinergic system is responsible for control of learning rate parameters, the noradrenergic system for degree of exploration (related to habituation), the serotonergic system for the time scale of evaluation (related to preferred probability), and the dopaminergic system for encoding prediction errors.

Structure of the environment: The mechanism we have described allows infants to learn social contingencies in a suitably structured environment. Where does this structure come from? We believe that parents provide this structure in one-on-one interactions, and that infants readily learn this structure and tune their responses accordingly. Kaye (1982) found that mothers structure feeding predictably even for neonates, timing and sequencing their actions and reactions non-randomly. Caregivers also predictably look at interesting things, such as other faces, or at their own acting hands (Hayhoe et al. 1999; Land et al. 1999). From this infants may readily learn that the pose of their parent's head predicts the location of interesting things.

2.1. Some Views About Shared Attention in Normal and Autistic Children

How do all these pieces fit together to produce shared attention? In our model, the disposition to select social experiences increases encoding social information. A preference for social experience then facilitates prediction of social events, and the positive emotional states induced by social interactions act as powerful reinforcers. Triadic interactions—object play, feeding, and perhaps bathing and diapering—provide infants with the critical input to learn associations between their caregiver's face poses and regions of space likely to contain interesting objects. Infants thus learn the event structure of the parent-infant interactions and the direction of parents gaze, resulting in shared attention behaviors.

We hypothesize that in autism the neural mechanisms for learning are set in a manner that is not optimal for the type of contingencies typically found in social interactions. For instance, because the triadic interactions that provide the input for the learning mechanism depend on habituation, we would predict that if the normal cycle of habituation and attention shifting is disrupted, infants will not learn to follow gaze. Indeed, Swettenham et al. (1998) showed that children with autism are significantly delayed in shifting and distributing attention between people and objects (see also Courchesne et al., 1994).

Our hypotheses about abnormal parameter settings of the learning mechanisms also fall in line with previous theories about autism. For instance, Gegerly and Watson

Figure 1: Example of following the "gaze" of a robot; Movellan & Watson (1987). The reflection of the robot can be seen in the mirror at the right of each image. Note also the child's positive affect.

(1999) hypothesized that infants with autism do not respond to the uncertainty levels and timing constants found in typical social interactions. They proposed that infants with autism prefer much higher – even perfect – predictability than is seen in typical social interactions. Preference for perfectly predictable event contingencies also might explain symptoms of autism such as steoreotyped motions. Maygar and Gegerly (1998) found support for the hypothesis by presenting normally developing toddlers and children with autism with two computer animated objects: one perfectly controlled by the infant's hand, and one only partially controlled by the hand. Whereas non-autistic toddlers preferred to look at the partially controlled object, children with autism preferred to look at the perfectly controlled one. Note that if a child cannot learn to predict typical (high-but-imperfect) social contingencies, it will be difficult to learn, for example, to follow a person's head turns in order to bring interesting sights into view.

3. Empirical Studies and Robotic Models: an Integrated Approach

How can we test this theory of development of shared attention as an emergent phenomenon? In the following we outline an ongoing research program to test our hypothesis. Because the environmental context of shared attention—physical setting, spatial arrangement of people and objects, etc.—is centrally important, we believe that the model must include an accurate description of the changing pattern of social input to a developing system. Thus, our research combines observational studies of humans and human-robot interactions to provide realistic input for both virtual and robotic systems. The benefits of a robotic approach for testing models of development have been discussed cogently by Breazeal and colleagues (Breazeal & Scassellati 1998, 2000), and we have adopted a similar outlook.

Our approach has three main components. The first uses carefully controlled observational studies to identify the normative input (i.e., interactions) between infants and caregivers. We have selected interactions that, we believe, allow typically developing infants to learn gaze following. The second component models the dynamics of these infant-caregiver interactions in virtual, and then robotic, simulations. This will allow us to first determine which parameter settings of the learning model facilitate or inhibit the emergence of shared attention and gaze following, and then implement learning trials with an embodied agent, or robot, that experiences systematic input delivered by a human "teacher." In the third component, normal, Down's syndrome, and autistic children interact with a humanoid robot that responds to the children with specific contingency schedules, roughly based on the results of the previous components. The goal is to determine whether certain timing and response probabilities in social interaction can facilitate autistic toddlers' response to a social agent. These three components are highly interactive, as data from each can help to inform and modify changes in modified versions of the others. This interaction is illustrated in Figure 2. Below, we go into some detail about the methods of our ongoing research.

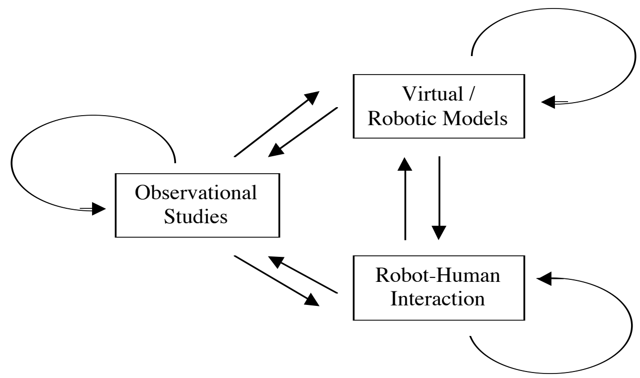

Figure 2: The dynamics of infant-caregiver interactions determined from observational studies serves as the input for virtual and robotic models. Theories developed in these models then serve as the basis for robot-human interaction experiments, which then motivate further observational and modeling studies.

3.1. Study of Triadic Interactions

Certain interactions that increase in frequency between 3 and 9 months are predicted to contribute directly to the emergence of gaze following and shared attention. These include face-to-face interactions with small, colorful objects (i.e. baby toys), held near the baby but to one side or the other. While holding an object, parents might occasionally shift gaze from infant to object, and the infant is predicted to learn to do the same, with increasing regularity from 5 to 12 months. Thus, gaze alteration, gaze shifting, and the correlation between parent's head pose (i.e., turned to left or right) and location of the object (to left or right) allow the infant to gradually learn a mapping between the caregiver's face poses and locations in space.

To better describe and understand these interactions, we are piloting an observational study of the structure and timing of infant-caregiver interaction. In this study, parents are seated facing the infant and asked to interact naturally with the infant. Cameras placed behind the infant and the parent, as well as a third overhead camera, record both partners' shifts in gaze, gaze direction, and parents' hand position, contents, pointing, and reaching. Several times during the experiment, the parent also attempts to re-direct the infant's attention to one of four visual targets spaced along the walls of the room. Finally, in a social orientation task (from Dawson et al, 1998), a confederate engages the infant with a toy. While the infant is engaged, the parent calls the infant, at a regular interval, until the infant orients to the parent.

Videos of face-to-face interactions are then coded for activity, gaze direction and head pose of caregiver and infant, vocalizations, affect, and timing and sequence of these events. The goal of this phase of the study is to precisely describe the statistics of social interactions: how mutual gaze, gaze shifting, and gaze following change between 6 and 12 months in normally developing infants and infants with Down's syndrome. This description might aid in early detection of social deficits. Analyzing the event structure (e.g., timing and sequential probabilities) of social interactions also will allow us to identify the normative input (i.e., interactions) that allows typically developing infants to learn gaze following.

3.2. Learning of Gaze Following by Virtual and Robotic Agents

We have proposed that infants' gaze following results from fairly generic learning mechanisms, in conjunction with preferences for social stimuli, operating in an appropriately structured environment. In order to test this hypothesis, we are building artificial agents, both robotic and computer animated, with learning processes and preferences modeled after 3-5-month-old infants, and providing them with an appropriately structured social environment, first via input from a virtual caregiver, and then from a human caregiver to robotic agents. If the systems learn gaze following, it suggests that infants might learn it in a similar way.

We are making the use of two complementary research platforms for this phase of the study: virtual humans and an anthropomorphic robot head. Virtual humans are software agents that act in computer generated virtual environments. Of course, the virtual environment is a simplification of the real world; therefore, we view the virtual humans as a platform for rapid prototyping of theories to be further refined and studied in an anthropomorphic robotic head (see below). Their advantage is that they allow for rapid testing and refinement of learning models because simulation speed is not restricted to the time course of real world social interactions.

Virtual caregivers are programmed to manipulate objects (e.g., toys) in the virtual world while looking at them, and to periodically look at the infant. The exact behavioral program of the caregiver is modeled after the statistics of human infant-caregiver interactions derived from the observational study. The virtual infant is likewise programmed with realistic preferences and capabilities described above.

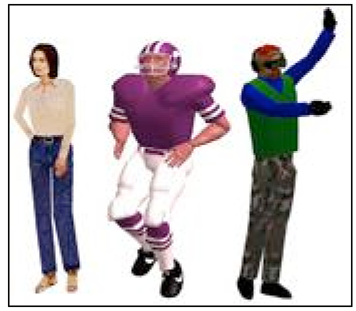

Figure 3: Some of the computer animated characters (from *www.bdi.com).*

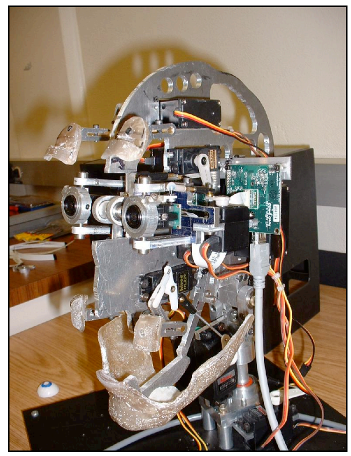

Figure 4: View of internal mechanisms of the robot head: cameras, interface boards, and servo motors.

In order to minimize the confounds inherent in using a virtual world as opposed to a real world, face detection and pose estimation is done with the same developmentally plausible models as used in the robot, operating on computer generated images from the 3-D virtual world. Meanwhile, the infant is designed using the neurally plausible learning algorithm described above, TD-learning, to learn to predict reward contingencies. This algorithm's efficacy depends on finding parameters that match the contingency structure of the environment, and much of our initial efforts are devoted to finding suitable parameters to enable learning in a plausible developmental time frame. These efforts in themselves might provide insights into parameter alterations that impair social learning in ways that resemble autistic behavior.

Once the appropriate parameters for the learning rule are determined in the virtual agents, the robot head will be endowed with perceptual and learning capabilities refined in the virtual infants. A human caregiver will then interact with the robot to try to teach it gaze following. We expect the hypothesized set of necessary learning algorithms, response tendencies, and stimulus preferences to allow a virtual infant to learn gaze following behavior, given appropriate social input. We expect perturbation of the learning parameters that alter the optimal contingency structure of social interaction to impair learning of gaze following. We speculate that a perturbed version of the system might acquire social attention deficits similar to those in autistic children.

3.3. Experimental tests of social contingency learning

Based on the results of the computer simulations we plan to conduct behavioral experiments with groups of toddlers with autism, with Down's syndrome, and with typical social abilities. These toddlers will interact with a humanoid robot that responds to the children with specific contingency schedules. The goal is to determine whether certain timing and response probabilities in social interaction can facilitate autistic toddlers' response to a social agent.

We expect non-autistic toddlers to learn to respond to robots more than autistic infants, and for toddlers with Down's syndrome to learn only slightly slower than normal (these children learn shared attention skills at near-typical times). We expect children with autism to interact less with a robot responding with the probability of typical interactions, but higher if the robot responds perfectly predictably. Normal infants might show autistic-like behavior towards a robot that is much less predictable than normal.

4. Conclusion

The development of shared attention is a cornerstone of human cognitive development. We have proposed a model of how shared attention emerges in normally developing infants through the interaction of only a few basic mechanisms: innate motivational biases, a habituation mechanism, and a contingency learning mechanism. The basic idea is that in the typical infant environment, the infant learns to predict the location of "rewarding" sights from the head pose (and eye direction) of the care giver, because care givers tend to look at "interesting" things like other faces or their own hands. These ideas have been proposed before, but no effort has been made to test them computationally. We are currently starting to test the implications of this idea in a three-pronged research effort that integrates observational, behavioral, and modeling studies.

Of central importance to our approach is the development of embodied learning models (computer animated characters and real robots) that interact with realistic care giver models or real care givers. The benefits of using robotic models to systematically study cognitive development have been highlighted recently by a number of authors (e.g., Brooks et.al. 1998; Asada et.al., 2001; Zlatev & Balkenius 2001) and also the potential of robotic models for furthering our understanding of developmental disorders have been emphasized (Balkenius & Bjoerne, 2001). With respect to shared attention, a number of researchers have endowed robots with innate shared attention capabilities (Breazeal & Scassellati, 2000; Kozima & Yano, 2001) but whether and how shared attention can be learned through interaction with the environment is still an open question.

By closely integrating robotic modeling with observational and behavioral studies, we hope to arrive at a more complete picture of the development of shared attention than any single methodology could provide. The plausibility of developmental theories can be elegantly tested using robotic models. At the same time, it is clear that modeling studies of social learning have to be informed by the statistics of real world social interactions. The extraction of data from real caregiver/infant interactions, on the other hand, must be interpreted in the light of putative learning algorithms that try to exploit the contingencies present in these interactions. Thus, we feel that a close integration of the mentioned research methodologies bring about a new way to understand cognitive and social development.

References

[1] Asada, M., MacDorman, K.F., Ishiguro, H., and Kuniyoshi, Y. (2001). Cognitive developmental robotics as a new paradigm for the design of humanoid robots. *Robotics and Autonomous Systems, 37,* 185-193.

[2] Balkenius, C. and Bjoerne, P (2001). Toward a model of attention deficit hyperactivity disorder (ADHD). In Balkenius et.al. (Eds.), *Proceedings 1st International Workshop on Epigenetic Robotics: Lund University Cognitive Studies, 85,* 61-69.

[3] Baron-Cohen, S. (1995). The eye direction detector (EDD) and the shared attention mechanism (SAM): Two cases for evolutionary psychology. In C. Moore & P. Dunham (Eds.), *Joint attention: Its origins and role in development* (pp. 41-59). Hillsdale, NJ: Erlbaum.

[4] Breazeal, C., & Scassellati, B. (1998). Infant-like social interactions between a robot and a human caretaker. *Adaptive Behavior,* Vol. 8, No. 1.

[5] Breazeal, C., & Scassellati, B. (2000). A context-dependent attention system for a social robot. In *1999 International Joint Conference on Artificial Intelligence.*

[6] Brooks, R.A., Breazeal, C., Irie, R., Kemp, C.C., Marjanovi_, M., Scassellati, B., and Williamson, M. (1998). Alternative Essences of Intelligence. AAAI/IAAI 1998, p. 961-968.

[7] Butcher, P. R., Kalverboer, A. f., & Geuze, R. H. (2000). Infants' shifts of gaze from a central to a peripheral stimulus: A longitudinal study of development between 6 and 26 weeks. *Infant Behavior and Development, 23,* 3-21.

[8] Cole, M., & Cole, S. (1996). *The development of children (3rd ed).* New York: Freeman.

[9] Corkum, V., & Moore, C. (1998). The origins of joint attention in infancy. *Developmental Psychology, 34,* 28-38.

[10] Courchesne, E. et al. (1994). Impairment in shifting attention in autistic and cerebellar patients. *Behavioral Neuroscience, 108*, 848-865.

[11] Dannemiller, J. L., & Stevens, B. K. (1988). A critical test of infant pattern preference models. *Child Development, 59,* 210-216.

[12] Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldi, J., & Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *Journal of Autism and Developmental Disorders, 28,* 479-485.

[13] Dayan, P, & Kakade S. & Read Montagne, P. (2000) Learning and selective attention. *Nature Neuroscience Supplement, 3,* 1218-1223.

[14] Deák, G. O., Fasel, I. R., & Movellan, J. R. (2001). The emergence of shared attention: Using robots to test developmental theories. In C. Balkenius et al. (Eds.), *Proceedings 1st International Workshop on Epigenetic Robotics: Lund University Cognitive Studies, 85,* 95-104.

[15] Deák, G. O., Flom, R., & Pick, A. D. (2000) Perceptual and motivational factor affecting joint visual attention in 12- and 18-month-olds. *Developmental Psychology, 36,* 511-523.

[16] DeCasper, A., & Fifer, W. (1980). Of human bonding: Newborns prefer their mothers' voices. *Science, 208,* 1174-1176.

[17] Dennis, W. (1973). *Children of the creche.* New York: Appleton-Century-Crofts.

[18] Doya, K. (2000). Complementary roles of basal ganglia and cerebellum in learning and motor control. *Current Opinion in Neurobiology, 10,* 732-739.

[19] Elman, J. L. (1990). Finding structure in time. *Cognitive Science, 14,* 179-211.

[20] Farroni, T., Johnson, M. H., Brockbank, M., & Simion, F. (2000). Infants' use of gaze direction to cue attention: The importance of perceived motion. *Visual Cognition, 7,* 705-718.

[21] Field, T.M., Cohen, D., Garcia, R., & Greenberb, R. (1984). Mother-stranger face discrimination by the newborn. *Infant Behavior and Development, 7,* 19-25.

[22] Hayhoe, M., Land, M., & Shrivastava, A. (1999). Coordination of eye and hand movements in a normal environment. Invest. Ophtalmol. & Vis. Sci., 40, S380.

[23] Hobson, P., Ouston, J., & Lee, A. (1988). What's in a face? The case of autism. *British Journal of Psychology, 79,* 441-453.

[24] Houk J.C., Davis J.L., Beiser, D.G. (1995). In *Models of Information Processing in the Basal Ganglia*. (pp. 215-232). Cambridge, MA: MIT Press.

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

[26] Jones, S. S. (1996). Imitation or exploration? Young infants' matching of adults' oral gestures. *Child Development, 67,* 1952-1969.

[27] Kaye, K. (1982). *The mental and social life of babies.* Chicago: University of Chicago Press.

[28] Kootz, J.P., Marinelli, B., & Cohen, D.J. (1982). Modulation of response to environmental stimulation in autistic children. *J. of Autism and Developmental Disorders, 12*, 185-193.

[29] Land, M.F., Mennie, N. & Rusted, J. (1999). *Eye movements and the roles of vision in activities of daily living: making a cup of tea.* Perception, 28, 1311-1328.

[30] Morales, M., Mundy, P. & Rojas, J. (1998). Following the direction of gaze and language development in 6-month-olds. *Infant Behavior and Development, 21,* 373-377.

[31] Movellan, J. R., & Watson J. S. (1987). Perception of directional attention. *Infant Behavior and Development: Abstracts of the Sixth International Conference of Infant Studies*. NJ, Ablex.

[32] Scaife, M., & Bruner, J.S. (1975). The capacity for joint visual attention in the infant. *Nature, 253,* 265-266.

[33] Schultz, W., Dayan, P., & Montague, P. R. (1997). A neural substrate of prediction and reward. *Nature, 275,* 1593-1599.

[34] Sutton R.S., & Barto, A.G. (1998). *Reinforcement Learning.* Cambridge, MA: MIT Press.

[35] Thelen, E., & Smith, L. B. (1994). A dynamic systems approach to the development of cognition and action. Cambridge, MA: MIT Press.

[36] Watson, J. (1972). Smiling, cooing, and "the game." *Merrill-Palmer Quarterly, 18,* 323-339.

[37] Watson, J. S., & Ramey, C. T. (1985). Reactions to response-contingent stimulation in early infancy. In J. Oates et al. (Eds.), *Cognitive development in infancy.* Hove, England: Erlbaum.

[38] Zlatev, J. and Balkenius, C. (2001). Introduction: Why 'epigenetic robotics?' In Balkenius et.al. (Eds.) *Proceedings 1st International Workshop on Epigenetic Robotics: Lund University Cognitive Studies, 85,* 1-5.