# Development of Socially Assistive Robots for Children with Autism Spectrum Disorders

David Feil-Seifer, USC Viterbi School of Engineering[1]
Matthew Black, USC Viterbi School of Engineering[2]
Elisa Flores, USC Keck School of Medicine[3]
Aaron St. Clair, USC Viterbi School of Engineering[4]
Emily Mower, USC Viterbi School of Engineering[5]
Chi-Chun Lee, USC Viterbi School of Engineering[6]
Maja Mataric', USC Viterbi School of Engineering[7]
Shrikanth Narayanan, USC Viterbi School of Engineering[8]
Clara Lajonchere, USC Keck School of Medicine[9]
Peter Mundy, UC Davis M.I.N.D. Institute[10]
Marian E. Williams, USC Keck School of Medicine[11]

[1] David Feil-Seifer, Dept. of Computer Science, University of Southern California, Henry Salvatori Computer Center #300, Los Angeles CA 90089-0781 USA, dfseifer@usc.edu, tel: 213 740 6245, fax: 213 821 5696

[2] Matthew Black, Dept. of Electrical Engineering, University of Southern California, EEB 400, 3740 McClintock Ave., Los Angeles CA 90089 USA, matthepb@usc.edu

[3] Elisa Flores, USC University Center for Excellence in Developmental Disabilities at Childrens Hospital Los Angeles, Los Angeles CA USA, elflores@chla.usc.edu

[4] Aaron St. Clair, Dept. of Computer Science, University of Southern California, Henry Salvatori Computer Center #300, Los Angeles CA 90089-0781 USA, astclair@usc.edu

[5] Emily Mower, Dept. of Electrical Engineering, University of Southern California, EEB 400, 3740 McClintock Ave., Los Angeles CA 90089 USA, mower@usc.edu

[6] Chi-Chun Lee, Dept. of Electrical Engineering, University of Southern California, EEB 400, 3740 McClintock Ave., Los Angeles CA 90089 USA, chiclee@usc.edu

[7] Maja Mataric' , Prof., Dept. of Computer Science, University of Southern California, Henry Salvatori Computer Center #300, Los Angeles CA 90089-0781 USA, mataric@usc.edu

[8] Shrikanth Narayanan, Prof., Dept. of Electrical Engineering, University of Southern California, EEB 400, 3740 McClintock Ave., Los Angeles CA 90089 USA, shri@sipi.usc.edu

[9] Clara Lajonchere, Asst Prof of Clinical Pediatrics, Keck School of Medicine, University of Southern California, 5455 Wilshire Bvd., Suite 2250, Los Angeles CA 90036-4234 USA, cmlajonc@agre.org

[10] Peter Mundy, Director of Educational Reserach, M.I.N.D. Institute, U.C. Davis, Sacramento, CA USA, pcmundy@ucdavis.edu

[11] Marian E. Williams, Asst Prof of Clinical Pediatrics, Keck School of Medicine, University of Southern California, Los Angeles CA USA, mwilliams@chla.usc.edu

# Development of Socially Assistive Robots for Children with Autism Spectrum Disorders

## Abstract

This paper describes the design and implementation of a robotic system for interaction with children with autism spectrum disorders (ASD). The system consists of a robot and supporting infrastructure, including a "smart room". The intent of this work is to explore and study the design of a therapeutic, minimally-restrictive environment that enables free-form human-human and human-robot interaction. Our primary design goals include: 1) using minimal structure with the participants in order to elicit natural behavior; 2) increasing a child's social interactions while utilizing minimal human-operated technology; and 3) facilitating human-robot interaction while requiring the child to wear no more than the minimum needed for effective signal detection. The robot system implemented in this study uses non-invasive methods for sensing and interpreting the child's behavior in order to interact with a child in free-form play while eliciting social behavior from the child. We present results from two feasibility studies with 12 children with ASD in order to validate the effectiveness of the robot system. We also discuss recommendations for the use of robot technology in ASD research settings.

Autism is a biologically-based disorder affecting social-communicative development (Dawson et al. 2002) characterized by symptoms often observable by 24 months of age, although the symptoms may present earlier (Stone, Coonrod, & Ousley, 1997), and currently affects between 1 in 150 children (CDC 2007). One of the primary social-communication disturbances characteristic of autism is an impairment in social attention and/or initiation of social behavior that manifests as a lack of spontaneous seeking to share experience, enjoyment, interests or achievements with other people (Mundy & Crowson, 1997). Children with autism have difficulties with both social interactions and verbal communication skills (Mundy and Crowson 1997). The paucity of these skills presents challenges in the implementation of existing behavioral interventions, since these interventions include reinforcing naturally occurring social behavior (Koegel, et al., 2003; Mundy, et al., 1997).

To date, the focus of much early intervention research in the autism literature has been to develop methods addressing early symptomatic impairments in joint attention (Kasari et al., 2006; Jones et al., 2006; Mundy & Crowson, 1997). Research in this area suggests that the early improvement of joint attention development in autism may ameliorate a negative developmental cascade of language and social-cognitive impairments that ensue when this pivotal arena of developmental disturbance is not effectively addressed (Kasari et 2006). However, interventions are still challenging to implement and do not result in improvements for children with autism. One explanation for this may be that children with autism are less intrinsically motivated to initiate social behaviors and therefore do not engage in treatment activities. Thus, one important component of an effective intervention for children with autism is a child-focused approach whereby adults identify the child's intrinsic interests in order to gradually engage them in interactions, which will maximize the child's attention (Kasari et al. 2006; Siller & Sigman, 2002).

Socially Assistive Robotics (SAR) seeks to provide assistance to users through social interaction (Feil-Seifer and Mataric' 2005), and is being studied for therapeutic use with children with ASD. It has been observed that children with ASD interact with robots differently than with people or toys and may show intrinsic interest in such machines. This intrinsic interest could be applied as a robot augmentation for an intervention for children with ASD. Werry, et al., (2001) conducted an experiment in which children with ASD and their teacher interacted with a robot and a non-animated toy. Behavioral coding of the experimental data demonstrated an increase in physical contact and eye gaze behavior with the robot during the robot condition when compared to the toy condition. Scassellati (2005) conducted an exploratory study with an expressive robot head that made pre-recorded facial expressions and sounds. It was reported that the children were engaged with the robot, demonstrated through smiling and making physical contact with behaviors not observed in the children's natural interactions with human interlocutors. Other work has shown that children with ASD will imitate a puppeteered robot (Robins 2005) more readily than with a human participant. These preliminary studies suggest that robots may act as intrinsically-rewarding social partners for children with autism.

It is important to identify and develop interactive platforms that are of intrinsic interest to children with ASD to further advance the development of effective intervention methods. Such platforms should be open to manipulation and modification in order to personalize their use with children across different developmental levels and severity of ASD symptoms. Robots and computers have been shown to have promise as potential assessment and therapeutic tools by providing real-time feedback. Due to their intrinsic value as interesting and rewarding entities, robots and computer characters can become a focus of shared attention and interaction with a teacher or parent. This effect has been observed in education tasks, such as for spelling word instruction (Grynszpan, et al., 2005), for facial and body imitation, (Duquette, et al., 2006; Kozima, et al., 2006), and for social and spatial skill development, such as exercises for joint attention development, or relative size discrimination (Lathan, et al., 2007). There are also scenarios where a robot has been used for a music therapy-type imitation game (Robins, et al., 2005). However, such studies have all largely used non-autonomous robot/computer systems, robots that are controlled by an operator. This poses a challenge for identifying what is responsible for the effects observed during the experiment. When a human puppeteer is controlling the behavior of a robot, and a benefit occurs, it's important to determine whether the effect is the result of the robot or the operator. Other systems have used significant use of sensors placed on objects in the environment (Merryman, et al., 2008) or the participants (Madsen, et al., 2008) in order to effectively be able to participate in the interaction. The design of a robot system that can autonomously interact in an experimental setting without requiring encumbering sensing to be placed in the environment or on the participants is a difficult goal.

The research we describe attempts to develop autonomous interactive robot experimentation platforms. The two preliminary studies presented in this paper develop autonomous and semi-autonomous robotic technologies and explore their potential for augmenting therapeutic interventions for children diagnosed with

ASD. We describe a robot system (robot and supportive off-board sensing and computing) designed for social interaction. We then present findings from two pilot experiments, with a total of twelve children (11 diagnosed with autism) who have varying levels of expressive language. In the experiments, the children freely interacted with different robot morphologies and robot social behaviors in order to evaluate how these design elements impact the social behavior of the children. Similar free-form play scenario techniques were used in work by Duquette., et al. (2006), Kozima, et al. (2006), and Salter, et al. (2007). Unlike these prior studies, our experiments use a mobile robot scenario where the robot has real-time knowledge of the relative locations of all people and objects in the room, necessary for exhibiting autonomous joint attention beyond a table-top (or otherwise stationary) scenario (Merryman, et al, 2008; Kozima, et al., 2006). This proxemic state information (distance and orientation between participants and the robot and/or objects in the scene as they interact) is a crucial first step for mobile embodied interaction between the child and the robot. When utilizing a mobile robot intended to exhibit and recognize joint attention behavior, knowledge of the locations of the interaction participants and objects of interest is necessary. The observed interactions during the studies were manually annotated with relevant social-communicative behavior codes to help quantify the implicit social role of the robot in the experimental settings. In the rest of this paper, we describe the development of the robot system, the supportive smart room system, and experimental design used for testing the robot system as a catalyst for social interaction in ASD.

## Robot Development

A free-form play scenario was chosen for its usability for children with varying levels of intellectual ability and verbal skill. In addition, the scenario had to involve reliable robot operation dependent on robust technologies; this excluded complex social behavior recognition and understanding. The intended role of the robot in these studies was as a catalyst for social interaction, both human-robot and human-human, thus aiding human-human socialization of ASD users, rather than as a teacher for a specific social skill. An important feature of the experiment design that reflects this role is that we always included at least one other person in the experimental setting (a parent, therapist, or experimenter) in order to allow for both human-human interaction (HHI) and human-robot interaction (HRI). In this section, the design of the physical robots and a description of the programmed behavior are presented.

Robot and toy designs

The experiment featured two different robot forms, a humanoid robot and a non-biomimetic robot. The humanoid robot's base consisted of a rectangular box mounted atop the standard Pioneer 2DX mobile robot base. A custom-designed humanoid torso robot was mounted on the box; the complete robot system weighed approximately 70 pounds and stood approximately three feet in height, with the head 2.5 feet (0.8 m) above the ground (see Figure 2, left). The robot's arms were capable of moving with similar, but lesser, articulation as human arms; the head could be tilted up and down, and the neck could move from side to side. Speakers mounted on the robot played pre-recorded non-verbal vocalizations to suggest an emotional state of the robot (e.g., excited, confused, encouraging). We combined head and arm movements with appropriate vocalizations to produce human-like gestures and actions (see Figure 3). Although the humanoid robot's face featured eyes and movable eyebrows and mouth, they were kept stationary in a happy/neutral face for this study. Two cameras were installed on the robot. One was secured to the base and directed upwards in order to capture the face of the child when positioned in front of the robot. The second camera was installed in one of the robot's eyes to provide a view from the robot's perspective. This second view can be used to determine whether human activity is affected by the robot's gaze, an area of future research that is specifically relevant when studying children with ASD. Two microphones were mounted below the arms of the robot, facing forward. A large button fashioned from a compact disc, was installed on top of the mobile base, which actuated a standard toy bubble-blower mounted on the front of the robot. A computer, also mounted on the robot base, connected wirelessly to a laptop computer in the observation room, so the robot could be monitored remotely while it was interacting with the child.

The non-biomimetic robot used a nearly identical hardware setup as the humanoid robot but without the humanoid torso mounted on top (see Figure 2, center). This robot was designed to have no anthropomorphized appearance characteristics; i.e., there were no human- or animal-like limbs or head-like features. The robot was equipped with the same type of onboard computer, button, bubble-blower, and speakers, which played the same sounds. In this way, we minimized the differences between the two robot morphologies to appearance only.

The non-mobile toy used for the control condition consisted of a rectangular box housing a computer. The same button used on the robots was mounted on top of the box, with the same bubble blower mounted on its front panel (see Figure 2, right). The control software actuated the bubble blower for a one- to two-second burst each time the button was pressed. During experiments, the box was placed along the wall opposite to where the parent was seated. This was the same position as where the robots were initially placed at the start of each experiment session.

Semi-autonomous robot control system

The robot control system for the contingent condition was semi-autonomous, relying in part on environment sensing in the smart room and in part on a human operator (who actively watched the interaction from the observation room). Robot autonomy is important in order to study how children with ASD are able to interact with robots that can be realistically engineered. Since the role of the robot in this experiment is part of a free-play scenario, the robot had to be able to correctly sense and interpret the child's activities, monitor the dynamics of the current social situation, and select and produce an appropriate behavior in response. We automatically tracked the position and movement of the child using an overhead camera mounted in the ceiling of the experimental room, which provided a birds-eye view of the entire room. Figure 5 is a snapshot of the overhead camera, used for person- and robot-tracking. We identified the location of the robot by tracking infrared markers mounted on the back of the robot. We tracked the child by identifying large regions in the image that were different from a stock background image (background subtraction) and using a pre-selected distinct color of the shirts worn by the child and the parent. Based on the video from the overhead camera, the child's position and movement (or lack thereof) relative to the robot's triggered programmed robot actions (see Table 1). During the experiment session, the robot's actions were triggered in real time using the Behavior-Based Behavior Intervention Architecture ($B^3$IA) specifically designed for the control of robots in behavioral experiments (Feil-Seifer & Mataric', 2008). $B^3$IA maps observed proxemic state (distance between people and the robot as they interact) and behavior of the user to robot actions executed in response.

The human operator could also trigger (or override) robot actions to ensure a reasonable flow of the interaction. This was especially important if there were problems with the automated sensing and/or in situations in which unforeseen and/or unpredictable child behaviors occurred during the ex[illegible]ents. As shown in Table 1, all robot actions based on child vocalizations were also handled by the operator, because real-time natural spoken language processing of children speech is challenging computationally expensive and requires hours of domain-matched training data to achieve re[illegible]ble accuracy levels. For this study, we did not take into account more complex social cues from the child (e.g., arm/body gestures, facial expressions, eye-gaze) when informing the robot's actions, because they are also very difficult to interpret by the robot and even by human observers. Thus the current implementation of the robot control system focused on a coarse analysis of proxemics (the measurable distances between people, or robots), as well as on limited operator-provided cues about the child's vocalization.

We incorporated a number of robot design elements to ensure the safety of the child and parent. First, we limited the speed of the robots to 2 feet/second (0.6 m/sec). At this speed, contact with the robot would be unlikely to result in injury. Second, we designed the robot's controller to avoid obstacles when moving, by using the overhead tracking system to maximize the robot system's field of view. Third, the operator in the observation room could press an emergency stop button at any point during the experiment to instantly shut down the robot's motors and cease all its movements.

## Procedures

### Description of experiment #1

The purpose of experiment 1 was to test the feasibility of the robot technology, and observe the interactions of children with the robot in order to identify hypotheses for further study and refinement of the technology.. In order to gauge the effectiveness of the social behavior of the robot, we collected video data from a small pilot experiment involving four children. This original study employed slightly earlier version of the non-biomimetic robot. The robot was similar to the one described above, but instead of a single button it featured two colored buttons that actuated the bubble-blowing mechanism (see Figure xx, left panel). The humanoid robot was a prototype of the upper-torso humanoid robot described above (see Figure xx, right). The experiment involved open-ended interaction between the child and the robot, giving the child complete freedom to move about the experimental room.

A set of three conditions were presented to the participating children, in randomized order; thus each child participated in three play sessions. Two of the conditions employed the non-biomimetic robot. In the first, contingent condition, the robot blew bubbles when the child pushed one or both of the buttons. In the second condition, the random condition, the robot blew bubbles at random intervals, not in response to the child's behavior. The final conditions employed the humanoid robot. It was used in a "Simon-Says" scenario; the robot made an exaggerated pose and verbally encouraged the child to imitate, as in the game Simon-Says. In this early pilot experiment, the robot had no way to detect if the child was interacting and imitating. After 2-6 seconds, the robot made a new pose.

Four boys diagnosed with ASD between the ages of 5 and13 (mean age 7.75 years old) participated in this experiment. The child, the robot, the parent, and an experimenter were all present in the room during the testing session. One child could not participate in all the experimental conditions due to a malfunctioning bubble blower.

The sessions were videotaped from a single camera angle behind the robot. The recorded video was then coded for speech (including who the speech was directed at), social orientation (the child moving in front of the parent or robot), button-pushes, and how often the child moved toward or away from the robot.

We observed that the children behaved differently in the contingent and random conditions. One of the children with autism actively engaged the robot during the contingent condition (and also interacted proactively with his parent). In contrast, during the random condition he sat back against the well and did not interact with his parents nor with the robot until the robot blew bubbles; then he moved to pop them and then leaned against the wall again. Similar behavior was consistently observed from for the two other children with ASD that could fully participate. In the humanoid condition, the robot speech was unsettling to two children with ASD, while it was encouraging to one, demonstrated by either placing hands over the ears, or asking the robot to stop talking. From the coded data, we observed that during the contingent condition, the children exhibited greater vocalizations, spoke more to their parents, and exhibited more directed actions (moving toward the robot, orienting themselves in front of the robot or the parent, speaking to the robot or the parent) than in the random condition (see Figure yy).

Results from this pilot study supported our initial hypothesis that the robot's behavior influenced the behavior of the child, and that the bubble-blowing was enjoyable and interesting for our sample of participating children. The children would show this verbally, and by repeatedly interacting with the bubble blower and encouraging their parent to interact with it. We noticed that the children who were less verbal or who were poor initiators of social behavior would repetitively press the buttons on the robot more mechanically than socially. In addition, we observed that the movement of the robot encouraged the children to move, while they tended to be still when the robot was not itself moving.

To better understand which social behaviors are affected, we redesigned the robot and experiment scenario based on feedback from this pilot study.

In particular, we decided to have the robot recognize more social behavior than merely button pushes. We decided to use automatically recognized proxemic information and operator recognized vocalization information to augment the sensing capabilities of the robot, so that the robot could react to other behavior of the child. In addition, we decided to implement more complex social behavior than merely blowing bubbles by using pre-recorded non-verbal vocalizations to express acknowledgement, approval, and disappointment, and added the humanoid upper torso, Bandit, for human-like gesture expression.

### Description of experiment #2

In this experiment, each child participated in up to four conditions/play sessions within a single visit to an outpatient clinic. During an initial introductory period, a clinical psychologist greeted the child and parent and reviewed informed consent and assent forms while the child became adjusted to the clinic office. The child and parent then put on colored shirts, which enabled automated visual data collection and real-time person tracking using environment sensing while in the experiment room (dimensions 9 feet by 12 feet by 10 feet high). Each play session consisted of a flexible-length introductory period. During the introductory period the therapist demonstrated the physical capabilities of the robot or toy, such as sample arm and head gestures and the use of the button to trigger the bubble-blowing behavior, and instructed the child to be gentle with the robot. The parent was asked to remain seated in a corner of the room, but to participate if requested by the child. The therapist left the room after orienting the child and allowed the child to play for five minutes without interruption.

The therapist observed the free-play session from the observation room (see Figure 1) separated by a one-way mirror. If the child became upset, anxious, or ceased interacting, all robot motion was stopped and the therapist reentered the experiment space and attempted to calm the child and reengage him with the robot, if possible. If this was not possible, the therapist removed the robot from the room. Each session was separated by a break during which the parent and child left the room and the experimenters prepared the next condition (the order of conditions was randomly assigned for each child). The study was designed to compare the effects of different robot forms and behaviors and to explore how children respond to the robotic compared to a toy.

The study participants included eight boys between the ages of 5 and 9 years (mean age = 7.2 years). All children had been diagnosed with autism using results from the Autism Diagnostic Observation Schedule and the Autism Diagnostic Interview-Revised (Lord, et al., 1994). The children's expressive language was categorized by psychologists based on the children's verbalizations observed both in the waiting room and during the experiment. When describing the results of the study, children were grouped into two categories: 1) "verbal" children were those (five) who were categorized as having sentences and complex language, and 2) "less verbal" children were those (three) who were preverbal or used word combinations.

## Experiment conditions

This exploratory study employed a comparison of three conditions, a mobile humanoid robot, a mobile non-biomimetic robot, and an immobile toy, in an effort to better understand the role of both the form and behavior of a robot in a therapeutic setting (see Figure 2). The original study design planned to compare the humanoid robot to the immobile toy only. The non-biomimetic robot was introduced for the final three children after two of the first five children reported that they were frightened by the appearance of the humanoid robot.

In addition to comparing the robots with a control toy, we also compared two different behavior strategies for the robots: random and contingent. In the random condition, the robot (i.e., either of the two forms) moved about the room without regard to the child's position and behavior, except to avoid collisions, and paused occasionally to turn in place. The robot autonomously blew bubbles and made randomly selected sounds at random intervals (typically about 5-10 times per minute). In random condition sessions employing the humanoid robot, arm gestures and head movements were also executed randomly, not in response to the child's location or behavior.

In contrast, the contingent condition was designed to encourage social actions from the child and make the robot behave more like a social partner to the child. Unlike the first pilot study, the contingent condition in the second study involved more complex robot actions (not just blowing bubbles). Table 1 lists the various contingent actions of the robot in response to the child's relative location (see Figure 4) and behavior. These actions were triggered by our semi-autonomous robot control system, described below.

### Participant selection

Participants were recruited from Autism Speaks' Autism Genetic Resource Exchange (AGRE) (Geshwind, et al., 2001). The AGRE program provides biomaterials and phenotype and genotype information of families with two or more children (multiplex) with ASD to the scientific community. Participants were eligible for inclusion in the study if they had been diagnosed with autism by AGRE researchers, were between the ages of five and ten, had verbal ability above 2.0 years of age on the communication subscale of the Vineland Adaptive Behavior Scale (Sparrow, et al., 1984), or were evaluated using either modules two or three of the Autism Diagnostic Observation Schedule (Lord, et al., 2001), and lived in the greater Los Angeles area. Fliers were sent to eligible families. Interested families responded to the call for participation by phone and email. Of the 70 families who were sent a flier, 13 responded (18.5%), and 13 children participated in the pilot study from 10 families. Because the AGRE database is made up of multiplex families, there were three families where two brothers were recruited for the study.

### Behavioral coding

In addition to the overhead camera, we installed several other audio-video sensors in the experimental room and on the robots, based on smart room set-ups (Busso, et al., 2005). Specifically, we installed four cameras in each corner of the room, positioned so as to capture the dynamics of the interaction from multiple angles (Figure 6). Each camera was paired with a high-quality condenser microphone, mounted near the camera to record the sound from that corner of the room.

Videorecorded data from the child-robot interactions were manually coded in multiple ways. We recorded the duration of each trial and the reason the session ended (time expired, technical malfunction, or child withdrawal). Two engineering graduate students used the Anvil software (Kipp 2001) to code the following variables: frequency and duration of child's turning to look at the robot or toy; fraction of time child spent proximal (within an arm's distance) to the robot/toy; verbal/vocal behavior directed to/about the robot/toy; verbal/vocal behavior directed to/about the parent; frequency of anthropomorphic verbalizations (child verbalized something while looking at the robot/toy that had anthropomorphic content).

## Results

### Overview

Table 2 lists the ages of all eight children in the study, along with the sessions in which they participated. The children showed a range of overall responses to the robot interactions. Fifty percent (4 out of 8) of the children engaged in the free-play scenario with at least one of the robot forms. The other 50% became uneasy or upset and withdrew from participation before the planned end of the experiment (see Table 2). For those children who became upset or frightened, there seemed to be idiosyncratic reasons depending on each child. For example, the robot's movement, the appearance of its face, the sounds it made (both motor sounds and pre-recorded human sounds), and its unexpected actions were all factors that influenced the participant's unease.

### Robot & toy operation evaluation

There were very few technical problems with the toy condition. It ran out of bubble solution at the very end of one session, and button-presses were occasionally not recognized resulting in a bubble-blower that did not activate. These problems were minor and infrequent. Both robot designs running in the random condition also performed as expected, successfully using the overhead camera to avoid obstacles and otherwise behave randomly.

We encountered difficulties during the contingent robot conditions, which made the robot appear less contingent than intended. First, although the tracking system worked the majority of the time, there were cases in which the robot's infrared markers were blocked or the child was too close to the robot and was misinterpreted as part of the robot. The marker problem occurred intermittently with one child, and the proximity problem occurred intermittently with two more children. This resulted in the vision system believing that the robot or child was elsewhere in the room, which then triggered the wrong robot action based on this false information. Such issues can be mitigated in future experiments, through improved filtering.

A more complex challenge arises from the children's expectations of the robot. The children with complex expressive language spent long periods speaking to the robot while standing in front of it. The robot, in turn, repeatedly blew bubbles, saying “woohoo!” and danced in place (see Table 1), as it was programmed to do. However, since the children expected the robot to respond to what they were saying (which it could not do, due to the technological limitations explained above), the robot's responses did not appear contingent to those children even if they were correct based on its programming. Finally, due to time delays between the child’s actions and the robot’s responses, many of the robot’s actions did not appear to be based upon the immediate behavior of the child. The cumulative effect of these issues may have caused the differences between the contingent and random conditions, resulting in them becoming inconclusive.

The only potential experimental safety concerns resulted from two cases of unwanted contact between the robot and the wall, both resulting in no damage to any person or the robot. One collision was caused by a hardware failure with the wireless connection between the laptop computer in the observation room and the on-board computer on the robot (and was the cause for one premature ending to the session as shown in Table 2). The other contact was due to the real-time robot-tracking lagging behind by a couple seconds. This problem was resolved by dividing the computational load between two computers in the observation room.

## Discussion and Future Work

The described feasibility studies provide insights and recommendations for further refinement of the experimental design. In summary, we found that different forms of robots can interact safely with children on the autism spectrum and that more verbal children showed more interest in interacting with the robots but also had more expectations for the social capabilities of the humanoid robot, which were not matched by the robot’s technological abilities.

The overhead camera system that is presented addresses many of the [illegible]nomy needs of the system while not requiring the participants to be encumbered with more than a t-shirt. The cameras also provided a real-time view of all relevant proxemic information, critical to effective interaction in a mobile context. However, there still exists a large trade-off between fully autonomous systems where it is possible to isolate the effects of the robot from the potential confounding effects due to the operator, and the controllability of experiments using an operator-controlled robot. Our work shows how minimal operator intervention can be used to preserve largely autonomous behavior, while still maintaining adequate safety and experiment control standards.

This study also elucidated the importance of correctly managing child expectations and the need for the enhancement of the complexity of the robot's social behaviors. Several children in this study quickly developed expectations of the robot. For example, one child tried to encourage the robot to wave by telling it to wave. Another child demonstrated a wave and expected the robot to imitate. Still another child told the robot to play "tag" and tried to explain and demonstrate how to play. Thus, some children assumed that the humanoid robot was capable of verbal understanding and fluent verbal communication and game-playing. When the robot was not able to meet those expectations, the children became disappointed. The expectations expressed by the study participants will be valuable for improving the interactive abilities of the robot for future ASD research. As a first step, we have revised our orientation session in which the therapist introduces the robot to the child, so that we will provide a more detailed representation and demonstration of the robot's capabilities to both the child and the parent in order to better manage expectations.

Based on the data recorded from these sessions, we developed several guidelines for observing expressive behavior using a multi-camera setup. Working from the angles shown in Figure 6, it was clear that the top two views (the eye-level views) were usable, but the bottom two view (the overhead views) were not. Ideally, we would mount a camera behind the robot able to see in front of the robot and some of the humanoid robot's head. This will give a first-person view from the robot's perspective, though perhaps the view from the robot's camera would be usable. Additionally, a similar eye-level view from behind the parent would provide a good view for observing social behavior. Finally, a movable human-operated camera capable of zooming in on the child's face would provide crucial information about affect.

There was also useful feedback from our team regarding the behavior of the robot. First, rather than programming the robot to move freely about the room, or follow the child, we should consider making the robot semi-stationary, turning in place and moving back-and-forth. Bubbles should not be used as a reward for vocalizations, the duration of this action (6 seconds) is too long. Instead, the robot should use light and tone combinations to acknowledge vocalizations. Bubbles should instead be used to reward smiles and laughter (promote positive affect). Finally, the robot's facial affect should be positive, not neutral, since the neutral affect seemed to be distressing to some children.

The findings in this study lend support to the notion that for robotics technology to be useful in of the development of interventions for ASD, guidelines for recording joint attention, counting child initiated verbalizations and child initiated social interactions, are required for mobile scenarios such as the presented free-play scenario. Our experiences with the camera coverage used the described smart room setup showed the importance of appropriate camera-placement in order to capture events as the robot and participants move about the room.

While several advances in human-robot interaction technology were noted in this study, there were also multiple challenges and limitations that were addressed. In particular, we found that the robot behavior constructed in a lab setting was not necessarily interpreted by children as intended during the experiment. In addition, the dual but potentially contradictory goals of (1) autonomous robot behavior and (2) designing a well-controlled experiment, can lead to long development times. There exists a potential trade-off between a well-controlled experiment defined by a ideally robot behaving as intended 100% of the time and an experimental design that incorporates an operator that can initiate or override autonomous robot behavior at key times during the experimental interaction. Finally, perhaps the most difficult challenge for the technology is to correctly recognize and interpret user behavior, even for simple interactions, yet this is [illegible]sary for sustaining engaging human-robot interaction even during short experiment sessions. This is an open challenge for human-robot interaction in general and a topic of active ongoing research world-wide.

In the presented work, the robot maintained a fixed positive expression on its face during the entirety of the interaction. Our future work includes dynamic facial affect manipulation to study the effects of facial emotional behavior on the social interaction patterns of the child participants. The robotic facial configuration allows us to display emotions ranging from happiness to anger, sadness, and fear. We hypothesize that creating emotional social facial gestures will have differing levels of effect depending on the child's level of functioning.

Future work will also explore the effect of facial and verbal emotional mismatch on the interaction patterns of children with autism.

The technology advances presented in this paper demonstrate how a semi-autonomous robot can be used in an experimental setting in the ASD context. In particular, this study shows that without needing to encumber the user, the robot is able to maneuver safely and interpret proxemic information along with minimal operator-provided input to interact in a free-play setting. However, the significant challenge presented by the children's expectations of the robot's social ability must be addressed in order to encourage sustainable human-robot interaction toward uses in augmenting ASD therapy.

## References

1. CDC, Centers for Disease Control and Prevention. (2007). Prevalence of Autism Spectrum Disorders. Morbidity and Mortality Weekly Report. 54. (SS-1). 12-28. Feb.
2. Chakrabarti, S. and Fombonne, E. (2005). Pervasive developmental disorders in preschool children: confirmation of high prevalence. American Journal of Psychiatry. 162. (6). 1133-1141.
3. Dawson, G., Webb, S., Schellenberg, G., Dager, S., Friedman, S., Ayland, E., et al. (2002). Defining the broader phenotype of autism: Genetic, brain, and behavioral perspectives. *Development and Psychopathology, 14*, 581-612.
4. David J. Feil-Seifer and Maja J. Matarić, (2005). "Defining Socially Assistive Robotics," In *International Conference on Rehabilitation Robotics*, pp. 465-468, Chicago, IL.
5. David J. Feil-Seifer and Maja J. Matarić, (2008). "B$^3$IA: An architecture for autonomous robot-assisted behavior intervention for children with Autism Spectrum Disorders," In *IEEE Proceedings of the International Workshop on Robot and Human Interactive Communication*, Munich, Germany.
6. A. Duquette, and H. Mercier, and F. Michaud. (2006). Investigating the use of a mobile robotic toy as an imitation agent for children with autism. Proceedings International Conference on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems. Paris, France.
7. O. Grynszpan and J. Martin and J. Nadel. (2005) Human Computer Interfaces for Autism: Assessing the Influence of Task Assignment and Output Modalitie. Proceedings of the ACM Conference on Human-

Robot Interaction. Portland, OR.

8. D. Geschwind, J. Sowinski, C. Lord, P. Iverson, L. Ducat, and S. Spence. (2001). The Autism Genetic Resource Exchange: A Resource for the Study of Autism and Related Neuropsychiatric Conditions. American Journal of Human Genetics. 69. 461-463.
9. Kasari, C., Freeman, S., & Paparella, T. (2006). Joint attention and symbolic play in young children with autism: a randomized controlled intervention study. *Journal of Child Psychology and Psychiatry, 47*, 611-620.
10. M. Kipp. (2001). Anvil - A Generic Annotation Tool for Multimodal Dialogue. Proceedings of the 7th European Conference on Speech Communication and Technology (Eurospeech). Aalborg.
11. Koegel, L., Carter, C., & Koegel, R. (2003). Teaching children with autism self-initiations as a pivotal response. *Topics in Language Disorders, 23*, 134-145.
12. H. Kozima and C. Nakagawa. (2006). Interactive robots as facilitators of children's social development. Mobile Robots: Toward New Applications. Pro Literatur Verlag.
13. C. Lathan, K. Boser, C. Safos, C. Frentz, and K. Powers (2007). Using Cosmo's Learning System (CLS) with Children with Autism. Proceedings of the International Conference on Technology-Based Learning with Disabilities. Dayton, OH.
14. C. Lord, S. Risi, L. Lambrecht, E. H. Cook Jr., B. L. Leventhal, P. C. DiLavore, A. Pickles, and M. Rutter. (2001). The Autism Diagnostic Observation Schedule-Generic: A Standard Measure of Social and Communication Deficits Associated with the Spectrum of Autism. *Journal of Autism and Developmental Disorders*. 30:3. 205-223.
15. C. Lord and M. Rutter and A. Le Couteur. (1994). Autism Diagnostic Interview-Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. Journal of Autism and Developmental Disorders. 24. (5). 659-685. Oct.
16. Madsen, Miriam and el Kaliouby, Rana and Goodwin, Matthew and Picard, Rosalind. (2008). Technology for just-in-time in-situ learning of facial affect for persons diagnosed with an autism spectrum disorder. Assets '08: Proceedings of the 10th international ACM SIGACCESS conference on Computers and accessibility. New York, NY, USA.

17. J. Merryman and A. Tartaro and M. Arie and J. Cassell. (2008). Designing Virtual Peers for Assessment and Intervention for Children with Autism. Conference on Interaction Design for Children: Children with Special Needs. Chicago, Il.
18. Mundy, P. (1995). Joint attention and social-emotional approach behavior in children with autism. *Development and Psychopathology, 7,* 63-82.
19. Mundy, P., & Crowson, M. (1997). Joint attention and early social communication: Implications for research on intervention with autism. *Journal of Autism and Developmental Disorders, 27,* 653-676.
20. Robins, B. and Dautenhahn, K. and Boekhorst, R.T. and Billard, A.. (2005). Robotic assistants in therapy and education of children with autism: can a small humanoid robot help encourage social interaction skills? Universal Access in the Information Society. 4(2). 105-120.
21. T. Salter, F. Michaud, D. L'etourneau, DC, Lee, and Werry, IP. (2007) Using proprioceptive sensors for categorizing interactions. In Proc. Human-Robot Interaction. Washington, D.C.
22. M. Siller and M. Sigman. (2002). The behaviors of parents of children with autism predict the subsequent behavior of the children's communication. J. Autism and Developmental disorders. 32. (2). 77-87.
23. Sparrow, S.S., Balla, D.A. and Cicchetti, D.V. (1984). Vineland adaptive behavior scales. Test Corp of Amer.
24. Stone, W., Coonrod, E., & Ousley, C., (1997). Brief Report: Screening tool for autism in two year olds (STAT): Development and preliminary data. *Journal of Autism and Developmental Disorders, 30*, 607-612.
25. B. Scassellati. (2005). Quantitative metrics of social response for autism diagnosis. In IEEE International Workshop on Robots and Human Interactive Communication (ROMAN), pages 585–590, Nashville, TN.
26. I. Werry, K. Dautenhahn, and W. Harwin. (2001). Investigating a robot as a therapy partner for children with autism. In Proceedings of the European conference for the advancement of Assistive Technology (AAATE), pages 3–6, Ljubljana, Slovenia.

*Figure 1*: Picture of the experimental room from the observation room through the one-sided window

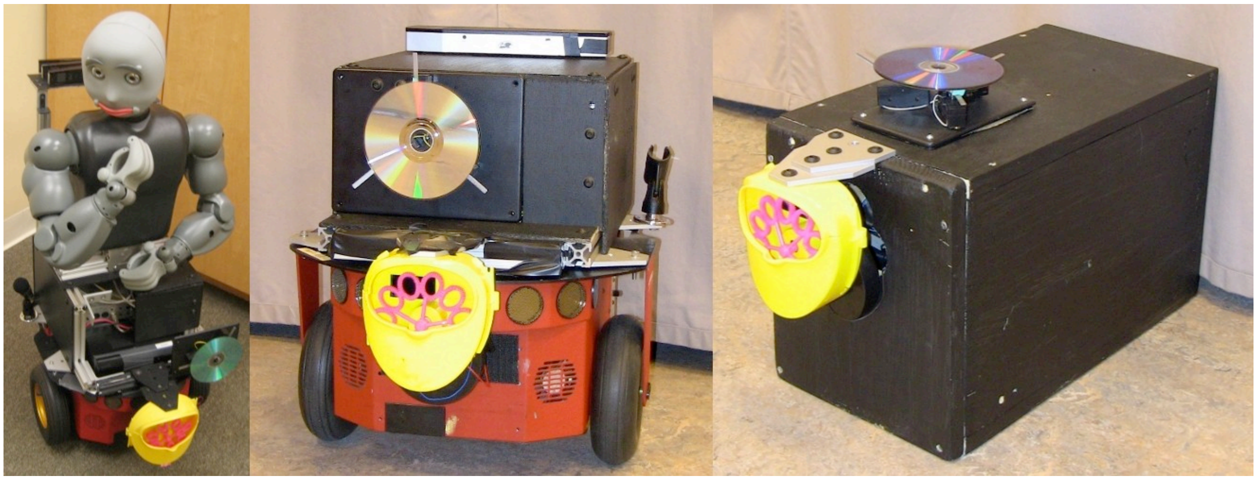

*Figure 2*: Pictures of the humanoid robot (left), non-biomimetic robot (center), and toy (right).

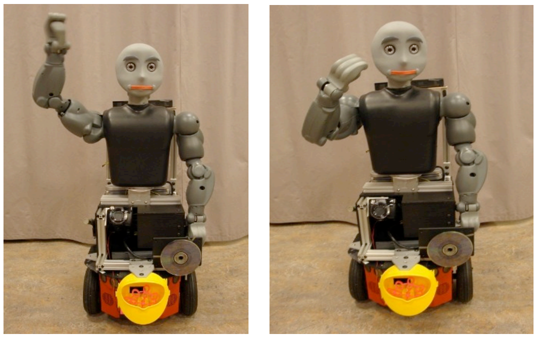

*Figure 3*:  Two examples of humanoid robot gestures: waving (left) and “come here” (right).

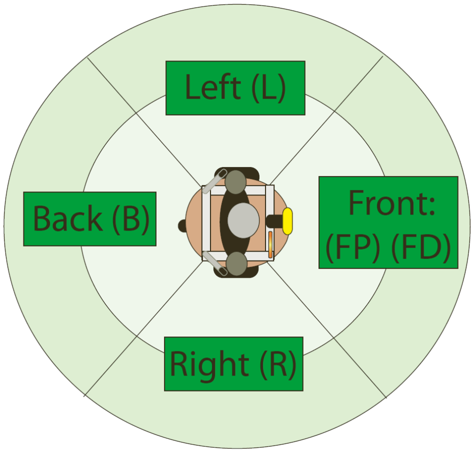

*Figure 4*: Child position relative to the robot was detected by the overhead camera and automatically discretized into 5 disjoint regions: front proximal (FP), front distal (FD), left (L), right (R), and behind (B). The threshold between the proximal and distal regions was heuristically set to 1m.

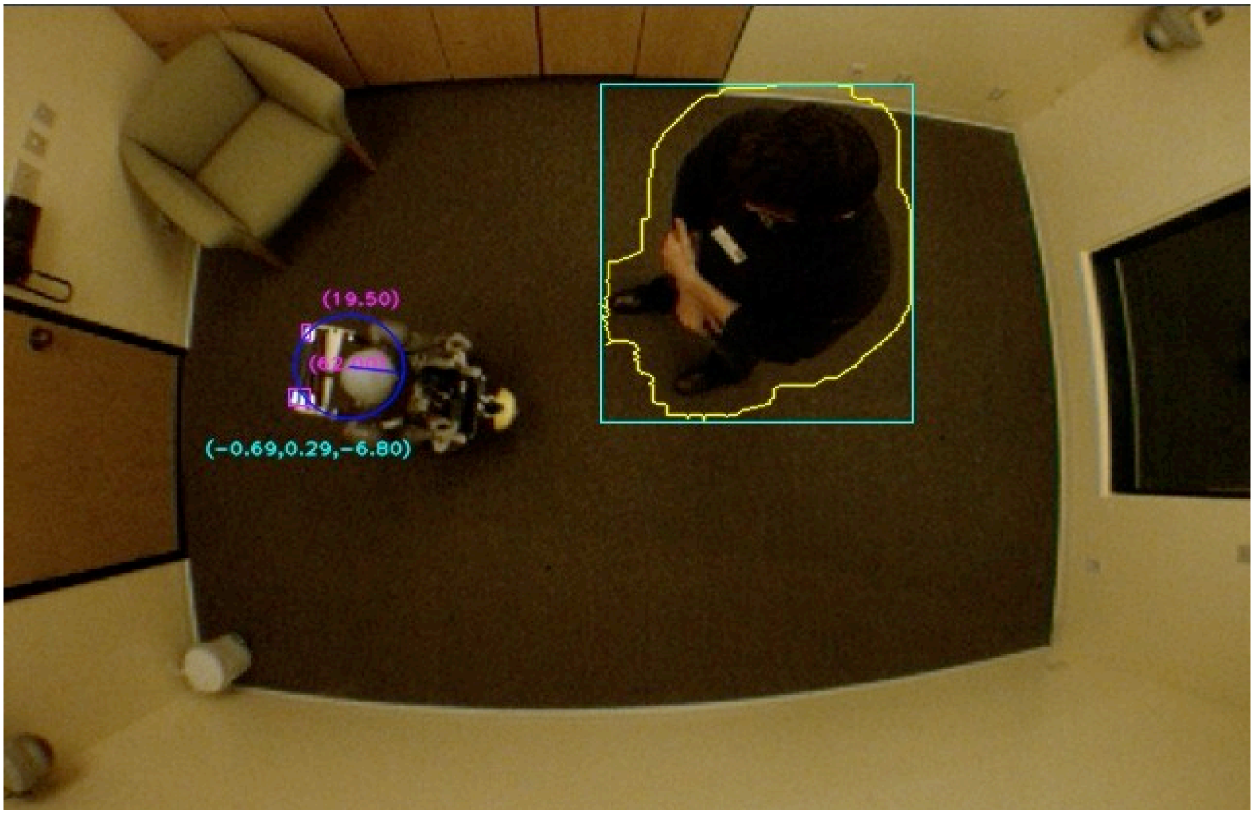

Figure 5: View from the overhead camera showing the computed positions of the robot and person.

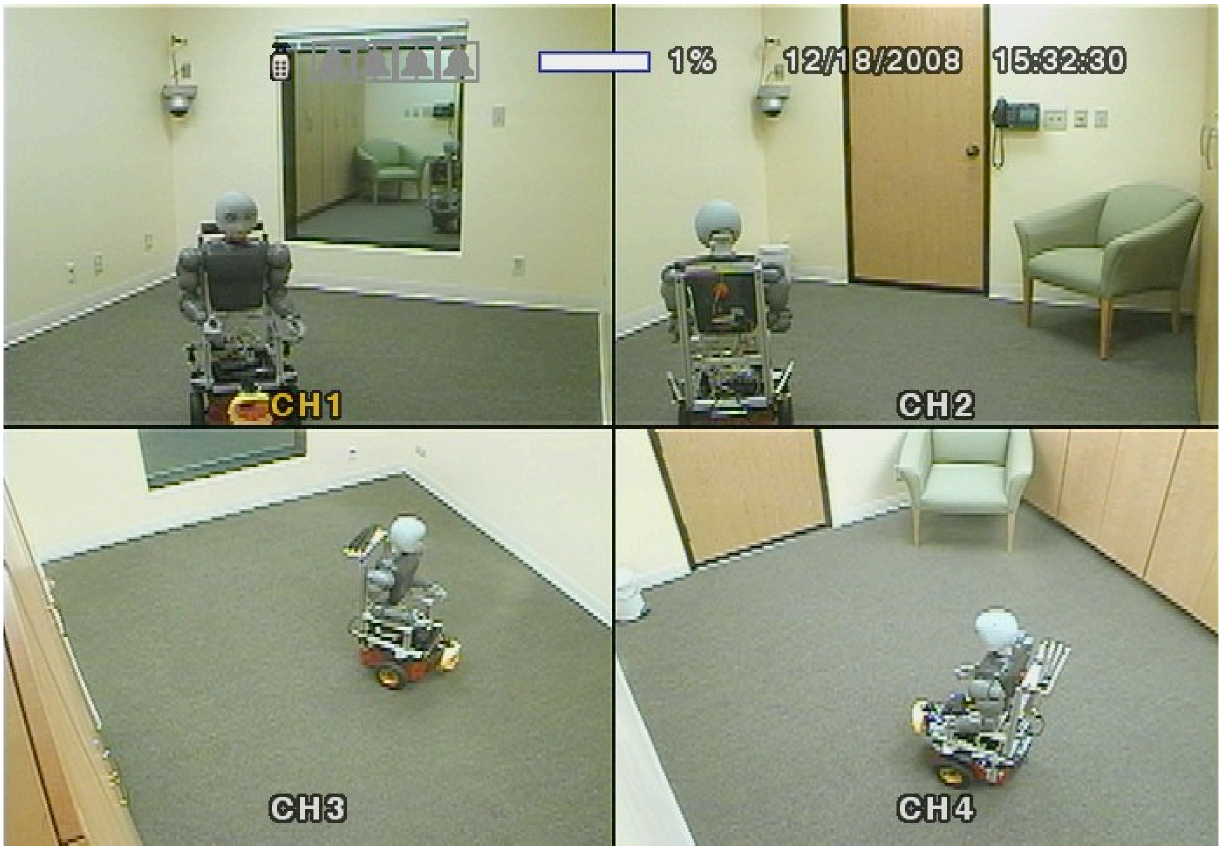

*Figure 6*: Screenshot of the cameras in the four corners of the experimental room. Cameras 1 and 2 (top) are five feet off the ground, and cameras 3 and 4 (bottom) are ten feet off the ground.

Figure xx, Left: The bubble-blowing robot used for experiment #1 Right: The humanoid robot used for experiment #1

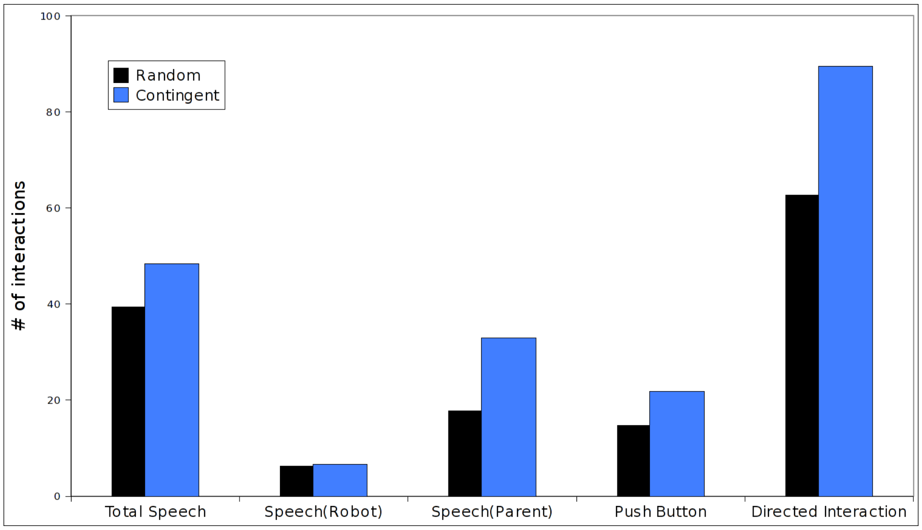

Figure yy: Results for experiment #1

| *Child behavior* | *Automatic sensing or human operated* | *Child location* | *Robot action(s)* |
|---|---|---|---|
| Null | Automatic | FP | 1) head nod; 2) head nod with encouraging vocalization; 3) "come here" gesture; 4) "come here" with encouraging vocalization; 5) "rocking" forward and backward motion |
| | | FD | approach child |
| | | L/R | turn, face child |
| | | B | 1) "rocking" motion, confused vocalization;<br>2) "look around" head movements |
| Approach | Automatic | FP/FD | nod head, approval vocalization |
| | | L/R/B | turn, face child |
| Retreat | Automatic | FP/FD | bow head, "sigh" vocalization |
| | | L/R/B | turn, face child, bow head, "sigh" vocalization |
| Vocalization | Human | FP | blow bubbles, "woohoo!" vocalization, dancing movement |
| | | FD | wave |
| | | L/R | turn, face child, nod head |
| | | B | turn, face child, "huh?" vocalization |
| Button press | Automatic | FP/FD/L/R/B | blow bubbles, "woohoo!" vocalization, dancing movement |

*Table 1*: Semi-autonomous robot control system used in the contingent robot condition. When more than one robot action is listed (Null child behavior and FP/B child locations), the actions cycle every ten seconds until the child's behavior and/or location changes.

| *Child* | *Age* | *Robot/Toy* | *Contingency* | *Session Length (sec)* |
|---|---|---|---|---|
| 1 | 9 years 5 months | Humanoid robot | Contingent | 343 |
| | | Humanoid robot | Random | 273 |
| | | Toy | Contingent | 309 |
| 2 | 9 years 5 months | Toy | Contingent | 289 |
| | | Humanoid robot | Random | 297 |
| | | Humanoid robot | Contingent | 334 |
| 3 | 5 years 10 months | Humanoid robot | Contingent | 326 |
| | | Toy | Contingent | 303 |
| | | Humanoid robot | Random | 320 |
| 4 | 8 years 11 months | Humanoid robot | Random | 318 |
| | | Humanoid robot | Contingent | 0[#] |
| | | Toy | Contingent | 289 |
| 5 | 7 years 1 month | Toy | Contingent | 305 |
| | | Humanoid robot | Contingent | 67[#] |
| 6 | 6 years 7 months | Humanoid robot | Contingent | 325 |
| | | Toy | Contingent | 315 |
| | | Non-biomimetic robot | Contingent | 48[#] |
| 7 | 6 years 4 months | Non-biomimetic robot | Contingent | 244[*] |
| | | Toy | Contingent | 295 |
| | | Humanoid robot | Random | 292 |
| | | Humanoid robot | Contingent | 299 |
| 8 | 5 years 4 months | Humanoid robot | Random | 0[#] |
| | | Toy | Contingent | 297 |
| | | Non-biomimetic robot | Contingent | 307 |
| | | Non-biomimetic robot | Random | 299 |

* Robot malfunction that ended the session prematurely

# Child withdrawal before or during the session

*Table 2*: Ages of the children and a description of the sessions in which each child participated. Of the eight participants, there were three sets of brothers (1-2, 4-5, 7-8).