

Social Robots as Embedded Reinforcers of Social Behavior in Children with Autism

Elizabeth S. Kim · Lauren D. Berkovits ·
Emily P. Bernier · Dan Leyzberg · Frederick Shic ·
Rhea Paul · Brian Scassellati

Published online: 31 October 2012
© Springer Science+Business Media, LLC 2012

Abstract In this study we examined the social behaviors of 4- to 12-year-old children with autism spectrum disorders (ASD; $N = 24$) during three tradic interactions with an adult confederate and an interaction partner, where the interaction partner varied randomly among (1) another adult human, (2) a touchscreen computer game, and (3) a social dinosaur robot. Children spoke more in general, and directed more speech to the adult confederate, when the interaction partner was a robot, as compared to a human or computer game interaction partner. Children spoke as much to the robot as to the adult interaction partner. This study provides the largest demonstration of social

human-robot interaction in children with autism to date. Our findings suggest that social robots may be developed into useful tools for social skills and communication therapies, specifically by embedding social interaction into intrinsic reinforcers and motivators.

Keywords Social robots · Assistive robots · Intervention · Embedded reinforcers

Introduction

Social deficits are among the primary characteristics associated with autism spectrum disorders (ASD; American Psychiatric Association 2000; Carter et al. 2005; Joseph and Tager-Flusberg 1997; Kanner 1943; Mundy et al. 1989). Specific deficits are widely heterogeneous across individuals with ASD, with some individuals never developing functional speech (Volkmar and Klin 2005). Even for individuals who develop strong formal language skills, difficulties persist in conversational interactions (Mesibov 1992; Paul 2008; Tager-Flusberg et al. 2005). For example, common deficits in conversational skills include difficulty managing turn-taking and topics of discourse, using inappropriate style of speech to fit conversation partners and settings, and trouble inferring what information is relevant or interesting to others (Paul 2008). Production and perception of affective expressions, as well as eye contact and other non-verbal attentional cues, can also be inappropriate, unconventional or deficient in individuals with autism (Mundy et al. 1986).

These deficits impact individuals' abilities to function independently in social, occupational and other important areas of life (American Psychiatric Association 2000). Many individuals need high levels of support and care

E. S. Kim · E. P. Bernier · D. Leyzberg · B. Scassellati
Department of Computer Science, Yale University, New Haven,
CT, USA

Present Address:

E. S. Kim (✉)
Child Study Center, School of Medicine, Yale University,
40 Temple St. Ste. 7D, New Haven, CT, USA
e-mail: elizabeth.kim@yale.edu

L. D. Berkovits · F. Shic · R. Paul
Child Study Center, School of Medicine, Yale University,
40 Temple St. Ste. 7D, New Haven, CT, USA

Present Address:

L. D. Berkovits
Department of Psychology, UCLA, Los Angeles, CA, USA

Present Address:

E. P. Bernier
Department of Psychology, Harvard University, Cambridge,
MA, USA

R. Paul
Speech-Language Pathology, Sacred Heart University, Fairfield,
CT, USA

throughout their lives, and early intervention is considered critical (American Psychiatric Association 2000; Klin et al. 2000; Mullen 1995; Sparrow et al. 2005; Volkmar et al. 2004).

Diverse human-delivered intervention approaches seek to improve social and communication skills in children with ASD (reviewed, for example, in Paul 2008; and Volkmar et al. 2004). Interventions tend to vary in terms of (1) the specific behaviors targeted; (2) whether targeted behaviors are allowed to be spontaneously initiated by the child, are motivated by naturalistically enticing the child (for instance, tempting a child to ask for help by keeping a toy out of reach), or are explicitly elicited through instruction (e.g., through repetition of highly structured interactions); and (3) the setting in which training or reinforcement takes place—anywhere from highly controlled clinical settings to naturalistic environments such as the child's home or classroom. In all cases, interventions reward children for performing targeted behaviors, whether by delivering edible reinforcers, providing access to preferred toys or objects, or by allowing the child to engage in a preferred activity (for instance, watching a favorite television program).

There is evidence that the use of child-preferred, or *intrinsic*, reinforcers leads to improvements in social engagement (reviewed in Paul 2008). Furthermore, *embedding* social interaction into the delivery of a child's preferred reinforcer (for example, singing a child's favorite song, rather than playing a video recording of the song) elicits greater social initiation, increased non-verbal (bodily) orientation to face an interaction partner, and more positive affect (L. K. Koegel et al. 1999; R. L. Koegel et al. 1987a, 2009).

The long-term aim of our research is to evaluate and fulfill the potential of social robots as *embedded reinforcers*, which elicit and reward social behavior in interventions for children with autism. Although there is ample evidence that children with ASD (as well as children and adults with typical development) will engage socially with robots, our long-term aim focuses on the ways social robots may support therapies that improve social interaction with other people (Duquette et al. 2008; Feil-Seifer and Matarić 2009; Kozima et al. 2009; Robins et al. 2005; Stanton et al. 2008). *Social robots* are designed to evoke social behaviors and perceptions in the people with whom they interact. There is promising case study evidence (discussed below) that robots, both socially evocative and not, can elicit social engagement from children, toward the robots themselves, and can mediate social engagement between children and adults. Whereas Koegel et al. (2009) have shown that embedding social interaction, *within the delivery* of preferred reinforcers, increases production of target behaviors, we are interested in further embedding social interaction, *within the reinforcer or motivator itself*. It is

our eventual hope that social robots can translate children's interest in novel technologies into increased motivation for participating in social interactions and social partnerships with people. Such an approach, if effective, could provide new methods to facilitate and augment behavioral, communicative, and social therapies that improve interactions between individuals with ASD and with other people (Scassellati 1996).

Non-human intervention tools have been explored for use with children with ASD. These include pet- (Martin and Farnum 2002; Redefers and Goodman 1989) computer- (Bosseler and Massaro 2003; Hetzroni and Tannous 2004), and virtual-reality-assisted therapies (Parsons and Mitchell 2002; Strickland 1997). Social robots have been investigated as assistive tools for elderly, or physically or cognitively impaired individuals (Scassellati et al. 2012; Tapus et al. 2007), and as supportive tools for social and communication skills therapy in children with ASD (Duquette et al. 2008; Feil-Seifer and Matarić 2009; Kozima et al. 2009, 2005; Robins et al. 2005; Scassellati 2005; Stanton et al. 2008; Werry and Dautenhahn 1999). Multiple studies have shown that children with ASD will interact with robots using social behaviors, e.g., by directing speech to the robot (Duquette et al. 2008; Feil-Seifer and Matarić 2009; Kozima et al. 2009; Robins et al. 2005; Stanton et al. 2008). Several of these studies have further demonstrated that children with ASD will interact with a parent, caregiver, or another human while engaged with a robot partner (Feil-Seifer and Matarić 2009; Kozima et al. 2009; Robins et al. 2005), for instance, by expressing excitement to a robot, and then turning to express this excitement to a parent (Kozima et al. 2009).

To date, the benefits of robotic interaction on social behaviors have been demonstrated over case studies of three or four individual children. However, there are few demonstrations over larger samples (for a review see Diehl et al. 2012). It has thus remained an open question whether the beneficial effects of social robots extend more broadly across the autism spectrum. The present study is designed to help answer this question using a randomized, controlled, crossover design, over a larger sample of children with ASD ($N = 24$), to examine the extent to which social robots can both elicit social engagement directed toward the robot itself, and motivate or facilitate social interactions with another person.

Methods

We designed a randomized, controlled, crossover experiment to compare the effects of interactions with a social dinosaur robot (Fig. 1) against the effects of interactions with a human or an asocial novel technology (a



Fig. 1 The socially expressive robot Pleo. In the robot condition, participants interacted with Pleo, a small, commercially produced, toy dinosaur robot. Pleo is about 21 inches long, 6 inches wide, and 8 inches high, and was designed to express emotions and attention, using body movement and vocalizations that are easily recognizable by people, somewhat like a pet dog. For this study we customized Pleo's movements, synchronized with pseudo-verbal vocalizations, to express interest, disinterest, happiness, disappointment, agreement, and disagreement

touchscreen computer game). Each participant in our study completed a sequence of three 6-min *interactional conditions*, in random order: one in which the interaction partner was a dinosaur robot, another in which the partner was an adult, and a third in which the partner was a touchscreen computer game. All interactional conditions (which we will also refer to simply as *conditions*) were guided and facilitated by a human confederate (different from the adult interaction partner) and took place in a standard clinical observation room.

Before the first, after the final, and between interactional conditions, each participant also completed 6-min, semi-structured *interview-and-play sessions*, which we will also refer to as *interviews*. Interview-and-play sessions gave participants rest from the more structured interactional conditions. They were conducted in another clinical observation room, different from the room where interactional conditions were administered. The interactional conditions and interspersed interviews are described in greater detail below (see “[Procedures](#)”).

We expected that children with ASD would find (1) the robot interactional condition social and engaging; (2) the human adult interactional condition social but less engaging; and (3) the computer game interactional condition engaging but not social. Thus we hypothesized that children with ASD would verbalize more while interacting with a social robot than while interacting with either a human adult or a computer game. Given evidence, from case studies (Kozima et al. 2009) and from our own pilot studies, that interaction with a social robot motivates high

levels of curiosity and increases social behaviors such as sharing expressions of excitement with an adult, we also hypothesized that children would direct more speech toward an adult confederate when the interaction partner was a robot, rather than when the partner was another adult or a computer game. We investigated these hypotheses in support of our ultimate goal—to understand the utility of social robots as reinforcers of social interaction with people (as opposed to robots).

Participants

Participants were recruited from two ongoing studies at a university-based clinic specializing in assessment, intervention, and educational planning for children with ASD. These included a multi-site comprehensive study of families with multiple children, only one of whom is affected by autism; and a longitudinal study of language development in children with ASD. Inclusion criteria included a chronological age of 4 to 12 years and a previous diagnosis of high-functioning ASD (defined as full-scale IQ ≥ 70 and verbal fluency with utterance production of phrases of at least 3 words).

Of the 30 initial volunteers for the study, two were excluded from participation due to below-threshold IQ measurement. Of the remaining 28 participants, four were excluded from analysis: one participant withdrew before completing the procedure; one was excluded for failing to meet ADOS criteria for ASD; and two were excluded due to technical recording problems that precluded speech annotation.

In the 24 participants that ultimately constituted our analytical sample, ages ranged from 4.6 to 12.8 years ($M = 9.4$, $SD = 2.4$). IQ eligibility was confirmed within 1 day of participation in this study using the Differential Abilities Scale (DAS-II: $M = 94.2$, $SD = 11.7$, $Min = 72$, $Max = 119$; Elliott 2007). Similarly, within 1 day of participation in this study, all participants completed the Autism Diagnostic Observation Schedule—Module 3 (ADOS—Module 3; Lord et al. 2000) with an experienced clinician, and diagnosis was confirmed by a team of clinical experts. Twenty participants met ADOS criteria for autism, and four for autism spectrum disorder. Of the 24 participants for whom analysis is presented in this article, three were female. Twenty participants were white (and not of Hispanic origin), two were black (and not of Hispanic origin), and two were Hispanic or Latino.

Materials

Video Recording

All interactional conditions and interviews were recorded using Mini-DV video cameras on stationary tripods from

distances of six feet and four feet from participants in the interactional conditions and interviews, respectively.

Robot, Robot Behavior, and Robot Control

The Pleo robot was used in the robot interactional condition because previous investigations have shown that healthy adults (Kim et al. 2009) as well as children with autism (pilot studies) readily engage socially with this robot. Pleo (Fig. 1) is an affectively expressive, toy dinosaur robot, recommended for use by children 3 years and older. It was formerly commercially produced and sold by UGOBE LifeForms; a larger, different model is now produced and sold by Inovo Labs (2012). Pleo measures approximately 21 inches long, 6 inches wide, and 8 inches high. It is untethered, battery-powered, and has 15 degrees of mechanical freedom. We extended UGOBE software to render Pleo controllable by a handheld television remote control, which communicates with Pleo via a built-in infrared receiver on the robot's snout, allowing us to instantaneously play any one of 13 custom, pre-recorded, synchronized motor and sound scripts on the robot. Pleo plays sounds through a loudspeaker embedded in its mouth.

We pre-programmed Pleo with 10 socially expressive behaviors, including a greeting, six affective expressions, and three directional (left, right, center) expressions of interest (to be directed towards nearby objects). All socially expressive behaviors were made up of motor movements synchronized with speech-like vocal recordings. We also pre-programmed three non-social behaviors: a bite (for holding blocks), a drop from the mouth (for letting go of blocks), and a forward walking behavior used when the robot interactional condition called for Pleo to interact with an object that was beyond its reach. Each of these 13 triggered behaviors endured for less than 2 s, and were initiated with the push of a button on Pleo's remote control.

When not executing one of the 13 triggered behaviors, Pleo continuously performed a background behavior, designed to maintain the appearance of its animacy. In the background behavior, Pleo randomly shifted its hips, bent and straightened its legs, and slightly nodded its head up and down, or left and right. Robot behaviors, and their carefully matched adult counterparts, are detailed in Table 1, and are further motivated below (see "Procedures").

We used hidden, *Wizard of Oz*-style, real-time, human remote control of the robot, a popular design paradigm in human-robot interaction research (Steinfeld, Jenkins, and Scassellati 2009), in order to elicit each participant's belief that Pleo was behaving and responding autonomously. In truth the adult interaction partner, who remained present for all interactional conditions, secretly operated the robot using a television remote control, hidden underneath a clipboard. The *Wizard of Oz* paradigm affords a robot with

the appearance of autonomous perception and behavior, with an accuracy and flexibility that currently only humans can produce. Under *Wizard of Oz* control, the Pleo robot has been shown to successfully impart an appearance of autonomous social interaction, both to adults with typical development (Kim et al. 2009) and to school-aged children with ASD (pilot testing).

The adult interaction partner was present for all three interactional conditions. In order to obscure the adult interaction partner's manual control of the robot, the confederate explained to participants that the adult partner would remain present for the robot condition, for the purpose of observing the robot's behavior. To maintain consistency with the robot condition, the confederate explained that the adult partner would remain present during the computer game, as well, for the purpose of ensuring that the computer worked. Throughout the robot and computer game conditions, the adult partner stood apart from the participant, confederate, and interaction partner, pretending to read papers on a clipboard and remaining silent unless addressed by the participant (see Fig. 2). In the robot condition, the adult partner hid the robot's television remote control beneath the clipboard.

It is important to note that most children, including those with typical development, largely or entirely ignored the adult interaction partner during the robot and computer game conditions. Only one participant voiced suspicion that the adult controlled the robot, and subsequently discovered the television remote beneath the clipboard at the end of the robot interactional condition. We included this participant in analysis nonetheless, because his discovery was made too late to affect his behavior while interacting with the robot.

Procedures

Adult and Robot Interactional Conditions

The adult, robot, and computer game interactional conditions were semi-structured and were completed by all participants in randomized orders. Interactional conditions took place on a 3-foot square table, with the participant and confederate sitting at adjacent sides. During the adult condition, the adult interaction partner sat to the other side of the participant, opposite the confederate (see Fig. 2). For the robot and computer game conditions, the adult's chair was left empty, and the adult stood several feet away from the table with clipboard in hand.

The adult and robot interactional conditions were designed to elicit social interaction, and were semi-structured closely in parallel to each other. The touchscreen computer game interaction was not designed to elicit social interaction, and thus did not match the interactional

Table 1 The Pleo robot's pre-programmed behaviors, and the adult partner's matched behaviors

Social intent expressed, or non-social activity	Robot		Adult	
	Movements	Pseudo-verbal vocalization	Movements	Minimally verbal vocalization
Greeting and satisfaction	Tail wags, head raises	“Hee!”	Smiles and looks at participant	An enthusiastic “Hi, < participant's name > !”
Selection of or interest in an object (in one of three directions for robot)	Head lowers toward left, right, or center	A prolonged, enthusiastic “Ooh!”	Looks in direction of object, points from afar	“Oooh!” or “That one!”
Yes	Head nods up and down	“Mm hmm!”	Looks at participant, nods	“Mm hmm!” or “Yes!”
Enthusiastic affirmative	Head raises, tail wags briefly, hips wiggle briefly	“Woohoo!”	Lifts head slightly or sits moderately upright and smiles moderately at participant	A slightly moderated “Nice!” or “All right!”
Elation	A dance: heads raises and moves left and right, hips wiggle, knees bend and straighten	An extended victory song	Sits upright energetically, smiles widely at participant, claps hands or puts hands in air	An extended and exaggerated “Woohoo!” or “Awesome!” or “Fantastic!”
No	Head shakes side-to-side	“Unh unh.”	Shakes head back and forth and frowns slightly	“Unh unh,” or “No.”
Dissatisfaction	Head and tail lower, mouth opens	“Ehhh.”	Frowns moderately, looks slightly downward, and hangs head slightly	“Ehhh.”
Intense disappointment	Head and tail lowers, head shakes slowly from side to side	A prolonged audible sigh, followed by a whimper	Slumps in chair or puts chin in hands, hangs head, looks downward	An audible sigh, followed by an extended, exaggerated “Awwww,” or “Oh man.”
Bite	Head raises, mouth opens for several seconds, then closes	“Aaaaahhhh...chomp.”		
Drop from mouth	Head lowers, mouth opens widely	<silence>		
Walk	Pleo takes four very short (0.5-inch) steps forward	“Hup, hup, hup. Hup!”		
Background animacy	Head occasionally moves up and down, and left and right; hips wiggle occasionally; knees bend and straighten occasionally	<silence>		

Ten robot behaviors were socially expressive, including a greeting, six affective expressions, and three directional (left, right, and straight ahead) expressions of attention, and were carefully matched with vague verbalizations in the adult interaction partner. In addition to the ten social behaviors, Pleo had three non-social behaviors (walk, bite, drop), and a “background” behavior to express animacy (i.e., that Pleo takes note of its environment and experiences feelings of boredom or interest). In the adult interaction partner, the robot's walk, bite, and drop matched an extended reach, a grasp, and a drop, respectively. All behaviors were carefully designed to be expressed multi-modally, through vocal prosody, and body and head movement

structure of the adult and robot conditions. Our intention was to compare participants' responses to two novel interesting technologies which provided contrasting amounts of social reinforcement. In all three conditions, children manipulated blocks: multi-colored, magnetically linking tiles in the robot condition; multi-colored, interlocking blocks in the adult condition; and tangrams, which the participant could move and turn by dragging or tapping the touchscreen with his or her finger (or a stylus, if preferred) in the computer game condition.

The adult and robot interactions were designed to elicit a host of social perception, reasoning, and interaction behaviors from participants. These included taking turns with the interaction partner; identifying the interaction partner's emotions or expressions of preference for one particular block or another; and shared, imaginative, and tactile play. The confederate's role was to guide the participant through an ordered, standard set of activities and cognitive probes, by instructing the participant, by subtly directing the adult or robot partner when to deliver pre-



Fig. 2 Three interactional conditions: adult (*top*), robot (*middle*) and touchscreen computer game (*bottom*). The confederate sits to the participant's right

scripted cues or affirmations, and by asking increasingly restrictive questions of the participant. In the robot and adult interactional conditions, one of each of the following probes and activities were completed, in order:

1. (Probe) The confederate instructed the participant to present one block at a time to the robot or adult interaction partner, and then asked the participant to identify whether the partner liked or disliked each block's color.
2. (Activity) The participant assembled the blocks into a structure of his or her own choosing. The participant and partner took turns selecting each next block to add to the structure.
3. (Probe) During each turn, the adult and robot interaction partners did not manipulate each chosen block directly. Instead, to indicate choice, the adult vaguely pointed at a block, saying, "That one!" or the robot

turned its head to look at a block, saying, "Oooh!" to choose it. The participant was asked to identify which block the adult or robot had chosen, and then was instructed to add that block to the structure.

4. (Probe) When the structure was completed, the adult or robot interaction partner expressed elation pseudo-verbally ("Woohoo!") and bodily (clapping hands or wagging tail, respectively), as further described in Table 1. The participant was asked to identify the partner's emotional state. Next, the confederate removed the blocks from the table, and the adult or robot interaction partner expressed disappointment (as described in Table 1). The participant was again asked to identify the partner's emotional state.
5. (Activity) Finally, the confederate invited the participant to pet the robot or invent a secret handshake game with the adult partner. In the robot condition, petting was included to give participants an opportunity explore the robot, while in the adult condition the secret handshake game was included to match the robot condition's tactile, interactive, and inventive petting activity. In the secret handshake game, each participant was instructed to tap or shake the adult partner's hand in any way he or she chose. The adult partner then presented his or her right hand as though to shake hands until the participant made contact, after which he or she exclaimed in delight, and then presented his or her hand open-palmed as if to give a high-five and again expressed delight when the participant made contact a second time. With the robot, participants were offered a chance to guess the robot's favorite spot to be petted. The robot exclaimed in delight after first contact, and participants were then told that the robot had another favorite spot. After being petted a second time, the robot expressed elation (its happy dance).

Items 1, 3, and 4, above, probed participants' perception and understanding of the robot and adult interaction partners' expressions of affect and preference. Each probe was delivered through a series of increasingly restrictive cues or presses. First the interaction partner would express an emotion or preference (e.g., lowering the head and sighing with prosody expressing disappointment), after which the partner and confederate waited silently for 2 s, giving the participant an opportunity to respond or comment spontaneously. If participants responded appropriately (some participants immediately comforted the robot or adult interaction partner), the confederate guided the interactional condition to the next activity or probe. Otherwise (some participants did not respond to the emotional or preferential expression), the confederate delivered a press, asking the child to interpret the behavior (e.g., "Why do

you think Pleo/Dan said that? How do you think he feels?”). If the participant did not appropriately respond to the confederate’s first press, the confederate delivered a second, more restrictive press, offering optional interpretations (e.g., “Do you think he’s happy? Do you think he’s sad?”). If the participant still did not respond appropriately, the confederate resolved the probe, stating the correct interpretation (e.g., “He seems sad.”). Finally, in response to the participant’s or confederate’s identification of the interaction partner’s emotional or preferential intent, the partner would affirm the correct interpretation (e.g., again expressing intense disappointment in the case of the robot, or nodding solemnly and saying, “Yeah, I’m sad,” in the case of the adult).

The robot and the adult stimuli’s social expressions were conveyed using body language, pseudo-verbal or minimally verbal (respectively), and vocal prosodic indications. The adult interaction partner was careful not to explicitly declare his or her communicative intent; for instance, rather than saying, “I feel disappointed,” she or he would sigh and say, “Oh, man.” (See Table 1).

Computer Game Interactional Condition

At the time of this study’s data collection (Spring through Fall 2010), touchscreen technology was relatively novel, only having recently emerged in consumer products. For instance, the first Apple iPad touchscreen computer was released in April 2010, and by November 2010, there only were an estimated 15.4 million iPhones (all touch-enabled) in use in the United States, out of a total of at least 234 million mobile phones in the US (Dediu 2011). We structured the computer game condition to involve little social interaction, in order to evaluate our broader hypothesis that in spite of shared novelty and sophistication in the touchscreen computer’s and robot’s technologies, contrasting amounts of social embedding in interactions with the technologies would elicit contrasting amounts and qualities of social behavior from participants.

In the computer game condition, the confederate explained the goal of the tangrams game, and showed the participant how to manipulate the tangram objects using his or her finger, or the touchscreen’s stylus if the participant requested, and then stopped initiating interaction, allowing the child to play the game at his or her own initiation and pace. If the participant asked for assistance, the confederate responded verbally or with minimal demonstration to answer the participant’s question. Also, even if the participant did not ask for help but apparently struggled to understand the puzzle, to strategize about a particularly challenging portion of the puzzle, or to manipulate a tile, then the confederate verbally offered assistance. All children were presented with the same three puzzles, in

consistent order of increasing difficulty, but were allowed to select alternate puzzles if they requested.

Interview-and-Play Sessions

We interleaved a total of four interviews before, after, and between the interactional conditions, beginning with an interview preceding the first interactional condition. Each participant interacted with a single experimenter for all four interviews. Interviews maintained consistent, loose structure, and concluded with imaginative play with miniature wooden dolls or with stuffed animal toys, and allowed participants rest from interactional conditions.

Dependent Variables

We counted the number of utterances participants produced during the interactional conditions, and judged to whom each utterance appeared to be directed. Number of utterances has been shown to be a useful metric in tracking the effects of social and communicative therapies (Koegel et al. 1987b; Maione and Mirenda 2006). An utterance was defined as a verbal production that either expresses a complete proposition (subject + predicate) or is followed by more than 2 s of silence. Utterances were transcribed from video recordings by the first author, and then were confirmed by an independent rater. Following transcription the first author judged the intended audience or recipient of each utterance to be the confederate, the adult partner, the robot, the computer game, some combination of the previous, the participant him- or herself, or indeterminable. Judgments of all utterances’ recipients were confirmed by an independent rater (agreement was 96 %, $K = 0.88$, $p < 0.0001$).

Results

More Speech While Interacting with Robot (Fig. 3)

A repeated-measures two-factor ANOVA (condition \times order, with condition repeating) revealed a main effect of interactional condition (robot, adult, or touchscreen computer game) on the total number of utterances produced by each participant within each interactional condition, $F(1.9, 33.4) = 8.13$, $p < 0.001$, but no main effect of order of presentation of interactional conditions, $F(5, 18) = 0.46$, and no interaction effect between interactional condition and order, $F(9.3, 33.4) = 1.12$.

One-tailed paired t tests showed that participants produced more utterances during the robot ($M = 43.0$, $SD = 19.4$) than the adult condition ($M = 36.8$, $SD = 19.2$), $t(23) = 1.97$, $p < 0.05$, and more in either

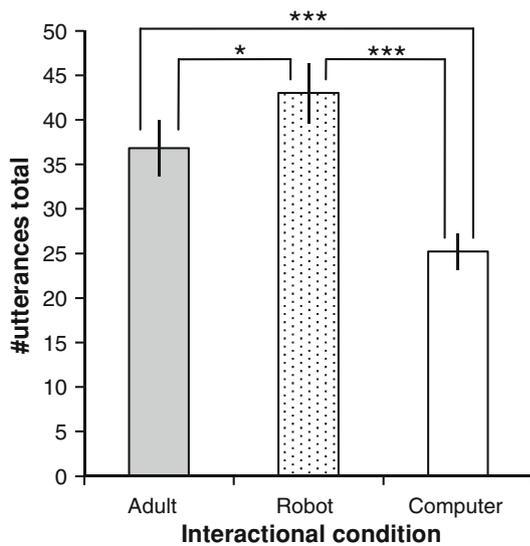


Fig. 3 Bars show means, over distributions of 24 children with ASD, of total number of utterances produced by participants in the adult (left), robot (center), and computer game (right) conditions. Error bars are ± 1 SE. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

the robot ($t(23) = 4.47, p < 0.001$) or adult conditions ($t(23) = 3.61, p < 0.001$) than in the touchscreen computer game condition ($M = 25.2, SD = 13.4$).

More Speech Directed Toward the Confederate, when Interacting with the Robot (Fig. 4)

The number of utterances directed toward the confederate varied with interactional condition, $F(1.8, 33.0) = 3.46,$

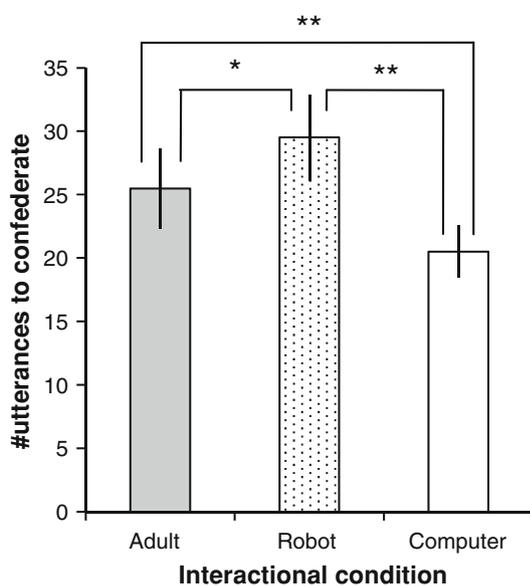


Fig. 4 Bars show means, over 24 children with ASD, of number of utterances directed toward the confederate, in the adult (left), robot (center), and computer game (right) conditions. Error bars are ± 1 SE. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

$p < 0.05$, demonstrated using a repeated-measures two-factor ANOVA (interactional condition \times order, with condition repeating). There was no main effect of order, $F(5, 18) = 0.48$, or of interaction between interactional condition and order, $F(9.2, 33.0) = 0.967$.

Children with ASD directed a higher number of utterances to the confederate in the robot ($M = 29.5, SD = 16.6$) than in the adult condition ($M = 25.5, SD = 15.5$), $t(23) = 1.87, p < 0.05$, and more in both the robot ($t(23) = 3.05, p < 0.01$) and adult ($t(23) = 2.15, p < 0.01$) conditions than in the touchscreen computer game condition ($M = 20.5, SD = 10.1$).

More Speech Directed to Robot and Adult than to Computer Game Interaction Partner; Amount of Speech Directed to Robot Comparable to Amount Directed to Adult (Fig. 5)

A repeated-measures two-factor ANOVA (interactional condition \times order, with condition repeating), revealed that the number of utterances directed toward the interaction partner (robot, adult, or touchscreen computer game) varied with interactional condition, $F(1.5, 26.9) = 15.20, p < 0.001$. However, there was no effect of order of condition presentation, $F(5, 18) = 0.86, p > 0.05$, or of the interaction between condition and order, $F(7.5, 26.9) = 0.50, p > 0.05$.

There were significantly more utterances directed toward the robot ($t(23) = 5.40, p < 0.001$; one-tailed t test) and toward the adult ($t(23) = 8.22, p < 0.001$; one-

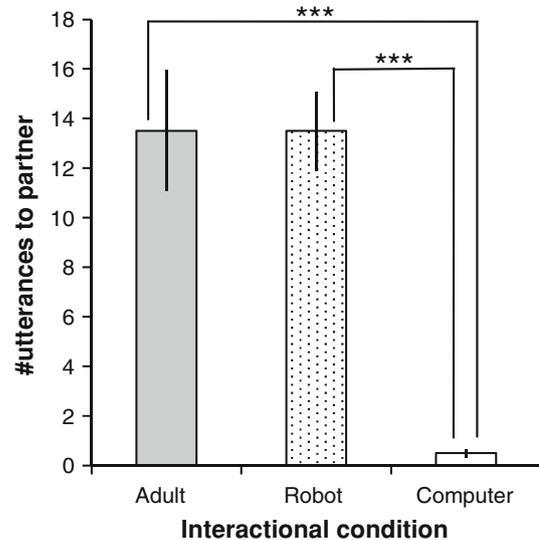


Fig. 5 Bars show means, over 24 children with ASD, of number of utterances directed toward the adult (left), robot (center), and computer game (right) conditions. Error bars are ± 1 SE. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Participants directed a comparable number of utterances to the adult partner as they did to the robot partner

tailed *t*-test) than toward the touchscreen computer game ($M = 0.5$, $SD = 0.8$). There was no difference, $t(23) = 0.02$, in the number of utterances directed toward the interaction partner in the robot condition ($M = 13.5$, $SD = 12.0$) as compared to the adult condition ($M = 13.5$, $SD = 7.8$).

Discussion

We found that children with ASD spoke more, in general, while interacting with a social robot than with another adult or a novel, touchscreen computer game. Utterance counts have been shown to be useful in measuring the effects of social and communicative therapies (Koegel et al. 1987b; Maione and Mirenda 2006). It should come as no surprise that the robot and adult elicited greater verbal interaction than the computer game, given that the computer game interaction condition was not designed to encourage social interaction. What is most interesting is our finding that a social robot elicits more speech than another human.

Between the adult and robot conditions, we found no difference in the amounts of speech children directed to the adult and robot interaction partners, respectively, and no difference in the number of utterances not directed to anybody. Rather, the increase in total speech found in the robot condition can be attributed to an increase of speech directed toward another adult, the confederate. One possible explanation for the absence of difference in the amount of speech to the robot and to the adult may be that the structure of the associated interactional conditions severely limited the speech the adult was allowed to produce in order to match the limited verbal capabilities of the robot. In this sense, the protocol was designed to support more verbal interaction with the confederate than with the interaction partners.

The robot's greater efficacy in eliciting utterances toward the confederate appears to be due to the excitement and interest (that is, preference) children spontaneously expressed for it, over the adult interaction partner. Qualitatively we observed that participants verbalized conjectures and asked questions about how the robot works, whether or not the robot "is real," and what the robot was doing throughout the robot condition. The children also spontaneously asked for permission to, or stated their interest in, touching or playing with the robot. In short, we attribute the robot's greater facilitation of utterances to the participants' greater curiosity about the robot than about the adult interaction partner, during respective conditions.

Heightened verbalization during the robot condition may also reflect the effects of the robot's embedding of social interaction into engagement with it. Our protocol was designed to reinforce interaction with both the

confederate and the interaction partner, but as explained previously, the controlled structure of the protocol allowed the confederate greater flexibility in speaking with participants than it did the adult or robot interaction partners. In this sense, this design better reinforced verbal interaction with the confederate than with interaction partners. This may explain why we saw a difference in the amount of speech to the confederate, between adult and robot conditions; and why we did not see any difference in speech to the respective interaction partners.

Our findings suggest potential utility in communication and social skill interventions for children with ASD. The ultimate goal of such interventions for children with ASD is to improve their ability to interact socially. We have shown that interaction with a social robot elicits speech directed socially toward an adult confederate, not just toward the robot itself, and not undirected speech. In other words, of the three interaction partners tested, the robot best motivates or facilitates an ecologically useful social behavior—interaction with another person—not just social interaction with objects.

This is the first controlled study, over a statistically powerful sample, to demonstrate a social robot's ability to facilitate social interaction with another person. This is also the first study to show this effect for older and higher-functioning children with ASD, whereas previous demonstrations have been presented in small-number case studies of younger children with lower functioning (Feil-Seifer and Matarić 2009; Kozima et al. 2009).

Social robots may draw a comparison with assistive animals, which also elicit social behavior during interaction. It is worth noting that robots have unique advantages over trained animals in that (1) robots can be highly customizable in form and behavior, (2) therapists and parents can control or (if need be) stop a robot instantly and with ease, and (3) robots can be produced in volume at potentially far smaller cost than that required to train assistive animals.

Previous studies of embedded reinforcers have demonstrated social improvements over the course of lengthy therapy sessions, repeated over several weeks. It is remarkable that the observed increases in verbal interaction afforded by a social robot occur immediately, during interaction with the robot. Further research must be conducted in the long-term durability of social robots' embedded reinforcing effects.

Limitations and Future Directions

One limitation of this study is that we examined only the quantity, and not the semantic content or communicative function, of utterances under different conditions. A

cursor examination suggested that the content of utterances varied across participants. For example, the number of spontaneous comments and questions about the robot (e.g., “Is he real?” “Did you build it?” “Does it have a battery?” “If there was another robot, they would be friends.”) ranged from zero to twelve. We plan more sophisticated pragmatic and semantic analyses in the future to better understand the nature of the increases in verbal production that we have observed in the robot over the adult condition.

It is also important to note that short-term effects of interaction with a robot do not necessarily predict long-term effects. This was demonstrated, for instance, in a 2-week field trial of school-aged, typically developing children’s daily interaction, with a social robot, in which most children’s interactions with the robot declined in the second week (Kanda et al. 2004). Because any effective therapy requires repeated opportunities to practice target behaviors, our study of short-term effects cannot alone indicate the utility of a social robot as a therapeutic tool. Long-term study of motivation, reinforcement, and pedagogical impact are required. While our study cannot speak to long-term effects, our encouraging short-term findings motivate investment in longitudinal studies. We are hopeful that as technology improves, social robots’ interactive behaviors will become increasingly complex and adaptable to relationships with individuals. Kanda et al. suggested that children who shared more “common ground” with their robot sustained interaction over time with the robot (Kanda et al. 2004). Thus, as robots increasingly support rich repertoires of social behavior, it is possible they will likewise increasingly support sustained interaction.

Our original intent in comparing a robot with an adult was to compare the robot against an agent operating at the upper limit of social capability. However, the adult interaction partner was unfamiliar to participants. A familiar adult might be considered even more capable socially, with respect to individual participants. Small numbers of children with autism have been observed to prefer interaction with a robot to that with an unfamiliar adult behaving like a robot (Dautenhahn and Werry 2004; Robins et al. 2006), and children with autism have also been shown to prefer interacting with their caregivers, to interacting with strangers (Sigman and Mundy 1989). Our work compared triadic therapy-like interactions with an unfamiliar adult and unfamiliar robot, and with an unfamiliar therapist-like confederate. Our study cannot speak to differences between a robot and a familiar adult, or to triadic interactions with a robot and a familiar therapist. Therefore, the effects of familiarity on interaction with an adult merit future investigation.

We chose the Wizard of Oz robot control paradigm in order to examine responses to a social robot operating at

the upper bounds of its social interaction capabilities. We share an aspiration, with many contemporary researchers in the field of human-robot interaction research, of eventually developing technologies that give social robots truly automatic perception of, and response to, their environments and interaction partners’ behaviors. At present, however, Wizard of Oz remains a standard design paradigm, given that state-of-the-art technologies do not yet afford highly reliable automatic speech recognition or other socially important perceptual capabilities, especially not for individuals with widely varying verbal and social abilities and behaviors. Currently, training any automatic perceptual system would be especially difficult, given the vastly heterogeneous presentations of social behaviors we expect to encounter among children with ASD; automatic perception must wait for advances in our understanding and description of typical and atypical social behaviors (Volkmar and Klin 2005). Demonstrations, like those presented in this article, of the utility of interaction with richly socially interactive robots, motivate further research into the automation of these rich social capabilities.

Previously the benefits had been shown (in small numbers of children with ASD) of using social interaction to deliver a preferred reinforcer (Koegel et al. 2009). We suggest that social robots may additionally enable a unique type of beneficial embedding, by which social interaction not only *delivers* the preferred reinforcer (e.g., a person presents a child with a robot), but also that the preferred reinforcer *is itself the object and source of social interaction*, not requiring an external social agent to deliver the preferred reinforcer. Social robots may bridge interest in novel technology with motivation for social interaction: if interaction with a social robot itself is rewarding to an individual child, then social interaction more generally may become more rewarding for that child. As technology develops to allow social robots greater and more flexible range of interaction, further research should explore whether they can elicit improved social behavior in children with low social motivation, and can then transfer this behavior to human social partners. Our sample population included only highly functioning individuals; future research should examine whether social robots offer a unique therapeutic support to children with lower functioning.

Finally, our work is just a first step in the larger goal of providing new tools for clinicians to use in interventions for individuals with ASD, not as alternatives to clinicians or trained peers, but as supplements. The true test of the efficacy of social robotics in facilitating social-communicative improvements in children with ASD will require larger field studies comparing long-term learning and skill generalization in the presence and absence of social robots. These studies are ongoing.

Conclusions

We have demonstrated that for a sample population of school-aged children with high functioning ASD, a social robot can elicit greater verbalization than a social (but less preferred) interaction partner, an adult human. We have shown that a robot elicits greater verbalization than a preferred but asocial interaction partner, a computer game. More importantly, a social robot increases social interaction with another person, more than an adult or a computer game does. These findings suggest that robots, with appropriate clinical guidance, may make useful supplements to communication and social skills interventions by facilitating social interaction with an adult, and by eventually being developed into uniquely embedded social reinforcers.

Acknowledgments This material is based upon work supported by Microsoft Research and the National Science Foundation under grants No. 0835767, No. 0968538, and No. 1117801, as well as by an NSF Expedition in Computing (award #1139078). We thank Kathleen Koenig for clinical insights into experimental design; Taylor Mae Brown and Hilary Rose Barr for meticulous video annotation; Julie Wolf, Elizabeth Schoen Simmons, and Maysa Akbar for clinical supervision of participants; Rebecca Loomis and Jonathan Tirrell for assistance with data management; Erin MacDonnell and Heidi Seib for assistance with recruitment; and Jonathan A. Kelley for insightful review of this manuscript.

References

- American Psychiatric Association. (2000). *Diagnostic and statistical manual of mental disorders: DSM-IV-TR*. Arlington, VA: American Psychiatric Association.
- Bosseler, A., & Massaro, D. W. (2003). Development and evaluation of a computer-animated tutor for vocabulary and language learning in children with autism. *Journal of Autism and Developmental Disorders*, 33(6), 653–672. doi:10.1023/B:JADD.000006002.82367.4f.
- Carter, A. S., Davis, N. O., Klin, A., & Volkmar, F. R. (2005). Social development in autism. In F. R. Volkmar, R. Paul, A. Klin, & D. J. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders* (3rd ed., Vol. 1, pp. 312–334). Hoboken, NJ: John Wiley and Sons.
- Dautenhahn, K., & Werry, I. (2004). Towards interactive robots in autism therapy: Background, motivation and challenges. *Pragmatics and Cognition*, 12(1), 1–35. doi:10.1075/pc.12.1.03dau.
- Dediu, H. (2011, 2008). *Nearly 75% of iPhones are in use outside the US | asymco*. Retrieved from <http://www.asymco.com/2011/01/08/nearly-75-of-iphones-are-in-use-outside-the-us/>.
- Diehl, J. J., Schmitt, L. M., Villano, M., & Crowell, C. R. (2012). The clinical use of robots for individuals with autism spectrum disorders: A critical review. *Research in Autism Spectrum Disorders*, 6(1), 249–262. doi:10.1016/j.rasd.2011.05.006.
- Duquette, A., Michaud, F., & Mercier, H. (2008). Exploring the use of a mobile robot as an imitation agent with children with low-functioning autism. *Autonomous Robots*, 24(2), 147–157. doi:10.1007/s10514-007-9056-5.
- Elliott, C. D. (2007). *Differential ability scales-II (DAS-II)*. San Antonio, TX: Pearson. Retrieved from <http://www.pearsonassessments.com/HAIWEB/Cultures/en-us/Productdetail.htm?Pid=015-8338-820>.
- Feil-Seifer, D., & Mataric, M. J. (2009). Toward socially assistive robotics for augmenting interventions for children with autism spectrum disorders. In O. Khatib, V. Kumar, & G. J. Pappas (Eds.), *Experimental robotics* (Vol. 54, pp. 201–210). Berlin, Heidelberg: Springer. Retrieved from <http://www.springerlink.com/content/12k004r536p73nl6/>.
- Hetzroni, O., & Tannous, J. (2004). Effects of a computer-based intervention program on the communicative functions of children with autism. *Journal of Autism and Developmental Disorders*, 34(2), 95–113. doi:10.1023/B:JADD.0000022602.40506.bf.
- Innvo Labs. (2012). *PLEOworld*. Retrieved February 22, 2012. From http://www.pleoworld.com/pleo_rb/eng/index.php.
- Joseph, R. M., & Tager-Flusberg, H. (1997). An investigation of attention and affect in children with autism and down syndrome. *Journal of Autism and Developmental Disorders*, 27(4), 385–396. doi:10.1023/A:1025853321118.
- Kanda, T., Hirano, T., Eaton, D., & Ishiguro, H. (2004). Interactive robots as social partners and peer tutors for children: A field trial. *Human-Computer Interaction*, 19(1), 61–84. doi:10.1207/s15327051hci1901&2_4.
- Kanner, L. (1943). Autistic disturbances of affective content. *Nervous Child*, 2, 217–250.
- Kim, E. S., Leyzberg, D., Tsui, K. M., & Scassellati, B. (2009). How people talk when teaching a robot. *Proceedings of the 4th ACM/IEEE international conference on human robot interaction, HRI'09* (pp. 23–30). New York, NY: ACM. doi:10.1145/1514095.1514102.
- Klin, A., Lang, J., Cicchetti, D. V., & Volkmar, F. R. (2000). Brief report: Interrater reliability of clinical diagnosis and DSM-IV criteria for autistic disorder: Results of the DSM-IV autism field trial. *Journal of Autism and Developmental Disorders*, 30(2), 163–167. doi:10.1023/A:1005415823867.
- Koegel, R. L., Dyer, K., & Bell, L. K. (1987a). The influence of child-preferred activities on autistic children's social behavior. *Journal of Applied Behavior Analysis*, 20(3), 243–252. doi:10.1901/jaba.1987.20-243.
- Koegel, L. K., Koegel, R. L., Harrower, J. K., & Carter, C. M. (1999). Pivotal response intervention I: Overview of approach. *Research and Practice for Persons with Severe Disabilities*, 24(3), 174–185. doi:10.2511/rpsd.24.3.174.
- Koegel, R. L., O'Dell, M. C., & Koegel, L. K. (1987b). A natural language teaching paradigm for nonverbal autistic children. *Journal of Autism and Developmental Disorders*, 17(2), 187–200. doi:10.1007/BF01495055.
- Koegel, R. L., Vernon, T. W., & Koegel, L. K. (2009). Improving social initiations in young children with autism using reinforcers with embedded social interactions. *Journal of Autism and Developmental Disorders*, 39(9), 1240–1251. doi:10.1007/s10803-009-0732-5.
- Kozima, H., Michalowski, M. P., & Nakagawa, C. (2009). Keepon: A playful robot for research, therapy, and entertainment. *International Journal of Social Robotics*, 1(1), 3–18. doi:10.1007/s12369-008-0009-8.
- Kozima, H., Nakagawa, C., & Yasuda, Y. (2005). Interactive robots for communication-care: A case-study in autism therapy. *IEEE International workshop on robot and human interactive communication, 2005*. ROMAN 2005 (pp. 341–346). Presented at the IEEE international workshop on robot and human interactive communication, 2005. ROMAN 2005, IEEE. doi:10.1109/ROMAN.2005.1513802.
- Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., et al. (2000). 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. doi:10.1023/A:1005592401947.
- Maione, L., & Mirenda, P. (2006). Effects of video modeling and video feedback on peer-directed social language skills of a child with autism. *Journal of Positive Behavior Interventions*, 8(2), 106–118. doi:10.1177/10983007060080020201.
- Martin, F., & Farnum, J. (2002). Animal-assisted therapy for children with pervasive developmental disorders. *Western Journal of Nursing Research*, 24(6), 657–670. doi:10.1177/019394502320555403.
- Mesibov, G. B. (1992). Treatment issues with high-functioning adolescents and adults with autism. In E. Schopler & G. B. Mesibov (Eds.), *High-functioning individuals with autism*, current issues in autism (pp. 143–156). New York: Springer.
- Mullen, E. M. (1995). *Mullen scales of early learning (AGS.)*. San Antonio, TX: Pearson.
- Mundy, P., Sigman, M. D., & Dawson, G. (1989). Specifying the nature of the social impairment in autism. *Autism: New perspectives on nature, diagnosis, and treatment* (pp. 3–21).
- Mundy, P., Sigman, M., Ungerer, J., & Sherman, T. (1986). Defining the social deficits of autism: The contribution of non-verbal communication measures. *Journal of Child Psychology and Psychiatry*, 27(5), 657–669. doi:10.1111/j.1469-7610.1986.tb00190.x.
- Parsons, S., & Mitchell, P. (2002). The potential of virtual reality in social skills training for people with autistic spectrum disorders. *Journal of Intellectual Disability Research*, 46(5), 430–443. doi:10.1046/j.1365-2788.2002.00425.x.
- Paul, R. (2008). Interventions to improve communication in autism. *Child and Adolescent Psychiatric Clinics of North America*, 17(4), 835–856.
- Redefer, L., & Goodman, J. (1989). Brief report: Pet-facilitated therapy with autistic children. *Journal of Autism and Developmental Disorders*, 19(3), 461–467. doi:10.1007/BF02212943.
- Robins, B., Dautenhahn, K., & Dubowski, J. (2006). Does appearance matter in the interaction of children with autism with a humanoid robot? *Interaction Studies*, 7(3), 509–542.
- Robins, B., Dautenhahn, K., te Boekhorst, R., & 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. doi:10.1007/s10209-005-0116-3.
- Scassellati, B. (1996). *Mechanisms of shared attention for a humanoid robot. Embodied Cognition and Action: Papers from the 1996 AAAI Fall Symposium* (Vol. 4, p. 21).
- Scassellati, B. (2005). Quantitative metrics of social response for autism diagnosis. *IEEE international workshop on robot and human interactive communication, ROMAN 2005* (pp. 585–590). doi:10.1109/ROMAN.2005.1513843.
- Scassellati, B., Admoni, H., & Mataric, M. J. (2012). Robots for use in autism research. *Annual Review of Biomedical Engineering*, 14, 275294.
- Sigman, M., & Mundy, P. (1989). Social attachments in autistic children. *Journal of the American Academy of Child and Adolescent Psychiatry*, 28(1), 74–81.
- Sparrow, S. S., Cicchetti, D. V., & Balla, D. A. (2005). *Vineland adaptive behavior scales (Vineland-II)* (2nd ed.). San Antonio, TX: Pearson. Retrieved from <http://psychcorp.pearsonassessments.com/HAIWEB/Cultures/en-us/Productdetail.htm?Pid=Vineland-II>.
- Stanton, C. M., Kahn Jr., P. H., Severson, R. L., Ruckert, J. H., & Gill, B. T. (2008). Robotic animals might aid in the social development of children with autism (p. 271). ACM Press. doi:10.1145/1349822.1349858.
- Steinfeld, A., Jenkins, O. C., & Scassellati, B. (2009). The oz of wizard: Simulating the human for interaction research. *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction, HRI'09* (pp. 101–108). San Diego, CA: ACM. doi:10.1145/1514095.1514115.
- Strickland, D. (1997). Virtual reality for the treatment of autism. In G. Riva (Ed.), *Virtual reality in neuro-psycho-physiology: Cognitive, clinical and methodological issues in assessment and rehabilitation* (pp. 81–86). Amsterdam: IOS Press.
- Tager-Flusberg, H., Paul, R., & Lord, C. (2005). Language and communication. In F. R. Volkmar, R. Paul, A. Klin, & D. J. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders* (3rd ed., Vol. 1, pp. 335–364). Hoboken, NJ: John Wiley and Sons.
- Tapus, A., Mataric, M. J., & Scassellati, B. (2007). Socially assistive robotics (Grand challenges of robotics). *Robotics Automation Magazine, IEEE*, 14(1), 35–42. doi:10.1109/MRA.2007.339605.
- Volkmar, F. R., & Klin, A. (2005). Issues in the classification of autism and related conditions. In F. R. Volkmar, R. Paul, A. Klin, & D. J. Cohen (Eds.), *Handbook of autism and pervasive developmental disorders* (3rd ed., Vol. 1, pp. 335–364). Hoboken, NJ: John Wiley and Sons.
- Volkmar, F. R., Lord, C., Bailey, A., Schultz, R. T., & Klin, A. (2004). Autism and pervasive developmental disorders. *Journal of Child Psychology and Psychiatry*, 45(1), 135–170. doi:10.1046/j.0021-9630.2003.00317.x.
- Werry, I., & Dautenhahn, K. (1999). Applying mobile robot technology to the rehabilitation of autistic children. *Proceedings of the 7th symposium on intelligent robotic systems (SIRS99)*.