



# HHS Public Access

Author manuscript

*Autism*. Author manuscript; available in PMC 2015 April 04.

Published in final edited form as:

*Autism*. 2014 July ; 18(5): 598–608. doi:10.1177/1362361313479454.

## Pilot clinical application of an adaptive robotic system for young children with autism

**Esubalew Bekele, Julie A Crittendon, Amy Swanson, Nilanjan Sarkar, and Zachary E Warren**

Vanderbilt University, USA

### Abstract

It has been argued that clinical applications of advanced technology may hold promise for addressing impairments associated with autism spectrum disorders. This pilot feasibility study evaluated the application of a novel adaptive robot-mediated system capable of both administering and automatically adjusting joint attention prompts to a small group of preschool children with autism spectrum disorders ( $n = 6$ ) and a control group ( $n = 6$ ). Children in both groups spent more time looking at the humanoid robot and were able to achieve a high level of accuracy across trials. However, across groups, children required higher levels of prompting to successfully orient within robot-administered trials. The results highlight both the potential benefits of closed-loop adaptive robotic systems as well as current limitations of existing humanoid-robotic platforms.

### Keywords

autism spectrum disorder; joint attention; robotics; technology

### Introduction

With an estimated prevalence of 1 in 88 (Centers for Disease Control and Prevention (CDC), 2012), effective early identification and treatment of autism spectrum disorders (ASDs) is a pressing clinical care and public health issue. The costs of ASD are thought to be enormous across the life span, with recent individual incremental lifetime cost projections exceeding US\$3.2 million (Ganz, 2007; Peacock et al., 2012). To address the powerful impairments and costs associated with ASD, a wide variety of potential interventions have been offered. The cumulative literature suggests substantial benefits of early, intensive, ASD-specific interventions; however, outcomes vary greatly for individuals, this variation is poorly understood, and most individuals continue to display potent impairments in many areas despite significant improvements (Howlin et al., 2009; Rogers and Vismara, 2008; Warren

---

© The Author(s) 2013

**Corresponding author:** Zachary E Warren, Vanderbilt Kennedy Center Treatment and Research Institute for Autism Spectrum Disorders (TRIAD), Vanderbilt University, PMB 74, 230 Appleton Place, Nashville, TN 37203, USA.  
zachary.e.warren@vanderbilt.edu.

#### Declaration of conflicting interests

The content is solely the responsibility of the authors and does not necessarily represent the official view of the National Institutes of Health (NIH).

et al., 2011). Given the present limits of intervention science and the powerful nature of early impairments across the life span, there is urgent need for the development and application of novel treatment paradigms capable of substantially more efficacious individualized impact on the early core deficits of ASD. Given rapid progress and developments in technology, it has been argued that specific computer- and robotic-based applications could be effectively harnessed to provide innovative clinical treatments for individuals with ASD (Goodwin, 2008).

A growing number of studies have investigated the application of advanced interactive technologies to ASD intervention, including computer technology (Goodwin, 2008), virtual reality (VR) environments (Bellani et al., 2011), and more recently, robotic systems (Diehl et al., 2012). Advances in robotic technology have certainly demonstrated the capacity for intelligent robots to fulfill a variety of human-like and neuro-rehabilitative functions in other populations (Dautenhahn, 2003), but well-controlled research focusing on the impact of specific clinical applications for individuals with ASD is very limited (Diehl et al., 2012). The most promising finding regarding robotic interaction to date has been a documented preference by some individuals with ASD, in certain circumstances, for technological interaction versus human interaction. Specifically, data from several research groups have demonstrated that many individuals with ASD show a preference for robot-like characteristics over non-robotic toys and humans (Dautenhahn and Werry, 2004; Robins et al., 2006) and, in some circumstances, even respond faster when cued by robotic movement than human movement (Bird et al., 2007; Pierno et al., 2008). While this research has been accomplished with school-aged children and adults, research noting the preference for very young children with autism to orient to nonsocial contingencies rather than biological motion suggests that downward extension of this technological preference may in fact be more salient and hold great potential for utilization of intervention paradigms (Annaz et al., 2012; Klin et al., 2009).

Despite suggested benefits of robotic technology for individuals with ASD, both potent methodological and existing system limits present challenges to understanding feasibility and ultimate clinical utility. To date, there have been very few applications of robotic technology for teaching, modeling, or facilitating interactions through directed intervention and feedback approaches (Diehl et al., 2012). In the only identified study in this category to date, Duquette et al. (2008) demonstrated improvements in affect and attention sharing with co-participating partners during a robotic imitation interaction task using a simplistic robotic doll. The study paired two children with robot mediator and another two with human mediator. Shared attention and imitation were measured using visual contact/gaze directed at the mediator, physical proximity, as well as rated facial expressions and gesture imitations. Although this study was conducted using a comparison group, the very small sample size, pre-programmed behaviors, and remote operation were limitations of the system that restricted interactivity and individualization.

Another limitation of the current evidence base regarding ASD robotic technology is the fact that most robotic systems studied have primarily been open-loop and remotely operated and unable to perform autonomous closed-loop interactions, where learning and adaptation are incorporated into the system. Open-loop systems utilize robots with pre-programmed

behavior, and at the time of interaction, they are either remotely operated by humans or execute the pre-programmed behaviors in simple form. Closed-loop systems, which are also referred as autonomous systems, utilize robots that alter their behavior in reaction to environmental interactions and sensor input based on control logic. Those autonomous systems that can not only autonomously react but also adapt their behaviors over time based on the interaction are referred to as adaptive robotic systems. Closed-loop systems have been hypothesized to offer technological mechanisms for supporting more flexible and potentially more naturalistic interaction but have rarely been applied to specific ASD applications. Moreover, the open-loop systems studied to date often require significant resources for operation by necessitating simultaneous involvement of both sophisticated robotic systems *and* specialized system administrators in addition to trained therapists. Hence, these systems may be very limited in terms of application to intervention settings for extended meaningful interactions.

A final limitation of the ASD robotic application literature is the fact that studies have yet to apply appropriately controlled methodologies with well-indexed groups of young children with ASD, and approaches have commonly assessed broad reactions and behaviors during interactions with robots (Kozima et al., 2005), rather than focusing on skills that relate to the core deficits of ASD (Diehl et al., 2012; Robins et al., 2004a, 2004b). Recent works have piloted specific closed-loop systems with potential applicability to ASD populations (Feil-Seifer and Mataric, 2011; Liu et al., 2008); however, these works have not yet examined impact of applications to relevant core deficit areas of the disorder.

In this study, we developed and tested a novel closed-loop adaptive robot-mediated architecture capable of administering joint attention prompts via both humanoid-robot and human administrators. The system automatically provided higher levels of prompts or contingent reinforcement via real-time, noninvasive gaze detection as a marker of response. In simpler terms, the system altered its function based on the child's response to the administrator's prompt for joint attention by providing an additional prompt, changing the type of prompt, or by providing reinforcement. We operationalized response to joint attention as the child's ability to follow an attentional directive to look toward an identified target area. We specifically examined response to joint attention prompts due to findings that deficits in both social orientating and joint attention are thought to represent core social communication impairments of ASD (Mundy and Neal, 2001; Poon et al., 2012), and these skills are often targeted in empirically supported intervention paradigms (Kasari et al., 2008, 2010; Yoder and McDuffie, 2006).

The primary objective of this study was to empirically test the feasibility and usability of a closed-loop adaptive robotic system with regard to providing joint attention prompts and within-system adaptation of such prompts. The secondary objective was to conduct a preliminary comparison of child performance between robot and human administrators. We hypothesized as follows: (a) our robotic system could administer joint attention tasks in a manner that would promote accurate orientation to target and (b) children with ASD would demonstrate increased attention to the humanoid robot compared to the human administrator. We also explored whether young children with ASD would be more accurate with robot prompts than human prompts.

## Methods

### Participants

A total of 12 children (6 ASD, 6 typically developing (TD)) participated in this pilot feasibility study. The age range for ASD group was 2.78–4.9 years, while that for the TD group was 2.18–4.96 years. Initially, a total of 18 participants (10 with ASD and 8 with TD) were recruited. Of the 18 participants, 4 children with ASD and 2 TD children were unable to complete the study. Among the ASD children who failed to complete the study, 3 were not willing to wear the hat and did not start the study protocol, with the other remaining child exhibiting distress during the initial robot-presented trials of the protocol. The 2 TD participants were also unable to start the study due to initial distress. Participants were recruited from an existing university-based clinical research registry as well as existing communication mechanisms attached to the university (e.g. telephone, online, and electronic recruitment). All children in the ASD group received a clinical diagnosis of autism based on *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.; *DSM-IV-TR*; American Psychiatric Association (APA), 2000) criteria from a licensed psychologist, met the spectrum cutoff on the Autism Diagnostic Observation Schedule (ADOS) (Gotham et al., 2007; Lord et al., 2000) administered by a research reliable clinician, and had existing data regarding cognitive abilities in the registry (Mullen Scales of Early Learning; Mullen, 1995). Although not selected a priori based on specific joint attention skills, children demonstrated varying levels of baseline abilities on the ADOS regarding formal assessments of joint attention (i.e. varied abilities on Responding to Joint Attention item of the diagnostic instrument—see Table 1). All parents were asked to complete both the Social Communication Questionnaire (SCQ) (Rutter et al., 2003) and the Social Responsiveness Scale (SRS) (Constantino and Gruber, 2009) to index current ASD symptoms. All children in the TD group fell well below formal risk cutoffs for ASD on both instruments. In all, 4 children with ASD recruited for participation would not tolerate wearing the cap involved in the protocol (described below) and were unable to fully participate in the experimental trials. Available descriptive data for all groups appear in Table 1.

### Apparatus

The system was designed and implemented as a component-based distributed architecture where systems interact via a network in real time. System components included the following: (a) a humanoid robot that provided joint attention prompts, (b) a head-tracker system for gaze inference composed of a network of spatially distributed infrared (IR) cameras, (c) a child-sized cap embedded with light emitting diode (LED) lights, (d) a camera processing module (CPM) capable of providing real-time gaze inference data, and (e) two target monitors that could be contingently activated when children looked toward them in a time-synched response to a joint attention prompt. Figure 1 illustrates experiment room setup.

**Humanoid robot**—The robot utilized, NAO (Figure 2), is a commercially available (Aldebaran Robotics Company) child-sized plastic-bodied humanoid robot (58 cm tall, 4.3 kg) equipped with 25 degrees of freedom (DOF), tactile and audio sensors, and wide variety of actuators (e.g. direct current (DC) servo motors and LED displays). In this work, a new

rule-based supervisory controller was designed within NAO with the capacities for providing joint attention prompts in the form of recorded verbal scripts, head and gross orientation of gaze shifts, as well as coordinated arm and finger points. Prompts were activated based on real-time data provided back to the robot regarding gaze inference as part of the closed-loop interaction paradigm.

**Head tracker for approximate gaze direction inference**—Eye trackers are often utilized to infer gaze, but such systems place potent restrictions on head movement and detection range. They are also expensive, generally sensitive to large head movements, require significant calibration, and often require participants to be in close proximity to targets. Given our focus on young children, we wanted to develop a low-cost system that would not be limited by these factors and could be utilized to track gaze direction approximated by head shift requiring large head movement. The marker-based head tracker used in this protocol is composed of four near-IR cameras placed across the intervention room and three arrays of IR LEDs sewn on the top and the sides of a child-sized cap. Each top and side camera has its own CPM for processing input images containing the LEDs' projections. The IR cameras were custom modified to monitor only in near-IR from inexpensive Logitech<sup>®</sup> Pro 9000 webcams via simple filters that blocked visible light. CPMs were equipped with contour-based image processors to detect the LEDs, provide image plane projections, and to calculate approximate gaze directions based on yaw and pitch angles.

The results of the head-tracker CPMs are combined by a software supervisory module to evaluate the approximate gaze direction represented by an imaginary vector extending normally from the middle of the forehead at the midpoint between the two eyes. This vector was used to perform intersection test between the vector and one of the targets to automatically determine whether the child was looking toward the target. The system was validated for accuracy and correctness in previous work using an IR laser pointer as ground truth. Validation of the head tracker was performed using 20 rectangular square grid patterns of 2 cm × 2 cm each distributed across the target monitors. A near-IR laser pointer was carefully sewn to the head tracker to project the approximate frontal head shift direction, and that was used as a ground truth to measure the accuracy of the head tracker. The head tracker approximated eye gaze with average validation errors of 2.6 cm and 1.5 cm at a distance of 4 feet (1.2° and 0.7°, respectively) in *x* and *y* coordinates (see Bekele et al., 2011). Simply, such a system approximates gaze direction via head movement. While this would potentially be susceptible to errors in averted gaze (e.g. peering without head movement), for the purposes of this study, we positioned the monitors within the test environment such that they required fairly large head movements and shifts in orientation to find the visual targets to avoid effects of small peering on the overall gaze shift accuracy.

**Target monitors**—Two 24-inch computer monitors were hung at identical positions on the left and right sides of the experimental room. The flat screen monitors not only displayed static pictures of interest at baseline but also played brief audio files and video clips based on study protocol. The target monitors were 58 cm × 36 cm (width × height). They were placed at locations 148 cm and 55 cm in the *x* and *y* axes with respect to the top central

camera frame of reference, which is located approximately at the top of the participant's sitting location (Figure 1). They were hung at a height of 150 cm from ground and 174 cm from the top camera (ceiling).

**Network architecture**—A sensory network protocol was implemented in the form of client–server architecture (Bekele et al., 2011). Each CPM of each camera had an embedded client. A central server monitored each camera for time-stamped head-tracking data. The central server, via a command from the software supervisory controller, enabled the CPMs to start monitoring the head movement of the child. The central server processed the raw data for the duration of a trial, produced measured performance metric data from tracking data at the end of each trial, and sent those data to the supervisory controller. The supervisory controller generated feedback and communicated with the humanoid robot and the target monitors.

### Design and procedures

Participants came to the lab for a single visit of 30–50 min. Informed consent was obtained from all participating parents. Participants were introduced to the experiment room and given time to explore the robot. The child was then seated in a Rifton chair at a table across the designated administrator space and then fitted with the instrumented baseball cap. The parent was seated behind the child. Parents were instructed to avoid providing assistance to the child during the study.

Four total blocks of joint attention tasks were presented by the human administrator (two blocks) and humanoid robot (two blocks) with quasi-randomized order of presentation across participants (i.e. Robot, Human, Robot, Human; Human, Robot, Human, Robot). Each trial block consisted of four trials presented in a randomly assigned order. Each trial began with the human or the robot administrator instructing the child to look at one of the two mounted computer monitors. Six predetermined prompt levels (PLs) provided rule-based decision making that cued the next level of prompt to be given by the administrator. Levels of the “least-to-most” hierarchical protocol are described in Table 2. The table shows the specific instructions given to each child within each PL and the accompanying gestures, audio, and video when necessary. Each trial's prompt was 8 s long, including a monitoring interval (approximately 5 s for prompting and a 3-s monitoring interval). If the participant did not respond based on robot/human prompt, an audio (approximately 5 s) and then a video (approximately 5 s), which did not directly address the participant, were used as additional attention capturing mechanisms. If the participant responded at any level of the prompt hierarchy, reinforcement was given via verbal feedback from the robot followed by a 10-s video.

Across all PLs, inferred gaze to target within an 8-s window of prompt administration was labeled as an accurate response. The hierarchy moved children from simple name and gaze prompts, to prompts also combining points, to prompts combining all plus audio and/or visual activation. The administrator proceeded through the hierarchy in this way until the child displayed success with the task. The last step provided the greatest degree of prompting with the added activation of the target audio clip and then video clip onset. For

both human and humanoid-robot administrators, decisions regarding PL administration were governed by the system such that conditions remained constant across administrators. Feedback was provided to the humanoid robot through the networked system. Feedback was also provided to the human administrator by the system via the addition of time-synchronous red and green lights visible only to the administrator (i.e. behind the child) to indicate timing and type of next prompt (i.e. when time of prompt interval had expired for failures). In each trial, a 30-s video clip was turned on contingent to the registration of child success by the system, or at the conclusion of the prompts. These video clips were short musical video segments of common preschool television programs (e.g. Bob the Builder, Dora the Explorer, Sesame Street, etc.) that were randomized across trial blocks and participants. In addition to the video onset, the administrator in each condition provided specifically designated verbal praise to the child for success.

### Primary outcome measures

The system continuously gathered inferred eye gaze data, including data regarding gaze within defined regions of the experimental room. We specifically examined the following performance metrics within task administration: duration of gaze to administrator (i.e. amount of time looking at the robot and human administrator during trials), necessary PLs, and ultimate target success, as well as hit frequency of looks to and away from target during success. Given quasi-randomized blocked presentation of trials across the human and robot conditions, we reduced data to performance across all robot trials and all human trials for individual participants. Given presumed non-normality of the distributed responses, we utilized a nonparametric analytic approach to examine within- and across-group differences. Specifically, within-group and specific ASD versus TD group differences were examined via Wilcoxon and Mann–Whitney tests, respectively.

## Results

### Gaze to administrator

We computed the total time that children's gaze was directed toward the region of interest for both the robot and the human administrator across all trials (Figure 3). These regions corresponded to the specific physical space occupied by the administrator within sessions (i.e. the region varied for robot and human based on size). Considering that the total time of trials varied due to performance (i.e. shorter duration of trials for children requiring limited prompting), we compared percentages of time across the total trial blocks. Children in the ASD group spent an average of 52.8% (standard deviation (SD) = 21.4%) of the robot trial blocks looking at the robot compared to an average of only 25.1% (SD = 18.7%) of the human administrator trial blocks looking at the human administrator. This resulted in mean difference of 27.7%,  $p < 0.05$ . Figure 3 shows a comparison of the time children in the ASD group spent looking at the robot and human administrators across trials.

This same pattern of increased duration of gaze toward robot was observed for the TD control group. The duration of gaze toward robot in the control group yielded a mean of 54.3% (SD = 17.7%) across all trials, whereas the duration of gaze toward the human administrator across all trials yielded a mean of 33.6% (SD = 16%), resulting in mean

difference of 20.7%,  $p < 0.05$ . Children in the TD control group spent slightly more time, 8.53%,  $p > 0.1$ , looking toward the human administrator during trials than the ASD group, but this difference was not statistically significant.

These differences were observed in all individual cases in both the ASD and TD groups. All participants spent a higher percentage of time looking toward the robot compared to the human administrator. Figure 4 shows that this is the case for all participants in both groups.

## PL

We also examined what level of prompting was required for children to successfully find the target across robot and human trials. Given variability in performance of individual participants across trials, we computed a cumulative percentage of prompts necessary to find the target (i.e. PL necessary to find target/total number of PLs). Total number of prompts is the maximum PL (6)  $\times$  total number of available number of trials for each participant (8 in each of human and robot sessions). Both children in the ASD group (Human: 26.4%, SD = 11.7%; Robot: 40.9%, SD = 20%;  $p < 0.05$ ) and TD groups (Human: 20.1%, SD = 4.8%; Robot: 29.5%, SD = 15%;  $p < 0.05$ ) required more PLs for accurate response in the robot conditions when compared to human conditions. The ASD group ultimately required 11.46%,  $p > 0.1$ , more PLs in the robot trials than the TD group, although the difference was not statistically significant. Furthermore, we computed the contribution of each PL to the overall accuracy within each condition to see which PLs were most effective in accurately orienting participants toward the target. Figure 5 shows the contribution of each PL in percentage to the overall accuracy for each condition. The plot indicates that the vast majority of success was achieved by the first PL with subsequent levels contributing 20% or less of the overall accuracy in each condition.

## Target success

We also examined children's abilities to successfully find the target based on human and robot administrator prompts (i.e. percentage of trials that resulted in correct response to target at any of six levels of prompts out of the total number of trials administered). Across human trials, success was at 100% for children in both TD and ASD groups, with all children responding in one of the six PLs. In the robot administrator condition, success measured at 97.9% for children in the TD group and 95.8% for children in the ASD group. The observed difference in performance was not statistically significant by group,  $p > 0.1$ . In order to determine success with robot and human administrator's prompts, rather than simple orientation with activation of dynamic target stimuli, we examined success rates (i.e. percentage of trials successfully completed prior to activation). In a majority of both robotic (ASD = 77.08%; TD = 93.75%) and human trials (ASD = 93.75%; TD = 100%), children were finding targets prior to target activation. However, the ASD group had significantly less accuracy prior to activation than the TD group (mean difference = 16.67%,  $p < 0.05$ ).

## Hit frequency

We also computed hit frequency, defined as the frequency of looks toward and away from the appropriate target during successful trials (i.e. hitting the target and looking away count as one hit). Hit frequency is of particular importance as large hit frequency may be



indicative of erratic gaze, and an atypically small hit frequency may be indicative of difficulties shifting attention within environments (e.g. “sticky attention”) (Landry and Bryson, 2004). Hit frequencies were not significantly different for the ASD group across conditions or between the ASD and TD groups. The ASD group showed an average hit frequency of 2.06 (SD = 0.71) in the human therapist sub-sessions and 2.02 (SD = 0.28) in the robot sub-sessions. Children in the TD group did demonstrate increased hit frequency of 2.17 (SD = 0.65) in the human therapist sub-sessions and 1.67 (SD = 0.35) in the robot sub-sessions ( $p < 0.05$ ).

## Discussion

In this pilot feasibility study, we studied the development and application of an innovative closed-loop adaptive robotic system with potential relevance to core areas of deficit in young children with ASD. The ultimate objective of this study was to empirically test the feasibility and usability of a robotic system capable of intelligently administering joint attention prompts and adaptively responding based on within-system measurements of performance. We also conducted a preliminary comparison of child performance across robot and human administrators. Both TD and ASD children spent more time looking at the humanoid robot and were able to achieve a high level of accuracy across trials. However, across groups, children required higher levels of prompting to successfully orient within robot-administered trials. No specific data suggested that ASD children exhibited preferences or performance advantages within system when compared to their TD counterparts within system.

Children with ASD and TD children were able to ultimately respond accurately to prompts delivered by a humanoid robot and a human administrator within the standardized protocol. Children with ASD also spent significantly more time looking at the humanoid robot than the human administrator, a finding replicating previous work suggesting attentional preferences for robotic interactions over brief intervals of time (Dautenhahn et al., 2002; Duquette et al., 2008; Kozima et al., 2005; Michaud and Théberge-Turmel, 2002; Robins et al., 2009). Furthermore, children with ASD displayed comparable levels of gaze shifts during correct looks, suggesting that differences in attention were likely not simply a reflection of atypically focused gaze toward the technological stimuli or random looking. In terms of tolerability, we anticipated a certain, if not large, fail rate across the ASD sample in terms of willingness to wear the LED cap even for a brief interval of time (i.e. less than 15 min). The completion rate of 60% for the ASD group was promising but ultimately highlights the need for the development of noninvasive systems and methodologies for realistic extension and the use of such technologies with a young ASD population with common sensory sensitivities (Rogers and Ozonoff, 2005). Similarly, such a noninvasive system may help overcome the challenges of head-tracking methodologies for marking and approximating gaze. In the current protocol, targets were placed out of peripheral range of vision to necessitate head movement for tracking; however, precise gaze detection would afford for more robust systems and methodologies in future investigations.

Collectively, these findings are promising in both supporting system capabilities and potential relevance of application. Specifically, preschool children with ASD directed their

gaze more frequently toward the humanoid-robot administrator, they were very frequently ultimately capable of accurately responding to robot-administered joint attention prompts, and they were also looking away from target stimuli at rates comparable to TD peers. This suggests that robotic systems endowed with enhancements for successfully pushing toward correct orientation to target either, with systematically faded prompting or potentially embedding coordinated action with human-partners, might be capable of taking advantage of baseline enhancements in nonsocial attention preference (Annaz et al., 2012; Klin et al., 2009) in order to meaningfully enhance skills related to coordinated attention. The current system only provides a preliminary structure for examining ideal instruction and prompting patterns. Future work examining PLs, the number of prompts, cumulative prompting, or a refined and condensed prompt structure would likely enhance future applications of any such robotic system.

While our data provide preliminary evidence that robotic stimuli and systems may have some utility in preferentially capturing and shifting attention, at the same time, both children with ASD and TD children required higher levels of prompting with the robot administrator when compared to a human administrator in this study. It is entirely plausible that such differences were related to unclear or suboptimal instructions within system or initial naïve response patterns of children, given that children had no previous exposure to an unfamiliar robot and copious exposure to human directives, prompts, and bids. In this context, children had to figure out what the robot was doing and, in turn, expecting them to do. The finding that this was both the experience of the TD and ASD children lends potential support to this explanation. If this was the case, improvements in performance over time might be seen with a refined robot system, including optimized prompts and instructions, and could yield greater success over time based on preferential attention. However, it is also entirely plausible that such a difference highlights the fact that humanoid-robotic technologies, in many of their current forms, are not as capable of performing sophisticated actions, eliciting responses from individuals, and adapting their behavior within social environments as their human counterparts (Dautenhahn, 2003; Diehl et al., 2012). Although NAO is a state-of-the-art commercial humanoid robot, its interaction capacities have numerous limits. Its limb motions (driven by servo motors) are not as fluid as human limb motions, it creates noise while moving its hand, which is not present in the human limb motion, flexibility and DOF limitations produce less precise gestural motions, and its embedded vocalizations have inflection and production limits related to its basic text-to-speech capabilities. In fact, our data ultimately suggest that all children fundamentally performed best with human prompting across all trials when compared to this type of humanoid-robotic interaction. As such, these data suggest that it is unlikely that the mere introduction of a humanoid robot that performs a simple comparable action of a human in isolation will drive behavioral change of meaning and relevance to ASD populations. Robotic systems will likely necessitate much more sophisticated paradigms and approaches that specifically target, enhance, and accelerate skills for meaningful impact on this population. Closed-loop technologies (Feil-Seifer and Mataric, 2011; Liu et al., 2008) that harness powerful differences in attention to technological stimuli, such as humanoid robots or other technologies, may hold great promise in this regard.

There are also several methodological limitations of this study that are important to highlight. The small sample size examined and the limited time frame of interaction are the most powerful limits of this study. As such, while we are left with data suggesting the potential of closed-loop application, the utilized methodology potentially restricts our ability to realistically comment on the value and ultimate clinical utility of this system as applied to young children with ASD. Eventual success and clinical utility of robot-mediated systems hinge upon their ability to accelerate and promote meaningful change in core skills that are tied to dynamic, neurodevelopmentally appropriate learning across environments. We did not systematically intend to assess learning within this system; rather, we indexed simple initial behavioral responses within system application. As such, questions regarding whether such a system could constitute an intervention paradigm remain open. Furthermore, the brief exposure of the current paradigm, in combination with unclear baseline skills of participating children, ultimately cannot answer questions as to whether the heightened attention paid to the robotic system during the study was simply the artifact of novelty or of a more characteristic pattern of preference that could be harnessed over time. In addition, the robot's nonbiological limb movements and inability to move the eyes separately from the head were some of the limitations that might inhibit exact comparison to a human therapist. Another important technical limitation was the approximation of gaze with three-dimensional (3D) head orientation. It must be emphasized that head orientation approximating gaze does not necessarily equate to actual eye gaze, particularly in circumstances of averted gaze. The requirement to wear a hat was also a major limitation, with 33% dropout rate overall. Although this dropout rate is similar or less than minimally invasive clinical devices such as physiological monitoring devices, it highlights the need to develop a noncontact remote eye gaze tracker. Finally, although we made attempts to ensure that children with ASD had received evaluations with gold-standard assessment tools (e.g. ADOS, clinician diagnosis), we did not have rigorous assessment data on the comparison sample on these same instruments nor baseline measurements of basic joint attention skills within system. As such, our ability to comment on the specific clinical characteristics matched with performance differences regarding this technology is limited.

Despite limitations, to our knowledge, this work is the first to design and empirically evaluate the usability, feasibility, and preliminary efficacy of a closed-loop interactive robotic technology capable of modifying response based on within-system measurements of performance on joint attention tasks. Few other existing robotic systems (Feil-Seifer and Mataric, 2011; Liu et al., 2008) for other tasks have specifically addressed how to detect and flexibly respond to individually derived, socially and disorder-relevant behavioral cues within an intelligent adaptive robotic paradigm for young children with ASD. Movement in this direction introduces the possibility of realized technological intervention tools that are not simple response systems, but systems that are capable of necessary and more sophisticated adaptations. Systems capable of such adaptation may ultimately be utilized to promote meaningful change related to the complex and important social communication impairments of the disorder itself.

Ultimately, questions of generalization of skills remain perhaps the most important ones to answer for the expanding field of robotic applications for ASD. While we are hopeful that future sophisticated clinical applications of adaptive robotic technologies may demonstrate

meaningful improvements for young children with ASD, it is important to note that it is both unrealistic and unlikely that such technology will constitute a sufficient intervention paradigm addressing all areas of impairment for all individuals with the disorder. However, if we are able to discern measurable and modifiable aspects of adaptive robotic intervention with meaningful effects on skills seen as tremendously important to neurodevelopment, or tremendously important to caregivers, we may realize transformative accelerant robotic technologies with pragmatic real-world application of import.

## Acknowledgement

The authors gratefully acknowledge the contribution of the parents and children to the AUTOSLab at Vanderbilt as well as the clinical research staff of the Vanderbilt Kennedy Center Treatment and Research Institute for Autism Spectrum Disorders. Specifically, the contributions of Alison Vehorn were instrumental to ultimate project success.

### Funding

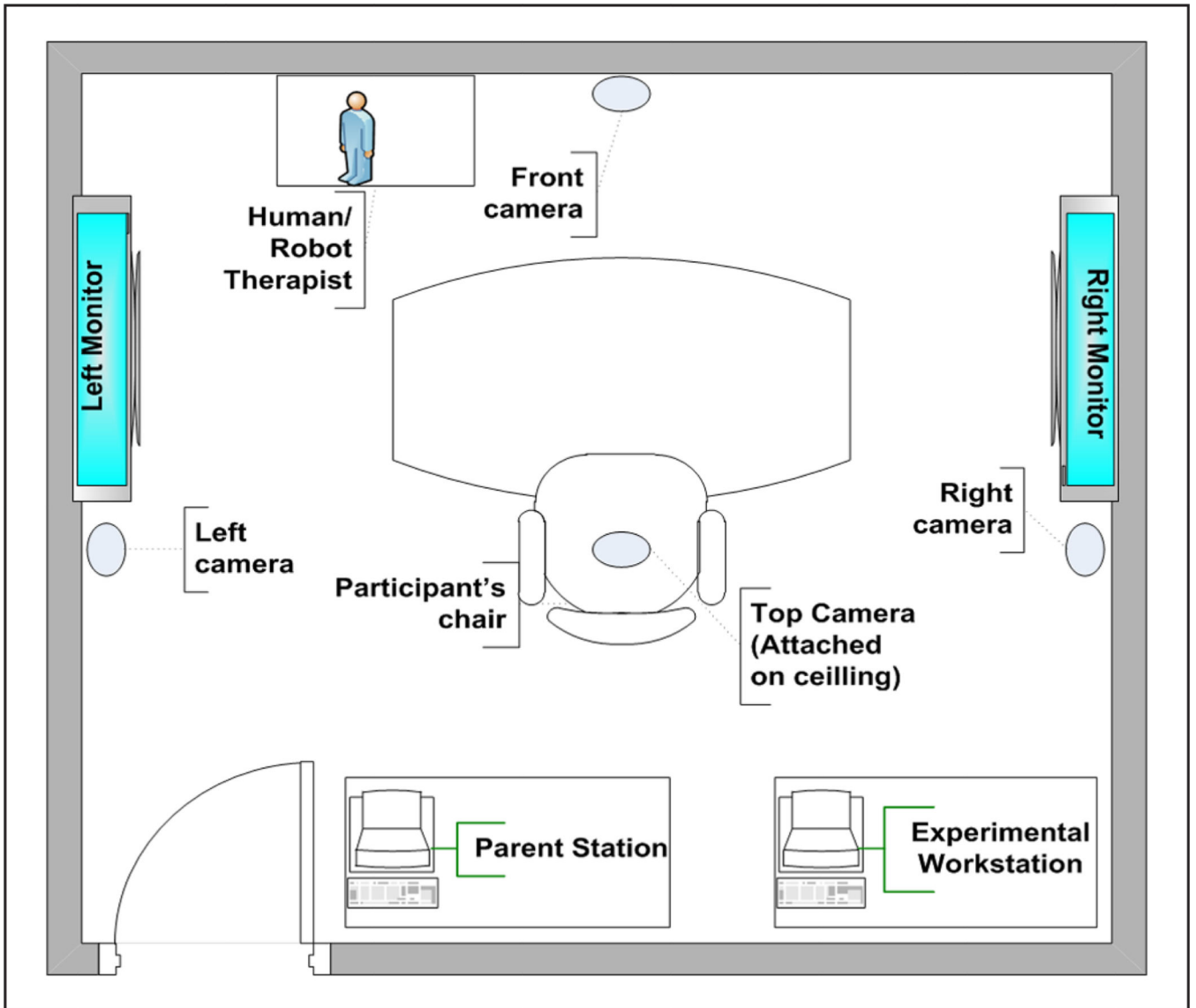
This study was supported in part by grants from Vanderbilt University Innovation and Discovery in Engineering and Science (IDEAS), National Science Foundation Grant [award number 0967170] and National Institute of Health Grant [award number 1R01MH091102-01A1] as well as the Vanderbilt Kennedy Center. This includes core support from the National Institute of Child Health and Human Development (NICHD; P30HD15052) and the National Center for Research Resources (UL1RR024975-01), now at the National Center for Advancing Translational Sciences (2UL1T000445-06).

## References

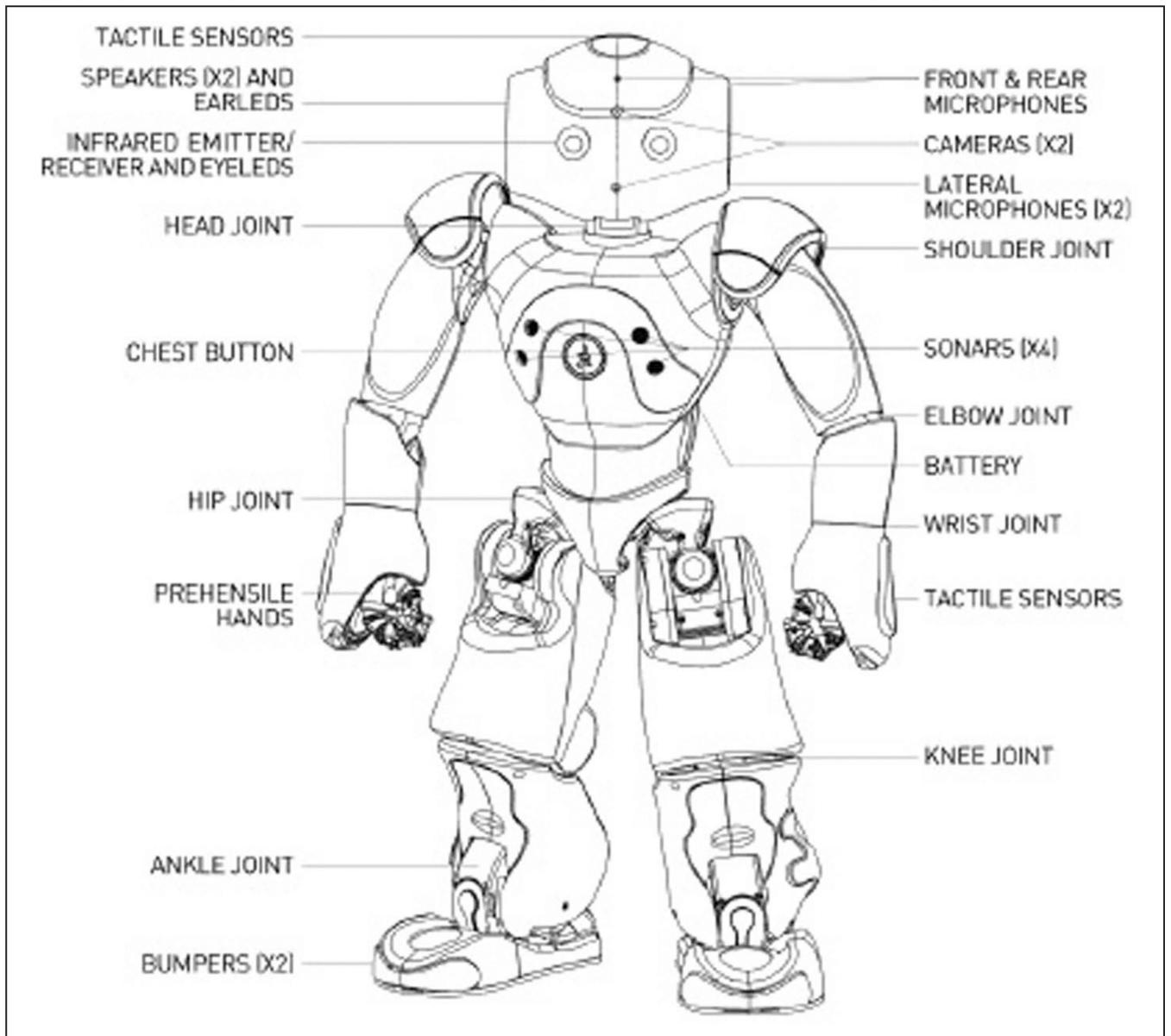
- American Psychiatric Association (APA). Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR). 4th ed. Washington, DC: American Psychiatric Association; 2000. text rev
- Annaz D, Campbell R, Coleman M, et al. Young children with autism spectrum disorder do not preferentially attend to biological motion. *Journal of Autism and Developmental Disorders*. 2012; 42(3):401–408. [PubMed: 21547412]
- Bekele, E.; Lahiri, U.; Davidson, J., et al. Proceedings of the 20th IEEE international symposium on robot man interaction. Piscataway, NJ: RO-MAN 2011 IEEE; 2011. Development of a novel robot-mediated adaptive response system for joint attention task for children with autism; p. 276-281. Atlanta, GA, 31 July–3 August
- Bellani M, Fornasari L, Chittaro L, et al. Virtual reality in autism: state of the art. *Epidemiology and Psychiatric Sciences*. 2011; 20(3):235–238. [PubMed: 21922965]
- Bird G, Leighton J, Press C, et al. Intact automatic imitation of human and robot actions in autism spectrum disorders. *Proceedings of the Royal Society of Biological Sciences*. 2007; 274:3027–3031.
- Centers for Disease Control and Prevention (CDC). Prevalence of autism spectrum disorders—autism and developmental disabilities monitoring network, 14 Sites, United States, 2008. *Morbidity and Mortality Weekly Report*. 2012; 61(3):1–19.
- Constantino, JN.; Gruber, CP. *Social Responsiveness Scale*. Los Angeles, CA: Western Psychological Services; 2009.
- Dautenhahn K. Roles and functions of robots in human society: implications from research in autism therapy. *Robotica*. 2003; 21(4):443–452.
- Dautenhahn K, Werry I. Towards interactive robots in autism therapy: background, motivation and challenges. *Pragmatics and Cognition*. 2004; 12(1):1–35.
- Dautenhahn, K.; Werry, I.; Rae, J., et al. Robotic playmates: analysing interactive competencies of children with autism playing with a mobile robot. In: Dautenhahn, K.; Bond, A.; Cañamero, B., et al., editors. *Socially Intelligent Agents: Creating Relationships with Computers and Robots*. Boston, MA: Kluwer Academic Publishers; 2002. p. 117-124.
- Diehl J, Schmitt L, Villano M, et al. The clinical use of robots for individuals with autism spectrum disorders: a critical review. *Research in Autism Spectrum Disorders*. 2012; 6(1):249–262. [PubMed: 22125579]

- Duquette A, Michaud F, Mercier H. Exploring the use of a mobile robot as an imitation agent with children with low-functioning autism. *Autonomous Robots*. 2008; 24(2):147–157.
- Feil-Seifer, D.; Mataric, M. Proceedings of the 6th international conference on human-robot interaction. New York, NY: ACM Press; 2011. Automated detection and classification of positive vs. negative robot interactions with children with autism using distance-based features; p. 323-330. Lausanne, Switzerland, 6–9 March
- Ganz ML. The lifetime distribution of the incremental societal costs of autism. *Archives of Pediatrics & Adolescent Medicine*. 2007; 161(4):343–349. [PubMed: 17404130]
- Goodwin MS. Enhancing and accelerating the pace of autism research and treatment: the promise of developing technology. *Focus on Autism and Other Developmental Disabilities*. 2008; 23(2):125–128.
- Gotham K, Risi S, Pickles A, et al. The Autism Diagnostic Observation Schedule (ADOS): revised algorithms for improved diagnostic validity. *Journal of Autism and Developmental Disorders*. 2007; 37:400–408.
- Howlin P, Magiati I, Charman T. Systematic review of early intensive behavioral interventions for children with autism. *American Journal on Intellectual and Developmental Disabilities*. 2009; 114:23–41. [PubMed: 19143460]
- Kasari C, Gulsrud AC, Wong C, et al. Randomized controlled caregiver mediated joint engagement intervention for toddlers with autism. *Journal of Autism and Developmental Disorders*. 2010; 40(9):1045–1056. [PubMed: 20145986]
- Kasari C, Paparella T, Freeman S, et al. Language outcome in autism: randomized comparison of joint attention and play interventions. *Journal of Consulting and Clinical Psychology*. 2008; 76(1):125–137. [PubMed: 18229990]
- Klin A, Lin DJ, Gorrindo P, et al. Two-year-olds with autism orient to nonsocial contingencies rather than biological motion. *Nature*. 2009; 459(7244):257–261. [PubMed: 19329996]
- Kozima, H.; Nakagawa, C.; Yasuda, Y. Robot and human interactive communication. Piscataway, NJ: IEEE International ROMAN; 2005. Interactive robots for communication-care: a case-study in autism therapy; p. 341-346. Nashville, TN, 1a3–15 August
- Landry R, Bryson S. Impaired disengagement of attention in young children with autism. *Journal of Child Psychology and Psychiatry*. 2004; 45(6):1115–1122. [PubMed: 15257668]
- Liu C, Conn K, Sarkar N, et al. Online affect detection and robot behavior adaptation for intervention of children with autism. *IEEE Transactions on Robotics*. 2008; 24(4):883–896.
- Lord C, Risi S, Lambrecht L, et al. The autism diagnostic observation schedule-generic: a standard measure of social and communication deficits associated with spectrum of autism. *Journal of Autism and Developmental Disorders*. 2000; 30(3):2055–2223.
- Michaud, F.; Théberge-Turmel, C. Mobile robotic toys and autism. In: Dautenhahn, K.; Bond, A.; Cañamero, L., et al., editors. *Socially Intelligent Agents: Creating Relationships with Computers and Robots*. Boston, MA: Kluwer Academic Publishers; 2002. p. 125-132.
- Mullen, EM. *Mullen Scales of Early Learning*. Circle Pines, MN: American Guidance Services Inc; 1995.
- Mundy P, Neal R. Neural plasticity, joint attention and a transactional social-orienting model of autism. *International Review of Mental Retardation*. 2001; 23:139–168.
- Peacock G, Amendah D, Ouyang L, et al. Autism spectrum disorders and health care expenditures: the effects of co-occurring conditions. *Journal of Developmental and Behavioral Pediatrics*. 2012; 33(1):2–8. [PubMed: 22157409]
- Pierro AC, Mari M, Lusher D, et al. Robotic movement elicits visuomotor priming in children with autism. *Neuropsychologia*. 2008; 46(2):448–454. [PubMed: 17920641]
- Poon KK, Watson LR, Baranek GT, et al. To what extent do joint attention, imitation, and object play behaviors in infancy predict later communication and intellectual functioning in ASD? *Journal of Autism and Developmental Disorders*. 2012; 42(6):1064–1074. [PubMed: 21858585]
- Robins, B.; Dautenhahn, K.; Dickerson, P. *Advances in computer-human interactions second international conference (ACHI)*. New York, NY: 2009. From isolation to communication: a case study evaluation of robot assisted play for children with autism with a minimally expressive humanoid robot; p. 205-211. Cancun, QR, Mexico, 1–7 February

- Robins B, Dautenhahn K, Dubowski J. Does appearance matter in the interaction of children with autism with a humanoid robot? *Interaction Studies*. 2006; 7(3):479–512.
- Robins, B.; Dautenhahn, K.; Te Boekhorst, R., et al. Proceedings of the universal access and assistive technology (CWUAAT). London: Springer-Verlag; 2004a. Effects of repeated exposure to a humanoid robot on children with autism; p. 225-236. Cambridge, UK, 22–24 March
- Robins B, Dickerson P, Stribling P, et al. Robot-mediated joint attention in children with autism: a case study in robot-human interaction. *Interaction Studies*. 2004b; 5(2):161–198.
- Rogers SJ, Ozonoff S. What do we know about sensory dysfunction in autism? A critical review of the empirical evidence. *Journal of Child Psychology and Psychiatry*. 2005; 46(12):1255–1268. [PubMed: 16313426]
- Rogers SJ, Vismara LA. Evidence-based comprehensive treatments for early autism. *Journal of Clinical Child and Adolescent Psychology*. 2008; 37(1):8–38. [PubMed: 18444052]
- Rutter, M.; Bailey, A.; Lord, C. *The Social Communication Questionnaire*. Los Angeles, CA: Western Psychological Services; 2003.
- Warren Z, McPheeters ML, Sathé N, et al. A systematic review of early intensive intervention for autism spectrum disorders. *Pediatrics*. 2011; 127(5):e1303–e1311. [PubMed: 21464190]
- Yoder, PJ.; McDuffie, AS. Treatment of responding to and initiating joint attention. In: Charman, T.; Stone, W., editors. *Social and Communication Development in Autism Spectrum Disorders: Early Identification, Diagnosis, and Intervention*. New York: Guilford Press; 2006. p. 88-114.

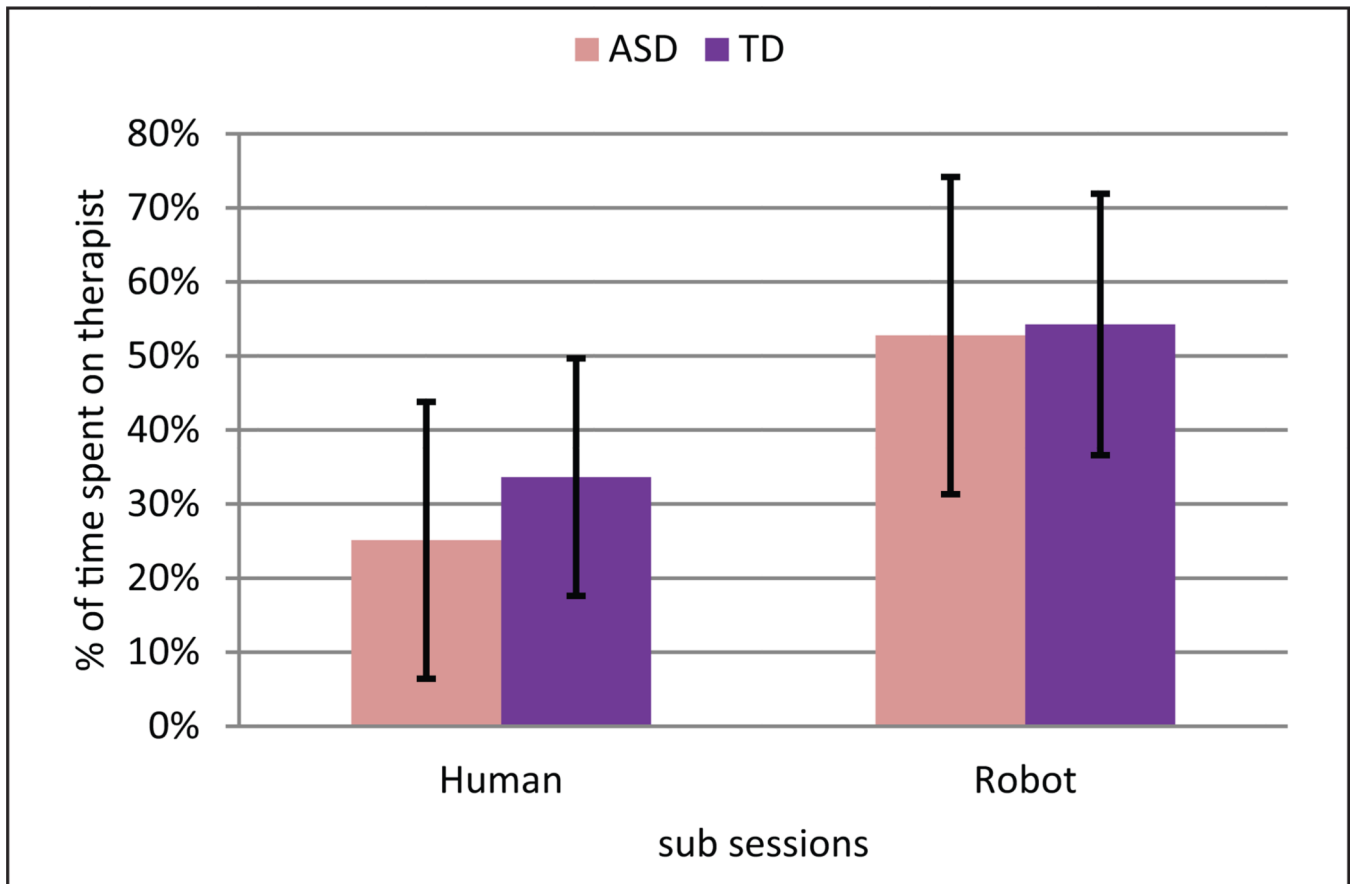


**Figure 1.**  
Apparatus and experiment room setup (not drawn to scale).

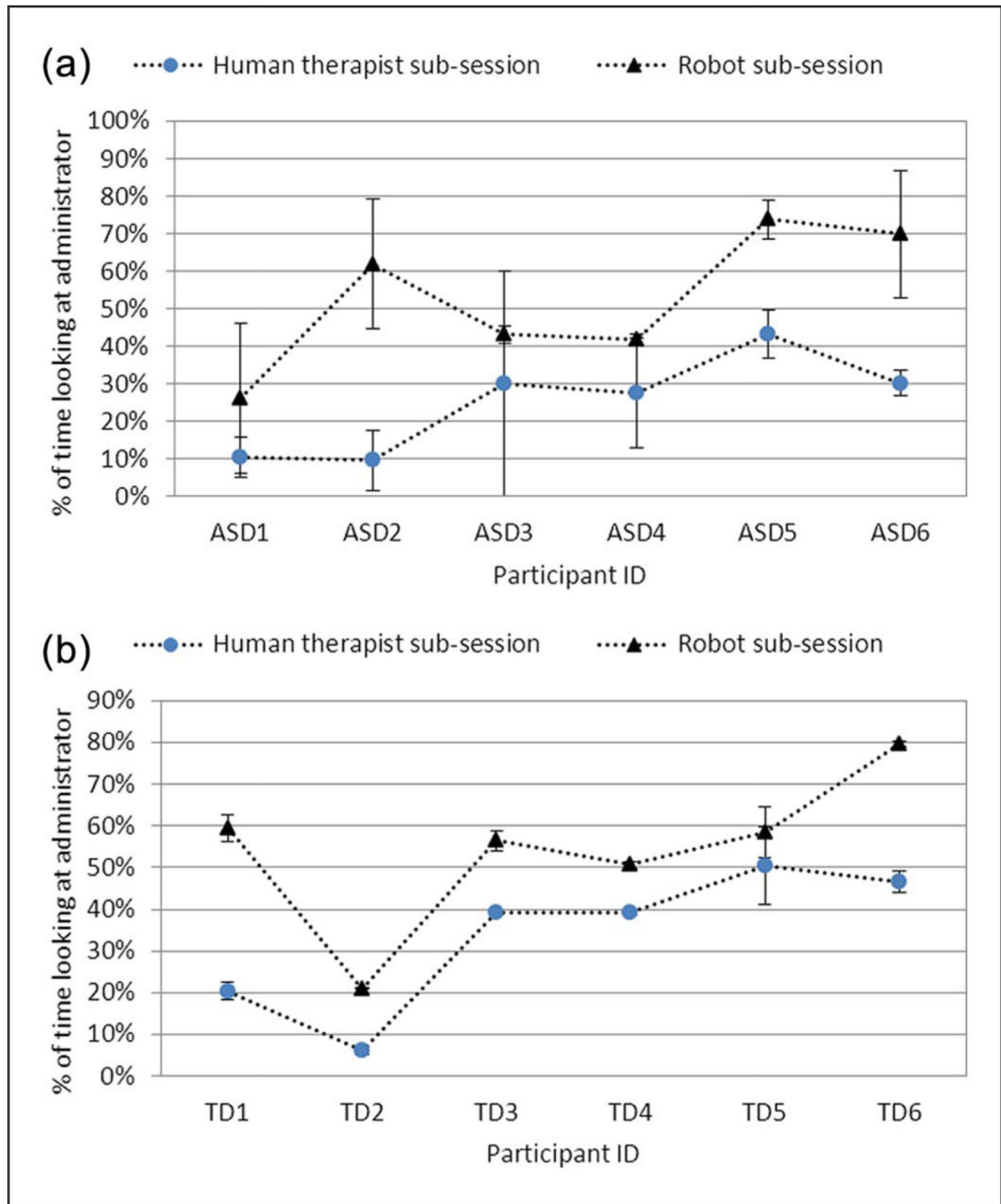


**Figure 2.**  
Humanoid-robot NAO utilized within protocol.

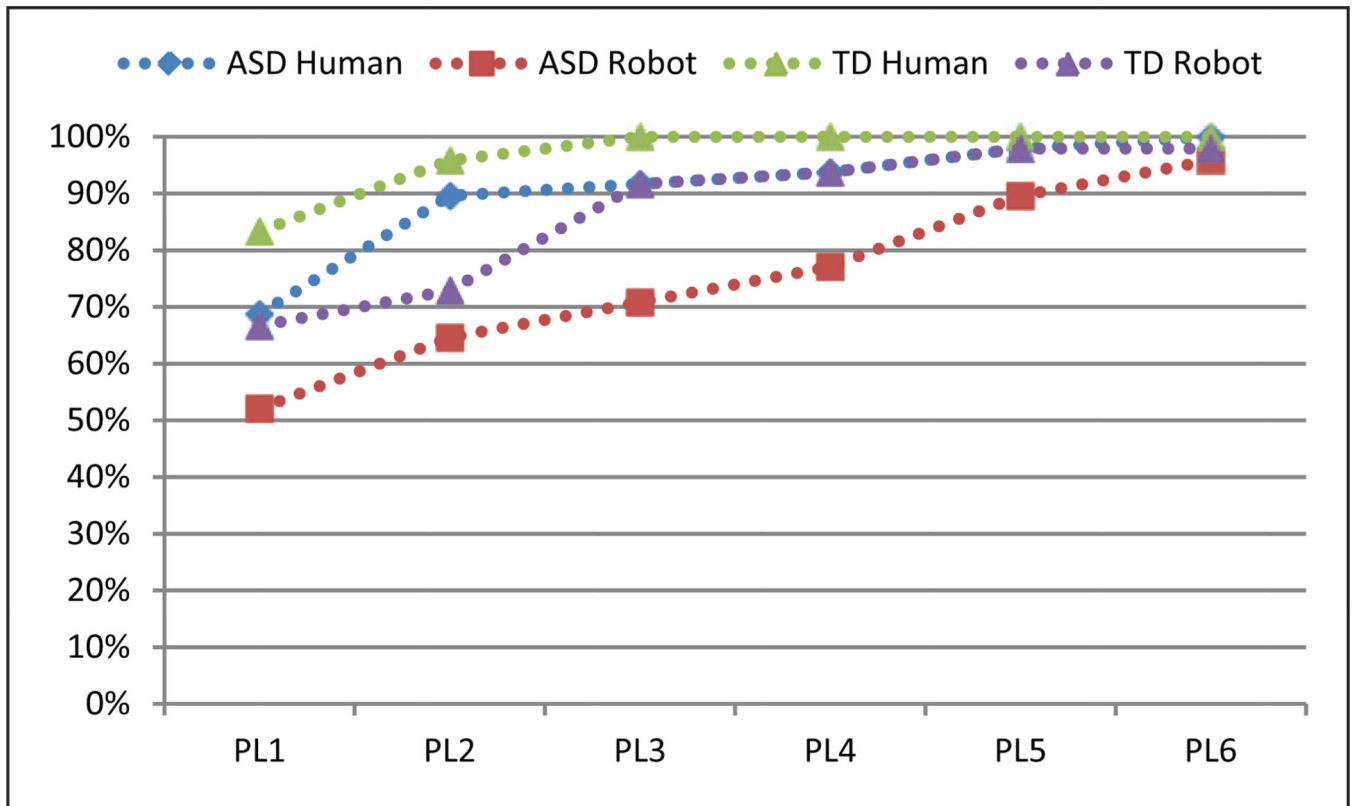




**Figure 3.**  
Percentage of time looking toward robot and human administrators across trials.



**Figure 4.** Percentage of time spent looking at the administrator in both the human and the robot sessions for individual participants in (a) ASD group and (b) TD group. ASD: autism spectrum disorder; TD: typically developing.



**Figure 5.** Percentage of correct responses achieved at specific prompt levels by group and condition.

**Table 1**

Participant characteristics (n = 16).

<b>Group M (SD)</b>			
	<b>ASD (n = 6)</b>	<b>TD (n = 6)</b>	<b>Failed to complete (n = 4)</b>
Gender (% male)	80%	60%	100%
Chronological age (years)	4.7 (0.7)	4.4 (1.15)	3.5
Mullen ELC	71.5 (22.65)	—	74.5 (21.1)
SRS total score	70.3 (12.2)	45.5 (3.3)	—
SCQ total score	13.3 (5.9)	3.8 (3.5)	—
ADOS total score	16.0 (6.6)	—	19.8 (5.4)
ADOS CSS	7.3 (2.0)	—	7.8 (2.3)
ADOS RJA item score	1.2 (1.1)	—	1.3 (1.3)

SD: standard deviation; ASD: autism spectrum disorder; TD: typically developing; ELC: Early Learning Composite; SRS: Social Responsiveness Scale; SCQ: Social Communication Questionnaire; ADOS: Autism Diagnostic Observation Schedule; CSS: Calibrated Severity Score; RJA: response to joint attention.

**Table 2**

Joint attention prompt hierarchy.

Prompt levels (PLs)	Administrator prompt
PL 1	“Name, look” + gaze shift to target
PL 2	“Name, look” + gaze shift to target
PL 3	“Name, look at that.” + gaze shift to target + point to target
PL 4	“Name, look at that.” + gaze shift to target + point to target
PL 5	“Name, look at that.” + gaze shift to target + point to target + audio clip sound at target
PL 6	“Name, look at that.” + gaze shift to target + point to target + audio clip sound at target + video onset for 30 s

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript