Auton Robot
DOI 10.1007/s10514-016-9554-4

A low-cost socially assistive robot and robot-assisted intervention for children with autism spectrum disorder: field trials and lessons learned

Laura Boccanfuso[1,2] · Sarah Scarborough[1] · Ruth K. Abramson[1] · Alicia V. Hall[1] · Harry H. Wright[1] · Jason M. O'Kane[1]

Received: 27 March 2015 / Accepted: 7 February 2016

Abstract Recent research has employed socially assistive robots as catalysts for social interaction and improved communication in young children with autism spectrum disorder (ASD). Studies describe observed therapeutic outcomes such as increased speech, social interaction, joint and directed attention, but few detail a robot-inclusive protocol which evaluates a set of robot tasks using widely-accepted, clinical assessments to evaluate the efficacy of the approach. In this study, we employed a low-cost, toy-like robot prototype with safety features such as a snap-off head and two snap-off arms, a camera for face, hand detection and session recording, two autonomous games and a teleoperated mode. We then developed and tested a new, robot-assisted intervention. Eight study participants and three controls diagnosed with ASD and a speech deficiency were recruited. The study group received pre-, post-intervention measures with the Vineland Adaptive Behavioral Scale II (VABS-II), mean length spontaneous utterance determination (MLSUD), motor imitation scale (MIS), unstructured imitation assessment (UIA) and Expressive Vocabulary Test 2 (EVT-2) and participated in twelve 30 min interventions. To explore the efficacy of the robot and new robot-assisted intervention we (1) measured improvements in spontaneous speech, communication and social skills using standard measures of performance, (2) compared improvements observed with a study group receiving the robot-assisted intervention with a control group receiving speech therapy but no robot-assisted intervention and, (3) validated a set of robot behaviors that may inform an integrated, cross-platform, approach for incorporating an autonomous, robot-assisted ASD intervention within a clinical methodology. Paired-samples t test results indicate significantly improved adaptive functioning in the VABS-II socialization and communication domains, MLSUD, UIA Social Interaction, UIA Requesting, and UIA Joint Attention domains. Between-group analyses also suggest significant improvement in VABS-II Play and Leisure, Receptive Language subdomains and trends in VABS-II Coping Skills and Interpersonal Scale subdomains.

Keywords Socially assistive robots · Autism · ASD · Interaction design · Communication · Socialization

1 Introduction

This multidisciplinary pilot study suggests that significant improvements in communication and social interaction can be achieved using an augmented intervention with a widely accessible, minimally-actuated socially assistive robot (SAR). This study also provides insight as to the viability of a new robot-assisted methodology and the definition of specific autonomous behaviors that may lead to longer-term, cross-platform utility of a SAR within a clinical intervention. Further, our results suggest that the presented social games-based paradigm can produce social exchanges that generalize from robot–child to child–other interactions.

Because the integration of robots into clinical methodologies is still mainly a theoretical premise, determining which robot behaviors are worth automating is challeng-

Electronic supplementary material The online version of this article (doi:10.1007/s10514-016-9554-4) contains supplementary material, which is available to authorized users.

This is one of several papers published in *Autonomous Robots* comprising the "Special Issue on Assistive and Rehabilitation Robotics".

✉ Laura Boccanfuso
laura.boccanfuso@yale.edu

1 University of South Carolina, Columbia, USA

2 Yale University, New Haven, USA

ing. Sustained, productive child-robot interactions leading to improved therapeutic outcomes will benefit from validating task-specific autonomous behaviors that can be widely integrated on existing robots, within a comprehensive therapeutic clinical protocol. Findings from this study identified several effective interaction model features which may be generalized to other robots and in other contexts are summarized in Sect. 10.

We designed and conducted a pilot field study to achieve three primary objectives: (1) to quantitatively assess the effectiveness of a new, robot-assisted intervention for increasing spontaneous speech, overall communication and social skills in children with ASD, (2) to compare communication and social skills increases obtained through therapy(ies) supplemented with a robot-assisted intervention to increases achieved without robot intervention and, (3) to identify specific, well-defined robot behaviors that augment the goals of clinical interventions for ultimately advancing robot autonomy in this domain. Our field study demonstrates improvements in social interaction, communication and speech, measured using widely accepted, established measures of adaptive functioning including the VABS-II (Vertue 2007), MLSUD, motor imitation scale (MIS) (Stone et al. 1997a), UIA (Ingersoll and Lalonde 2010a) and Expressive Vocabulary Test 2 (EVT-2) (Gurley 2011).

The remainder of the paper includes a summary of related work in Sect. 2 and a brief description of the robot prototype, previously-implemented software and interactive games developed for conducting field trials in Sect. 3. A description of the Experimental design is included in Sect. 4, where we describe the target population, pretest procedures, initial visit and data collection and second pretest visit. In Sect. 5 we detail field tests conducted over the course of 6 months, the lessons we learned and the substantive changes that were made to improve the interaction design. Within-group and between-group analyses are described in Sect. 6, results are presented in Sect. 7 and a discussion of results is presented in Sect. 8. Finally, limitations of the study and conclusions are included in Sects. 9 and 10, respectively.

2 Related work

The robotics community has recently seen a rapid acceleration of research in the development and testing of SAR for children with ASD (Warren et al. 2014; Grecez et al. 2014; Bekele et al. 2013; Scassellati et al. 2012; Robins and Dautenhahn 2014). Much of this work effectively contributes to diminishing the existing gap between the design and development of robots that perform pro-social and pro-communicative behaviors and generalizable therapeutic outcomes (Huskens et al. 2013; Kim et al. 2013) and (Kim et al. 2012).

These recent studies contribute to the broader corpus of work in this domain in three primary ways. First, by targeting characteristic social and communicative difficulties through robot-assisted activities and game play, research in this domain moves closer to converging on a set of validated, robot behaviors to effectively promote specific therapeutic goals. Second, by ultimately defining a set of validated robot behaviors, we will be more able to translate exploratory teleoperated activities to a suite of autonomous interactions that can be implemented on a wide range of new and existing robots to achieve common treatment objectives. Finally, the contribution of descriptive field studies which use validated objective measures appropriate to the therapy for which they are intended, will leverage the utility of robots for the clinical community and lead to more aptly designed clinical methodologies.

Early work by researchers at the University of Hertfordshire helped to lead the way to defining the role of robots in autism therapy by exploring the evolving role of robots and the types of play scenarios that are most conducive to an effective intervention for children with ASD (Dautenhahn 1999) and (Project 2011). Although robots have traditionally been designed to perform a task or to carry out a specific set of actions, researchers involved in the AURORA project made the distinction that the emphasis for robots in autism therapy will be on the interactions and expression of actions, not necessarily the completion of a given task. This is a fundamental departure from the standard design of robots and means that even simple tasks or actions can have significant impact if they are successful in producing the intended response.

A number of field studies describe the therapeutic benefits of using interactive robots in therapy such as increased speech, social interaction, joint and directed attention, (Yun et al. 2015; Welch et al. 2010) and (Feil-Seifer and Matarić 2008), but few studies exist which quantify observed communication increases using assessment instruments accepted by the autism and speech therapy communities (Diehl et al. 2012). Even fewer provide a statistical evaluation comparing the benefits received through speech therapy with those obtained through an additional robot intervention. No field studies to our knowledge have employed a robot prototype that is sufficiently robust and reasonably easy to operate for cooperative use by the therapist in the clinic, the family at home, and the special education teacher at school for promoting generalizable results. Further, although a fairly exhaustive search of robots previously used or in active use for ASD research yields close to 25 robots (Cabibihan et al. 2013), there is a paucity of data available for robot-assisted studies which explore employing a robot as an integrated part of any conventional intervention.

Our research presents a new, field-tested, robot-assisted methodology for promoting communication and social skills in children with ASD. This multidisciplinary effort lever-

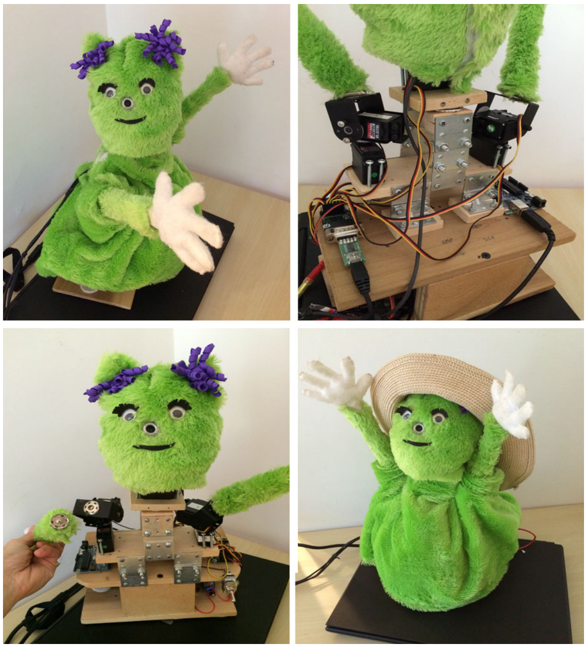


Fig. 1 [*top left*] CHARLIE the robot. [*top right*] Internal structure showing servos, pan/tilt assembly, servo controller and microcontroller. [*bottom left*] Snap-off arm [*bottom right*] CHARLIE with hat

ages the expertise of a clinical speech pathologist, ASD experts and computer scientists to deliver a novel interaction schema demonstrated to help promote standard therapeutic outcomes. While a single approach will not be effective with all children affected with ASD, this paper describes field study results that define salient features of an interaction paradigm that were shown to be effective for the studied group of children with ASD.

3 Approach and methodology

The robot prototype used for this study is named CHARLIE (Fig. 1) and, along with a set of interactive games, was introduced in our earlier research (Boccanfuso and O'Kane 2011). In this field study, we employ CHARLIE and previously developed interactive imitation games and introduce a new set of teleoperated games motivated by observations made during preliminary testing. A summary of the robot's hardware, software implementation and interaction design is provided below.

3.1 Robot prototype

3.1.1 Motivation for robot design

The design of the robot was motivated by three major objectives. First, the fundamental structure was designed to be kinematically simplistic, with few degrees of freedom. By providing just the minimal range of motion needed to perform imitation actions, the noise generated by onboard servo motors was limited and the extent of potential mechanical damage resulting from physical manipulation was reduced. Second, given the critical role of free play during a child's early social development, we sought to maximize opportunities for children to freely explore and manipulate the robot without excessive concern for the physical integrity of the robot. To this end, we integrated snap-off arms and head on the robot and included a base that could be secured to a table. Finally, the outward appearance of the robot was designed to be toy-like to invite the attention of young children with ASD and to avoid being intimidating to the greatest extent possible.

3.1.2 Hardware design

Structural design The robot's hardware includes 6 servos, 3 pan-tilt platforms, an 8 channel servo controller, a consumer-grade webcam and a lithium-polymer battery. The arms and head are each mounted on a pan-tilt platform using large metal snap fasteners and each platform is controlled by two servos. The resulting two degrees of freedom in the robot's arms allow for sufficient range of motion to imitate a wide range of hand poses and the two degrees of freedom in the head allow the robot to effectively track the face each participant.

Safety and feedback features The robot's body is padded for safety, and its outer surfaces are covered with a bright green, fur-like material to achieve a non-threatening appearance. During active game play the child's attention is typically focused near CHARLIE's hands, so one LED is embedded in each of the hands to provide positive feedback during interactive games. A speaker is also included in the CHARLIE's body in order to provide optional auditory instructions for playing interactive games and positive feed-back. Exclusive of the computing hardware, the retail cost of the robot's components is approximately 200 USD. In a production version of this robot, a computer could be integrated into the robot's body, or users could connect via USB to a standard laptop or desktop PC.

3.2 Previously implemented software

The initial interactive software was designed to promote two fundamental skills known to be closely linked to communication—turn-taking and imitation. Three auto nomous game modes were implemented to appeal to children with varying levels of communication and social skill. To accommodate children who were reluctant to play with an autonomous robot or for those who would benefit from an extended period of exploration before engaging in the interactive games, a teleoperated mode was implemented. For

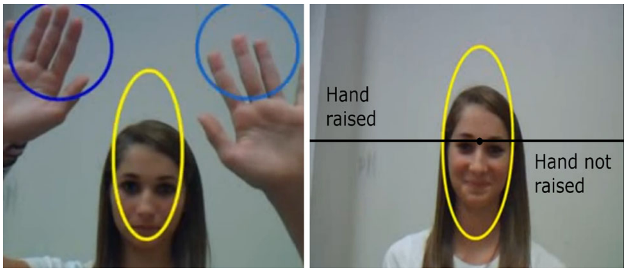


Fig. 2 [*left*] Face and hand detection from the robot's perspective. [*right*] Face centroid and *y*-axis used to determine a hand raised event

those who were ready to play directly with the robot, but who were not necessarily ready to play a cooperative game with another person, the single-player interactive game called "Imitate Me, Imitate You", was created. Finally, a two-player interactive game called "Pass the Pose" was created to appeal to those children with an established level of basic imitation and turn-taking, but required more practice with multiple-player games using the robot as a social mediator.

Face and hand detection, tracking The set of auto nomous, interactive games used during the preliminary trials of this study, employed face and hand tracking and is described in detail in our earlier work (Boccanfuso and O'Kane 2011). A brief summary is provided for the reader's reference. First, the robot performed face detection to locate the child's face and to compute the centroid of the detected face in order to orient its head appropriately throughout each interactive game. Next, to enable pose detection and imitation, the robot performed hand detection using a classifier we trained and presented previously (Fig. 2). Finally, the y-axis coordinate of the hand centroid was compared to the y-axis coordinate of the face centroid to determine if each detected hand was raised and x-axis coordinates were compared to determine which hand was raised (Fig. 2). In so doing, the robot detected each instance of a hand raise and responded accordingly.

3.2.1 Preliminary interaction design

Research in robot-assisted autism therapy typically emphasizes specific objectives for ideal human–robot interaction including an increased joint attention, eye contact, child-initiated interactions, verbal and non-verbal communication, turn-taking, imitative game playing and overall use of language. We designed and implemented a teleoperated mode and two interactive games to promote joint attention, imitation and turn-taking.

In the single-player, interactive "Imitate You, Imitate Me" game, the child may either initiate a pose for the robot to imitate ("Imitate Me") or the child may follow the robot's pose ("Imitate You"). During the robot-initiated game mode, the robot assumes a pose with one or both hands raised or peek-a-boo, and actively looks for the child to properly imitate its pose. Once the robot detects a correct mirroring of its position, the robot assumes another randomly-determined pose. In the child-initiated game mode, the robot actively searches for the child to raise one of both hands. Once detected, the robot mirrors the pose. If a pose is not detected within 10 s after the start of game play or after the last pose was assumed, the robot transitions to robot-initiated mode.

The two-player, interactive game begins with the robot describing how to play "Pass the Pose" and asking the first player (seated to the right of the robot) to assume a pose. Once it has detected the pose, the robot indicates that it has learned the pose by saying "O.k., I got it. Now let me try", turns to the second player (seated to the left of the robot), asks the child to imitate and then assumes the same pose learned from the first player. If the second player successfully imitates the pose, the robot responds by saying "You got it!", claps its hands and giggles. If the player does not immediately imitate the correct pose, the robot will ask the child to try again. If the child does not correctly assume the pose after three tries, the robot asks the current player to initiate a new pose and the game continues, this time with the second player initially "passing" the pose to the robot.

3.3 New software

Observations made during preliminary field testing in this study led to the revision of activities included in the final study protocol. A set of autonomous interactive games were previously pilot tested with a small group of typically developing children in a lab setting and were well-received. However, the children who participated in this study's preliminary testing did not find the same games as engaging. One participant expressed initial interest in the imitation games but his attention was not sustained for the duration of the session. Another two participants did not seem to recognize or enjoy the causality of the robot's actions at all. In general, we had a great deal of difficulty getting young children with ASD to sit long enough for the robot to reliably detect their hand movements, to maintain an interaction and to predict between different children and with repeated trials how long each child's attention to task would be. To address these concerns and to pursue a sustainable engagement across multiple sessions, a new protocol and accompanying software were developed and implemented.

3.3.1 Teleoperated games

Before developing and implementing a new set of autonomous robot tasks, particularly due to the exploratory nature of the novel interaction design, we opted to first use a teleoperated mode to control the robot's actions and validate the effectiveness of the new proposed approach with a population of children with ASD. The robot's new role relied heavily on responding contingently to the child's actions, performing salutations and coordinated movements with music. Although these new tasks were teleoperated in this study, each of the included activities may be automated with reasonable modifications. For example, using the face and hand detection software already implemented, the robot can perform salutations autonomously. With the inclusion of simple radio frequency identification (RFID) tags, the robot's contingent responses during the added activities can also be automated. The new game design and a full, detailed description of added tasks are included in Sect. 5.3.

4 Experiment design

The study design was developed to elicit child-led communication and socialization using a set of new, interactive games with a robot prototype. The experiment procedures were carefully devised to ensure an ASD diagnosis verified by a widely accepted, validated diagnostic instrument and to record multiple measures of each child's developmental, communicative and socialization ability. Institutional Review Board approval was sought to conduct this study and obtained on February 19, 2013 under study identification number Pro00-023119, and the study title: "Effectiveness of CHARLIE the Robot for Improving Verbal and Nonverbal Skills in Children with Autism."

4.1 Target population

A group of eight children, between 3 and 6 years of age, who were diagnosed with ASD and a speech deficiency as confirmed by the Autism Diagnostic Observation Schedule (ADOS) and a speech pathologist, respectively, participated in this study. Preliminary testing to evaluate the efficacy of interactive games was also conducted with a group of three children. Children were invited to participate based on the information provided in a prescreening questionnaire, confirmation of an ASD diagnosis and a speech or language delay. A description of the prescreening questionnaire is provided below. A control group consisting of three children between 3 and 6 years of age, diagnosed with an ASD and a documented speech or language delay, also participated. Improvements achieved through speech therapy alone are compared with those achieved with an additional robot-assisted intervention. Children in the control group received speech therapy for the entire 6-week participation in the study but did not receive the robot-assisted intervention whereas most of the children in the study group received the robot-enabled intervention in addition to speech therapy.

4.2 Pretest procedures

To determine whether each child met criteria for the study, a prescreening questionnaire was administered to the prospective participant's caregiver over the phone. The questions on the prescreening form were used to confirm that: (1) a formal diagnosis of an ASD was received (also, who made the diagnosis and when it was made), (2) the child's language ability was delayed for their chronological age (and to what degree), (3) the child's nonverbal communication ability (pointing, shaking/nodding head, etc) was delayed for their chronological age (and to what degree), (4) the child did not have any diagnosed hearing impairment and, (5) the child's current therapy schedule would allow for additional therapy sessions with the robot. Most of the children in the study were concurrently receiving other forms of intervention, but 3 of the 8 children were not receiving any other intervention at the time of their participation in the study.

If the child met criteria for the study, the child was invited to participate in the study and the first face-to-face meeting with her/his caregiver was scheduled at the University of South Carolina (USC) School of Medicine, Department of Neuropsychiatry and Behavioral Sciences. The second meeting and all subsequent sessions with the robot took place at the USC Speech and Hearing Research Center.

4.2.1 Initial visit and data collection

The first meeting with the caregiver(s) of each child participating in the study included the completion and signing of the informed consent, signing of releases for medical records form(s) documenting a diagnosis of an ASD and a speech impairment, completion of the Vineland Adaptive Behavior Scale II (VABS-II) (Vertue 2007), and the Social Communication Questionnaire (SCQ) (Services 2013).

The VABS-II is designed to measure personal, communication and social skills especially for special needs populations such as individuals with mental retardation, autism, and attention-deficit/hyperactivity disorder (ADHD). It is comprised of four domains: communication, daily living skills, socialization, and motor skills. The Communication Domain measures receptive, expressive, and written communication; the Daily Living Skills Domain assesses personal, domestic, and community skills; the Socialization Domain measures interpersonal relationships, play and leisure time, and coping skills; and the Motor Skills Domain measures gross and fine motor skills. Scores from the Communication and Socialization Domains in the VABS-II were used to establish a baseline score for each participant and stored for later comparison with scores recorded at the end of the study.

The SCQ is a brief questionnaire which aids in the evaluation of communication skills and social functioning in children who may have an ASD (Services 2013). Typically, the SCQ is used as a fast way to determine if an individual should be referred to a qualified professional for a complete diagnostic evaluation. The SCQ was used in this study as an additional measure to confirm the child's ASD diagnosis. A cutoff score of 15 or greater was used as an indication of possible ASD; any participant scoring below 15 on the SCQ was not invited to participate in the study.

To confirm the diagnosis of ASD, results from the Autism Diagnostic Observation Schedule (ADOS) (Lord et al. 1989) were requested and obtained. VABS-II questionnaires were also collected from the caregivers of control group participants immediately after consenting to participate in the study and again after their child received a minimum of 1 hour of speech therapy per week for 6 weeks. Composite communication and socialization scores were extracted from each pre- and post-test questionnaire to perform comparative statistical analyses for: within study group data, within control group data and between study group and control group data. Additionally, a total of four mean length spontaneous utterance determination (MLSUD) measures were collected; scores were calculated at the first meeting with the child, during sessions 4 and 8 and at the very end of the 6-week intervention period. Only MLSUD scores from the preliminary and final evaluation sessions were used to perform assessments of increased speech since intermediate sessions featured a significant amount of scripted speech that was used to engage in games with the robot and with co-present others.

Data collection also included videotaping each session in order to assist in documenting each child's progress. Video-recorded sessions were especially useful for calculating MLSUD (Brown 1973) measures throughout the 6-week study period for each child and for making note of any significant changes in the child's response as new parts of the intervention were introduced. The duration of the initial meeting was approximately one hour.

4.2.2 Second visit

The second meeting took place at the USC Speech and Hearing Research Center where three additional screenings were conducted. To assess motor imitation ability the MIS (Stone et al. 1997b) and the Unstructured Imitation Assessment (UIA) (Ingersoll and Lalonde 2010b) were administered. EVT2 (Williams 1997) was also administered to assess expressive vocabulary and word retrieval ability. At the conclusion of the second meeting, the video-recorded session was reviewed in order to compute a baseline measure of verbal utterances using the MLSUD. The MLSUD provides a total score for spoken meaningful language during the 1.0-1.5 hours assessment period. The MLSUD score is derived by assigning one point for each spoken morpheme divided by the total number of utterances in the session.

Fig. 3 Therapy room at the USC Speech and Hearing Research Center

5 Field study

Following the second meeting, each child received two 30-min sessions per week for a total of 6 weeks, or 12 total sessions of intervention with the robot. The room designated for the study sessions contained one child-sized table, two child-sized chairs, one or two adult-sized chairs, the robot and several hats and accessories for game play (Fig. 3). One or two researchers were also present during the interventions. The senior clinical instructor from the USC Speech and Hearing Research Center was regularly present (with few exceptions) to provide guidance and intervention expertise throughout each child's session. Additionally, a computer scientist was always present to provide continual monitoring and periodic operation of the robot during each session.

Recognizing that children with an ASD tend to experience high levels of stress in new situations and during certain activities, we also provided a second room where children could go to jump on a trampoline, roll on a large ball or read a book. The "break" room was used by participants on as as-needed basis at the clinician's discretion.

Six phases were initially identified for introducing the child to and engaging the child with the robot. Each phase was designed to address specific therapeutic goals, including increased speech and social skills, for interactions between study personnel, the child and the robot. However, upon completing sessions with the first two participants, the intervention strategy was revised to achieve improved study outcomes. A detailed discussion of the initial and revised intervention methodologies and the rationale for doing so, is described below.

5.1 Preliminary study protocol

The initial study procedure was developed over a series of weekly group meetings consisting of experts from computer science, autism diagnosis and treatment and speech therapy. Before introducing activities in the study procedure to promote specific speech and social skills and directly engage a child through interactive play, the group agreed that including exercises to facilitate the child's trust of the robot should precede any direct engagement. Therefore, the first two phases detailed in the study procedure describe objectives that encourage the child to manipulate the robot physically and explore controlling the robot's motions through teleoperation. This affords the child the opportunity to observe the robot's range of motion, kinematics and hear the sound(s) of the servo motors in a manner that gives the child control over the robot to the greatest extent possible. The objectives in Phases III–VI focus on promoting foundational skills required for communication and socialization through robot-assisted play. A detailed description of each initial phase is provided below:

Phase I The robot is situated in the room where the intervention will take place. For the first session (or at the clinician's discretion) the robot will be placed in stationary mode to allow the child the opportunity to physically explore the robot and its components before introducing movement.

1. On arriving and leaving each session, the subject will briefly make eye contact with the two researchers and the robot as part of his/her greeting. The child will be ready for Phase II after the child has been observed to:
2. Approach the robot.
3. Touch the robot.
4. Move the robot's arms.

Phase II The child, with guidance from one of the two researchers present (if required) will be given the opportunity to control the robot's arms and head with a remote control.

1. On arriving and leaving each session, the child will briefly make eye contact with the two researchers and the robot as part of his/her greeting.
2. During follow directions task, the child will point/operate remote/follow direction in order to lead the robot through an activity at least once during the session. We will be ready to approach Phase III after the child has been observed to:
3. Use the remote control to move the robot.

Phase III This phase introduces interactive play between the child and the robot where the robot plays "If you're happy and you know it" and "Wheels on the bus" while autonomously performing appropriate hand/arm motions. The researchers direct their attention toward robot, perform the appropriate hand motions and actively encourage the child to imitate as well.

1. On arriving and leaving each session, the child will briefly make eye contact with the researchers and the robot as part of his/her greeting.
2. During follow directions task, the child will point/operate remote/follow direction in order to lead the robot through an activity at least once during the session.
3. During song activities, the child will participate in fingerplay/gestures with the robot for 80 % of the opportunities presented (or at the clinician's discretion).
4. Once the robot has imitated child's movement, the child will continue to move/interact with the robot through X turns (X to be determined from performance on baseline/previous session). The child will be ready to approach Phase IV after the child has been observed to:
5. Respond to song with appropriate fingerplay/gesture.
6. Move in response to the child's prompt/action.

Phase IV The child, with guidance from one of the researchers (if required) will be given the opportunity to play imitation games with the robot. One-on-one games include just the child and the robot. The Pass the Pose game includes one of the researchers, the child and the robot.

1. On arriving and leaving each session, the child will briefly make eye contact with the researchers and the robot as part of his/her greeting.
2. During follow directions task, the child will point/operate remote/follow direction in order to lead the robot through an activity at least once during the session.
3. During song activity, the child will participate in fingerplay/gestures with the researchers for 80 % of the opportunities presented (or at the clinician's discretion).
4. Once the robot has imitated child's movement, the child will continue to move/interact with the robot through X

turns (X to be determined from performance on baseline/previous session).

5. When offered a choice of activities, the child will clearly make his/her performance known to others in session for 80 % of trials (or at the clinician's discretion).
6. Throughout the therapy session, the child will cooperate with a turn-taking task with the robot, caregiver and/or researchers through (2) turns (this number is expected to change as child progresses) each. We will be ready to approach Phase V after the child has been observed to:
7. Imitate the robot movements on 80 % of trials.
8. Imitate movements with another person in the intervention room.

Phase V The child will be given the opportunity to select from various modes of play with the robot.

1. On arriving and leaving each session, the child will briefly make eye contact with the researchers and the robot as part of his/her greeting.
2. During follow directions task, the child will point/operate remote/follow direction in order to lead the robot through an activity at least once during the session.
3. During song activity, the child will participate in finger-play/gestures with the robot for 80 % of opportunities presented.
4. Once the robot has imitated the child's movement, the child will continue to move/interact with the robot through X turns (X to be determined from performance on baseline/previous session).
5. When offered a choice of activities, the child will clearly make his/her performance known to others in session for 80 % of trials.
6. Throughout the therapy session, the child will cooperate with a turn-taking task with the robot, caregiver and/or researchers through (2) turns (this number is expected to change as child progresses) each.
7. During interactive games and songs with the robot, the child will participate in a structured reciprocal play routine for (2) minutes on (3) occasions (this number is expected to change as child progresses).

Phase VI If the child progresses through the previous five phases, s/he will be given the opportunity to select from the various available modes of play during Phase VI.

1. On arriving and leaving each session, the child will briefly make eye contact with the researchers and the robot as part of his/her greeting.
2. During follow directions task, the child will point/operate remote/follow direction in order to lead the robot through an activity at least once during the session.
3. During song activity, the child will participate in finger-play/gestures with the robot for 80 % of the opportunities presented.
4. Once the robot has imitated the child's movement, the child will continue to move/interact with the robot through X turns (X to be determined from performance on baseline/previous session).
5. When offered a choice of activities, the child will clearly make his/her performance known to others in session for 80 % of trials.
6. Throughout the therapy session, the child will cooperate with a turn-taking task with the robot, caregiver and/or researchers through (2) turns (this number is expected to change as child progresses) each.
7. When presented with communication opportunities by the researchers, the child will use gestures, vocalizations, or verbalizations for a variety of communicative intents on 80 % of opportunities presented.
8. When the child desires to initiate, change or discontinue activities within the last session, s/he will make eye contact as appropriate with the researchers or caregiver before communicating the message.

An important aspect of the study protocol is that each preceding phase provides the opportunity to build the basic skills necessary for succeeding in subsequent phases of the intervention. By first engendering trust and confidence, the exercises which follow can focus on scaffolding increasingly more challenging social and communication skills that rely on a well-established protocol with which the child has already become familiar.

5.2 Lessons learned

Early in the field study, study personnel made several significant observations that were both contributing to and limiting the success of the robot intervention.

First, the exercises described in Phases I and II for promoting trust by familiarizing the child with the robot's kinematics were accomplished in a much shorter period of time than anticipated. Some of the children participating in the study completely bypassed the first two phases and were ready to engage directly with the robot from the outset, while the remaining children progressed through the first two phases within 10 min of their first session. Second, the music and hand play featured in Phase III were very effective in capturing the attention of study participants. Almost all of the children in the study seemed to enjoy the music, directed their attention to the robot and engaged in at least some imitation.

Conversely, the "Imitate You, Imitate Me" and "Pass the Pose" games, as designed, were ineffective in practice. These two motor imitation games, which did not feature any kind of fun, positive sensory reinforcement, failed to maintain the child's attention or interest.

As a potential remedy for these observed limitations, we took inspiration from one of our nonverbal participants and redesigned the interactive games. From the outset of his participation in the study, this young child would look at the researchers and shake his head. When we began imitating him, he would laugh and seemed to derive great pleasure from this imitation game. We decided that trying this new movement and amusing response within the context of a new robot-facilitated game might promote longer interactions for this participant and, possibly, for others. To this end, a new game called "the Hat Game" was developed.

Ultimately, what we observed is that a key component driving productive, interactive game play with each child was the robot's engaging *reaction*. The increases in attention, eye contact, communication, speech and social interaction we observed immediately after the robot's amusing reaction were significant. Further, while other studies have shown that a robot can be used to effectively catalyze communication by encouraging attention, motor imitation and turn-taking, generalizing these behaviors to co-present others has remained challenging. In this study, generalizing robot–child games to child-co-present others was facilitated when caregivers and siblings in the room engaged in game play using the same amusing reaction performed by the robot. Additionally, although this study did not formally assess the generalizability of specific communication and social skills in other contexts, anecdotal reports from caregivers indicated that several participants engaged, and even initiated, the simple activities practiced during the intervention in their homes. A detailed description of our approach is included below.

5.3 Improved study protocol

As a result of these important observations, the original study procedure was modified. While some of the phases in the initial procedure remained unchanged, new software was written to deliver a new interactive game, the existing teleoperated imitation game was modified, less emphasis was placed on Phases I and II, and more emphasis was placed on child-directed play. This last modification effectively led to more child-initiated creative play, where speech introduced by study personnel to play robot-assisted games was often spontaneously generalized by the child to communicate a need or to engage in a different (but related) game created by the child. All modifications were made to introduce or emphasize positive sensory reinforcement received by the child when s/he successfully engages in interactive social play.

The description of Phases I and II include the child making eye contact with and greeting the robot and others in the room, touching and moving the robot's arms. These exercises remain unchanged from the original study procedure. However, instead of devoting one or two sessions to this activity these are all included as part of Phase I (and because the approach is based on scaffolding these skills, are also integrated into each subsequent session). Additionally, the improved approach features the robot responding by saying "hello" or "goodbye" and waving its hand as part of the greeting and parting process. Requiring that the child say "hello" and "goodbye" and/or wave to the robot and co-present others encourages social interaction, motor imitation and verbalization. Several participants enjoyed being able to "cause" the robot to respond in this way so much that they would practice this greeting repeatedly, seemingly enjoying their control over the robot and the predictability of its response.

The revised Phase II also features the addition of the simple game of peek-a-boo. Because this game requires some manipulation of the robot's arms and the game is at or below the developmental level of all the children participating in the study, this is an activity that the participants can master and enjoy early in their intervention; especially, when paired with the amusing response the child receives when the robot says "boo!". Moreover, the game generalizes to co-present others and establishes the practice of rehearsing games with the robot and then immediately generalizing them to others in the room.

Phase III remains unchanged except for the modification of the teleoperated game of imitation. While the original study procedure describes allowing the child to fully control the robot's motions with the remote control to familiarize her/him with the robot's kinematics and sounds and to encourage the child's trust of the robot, Phase III of the improved study procedure places more emphasis on reinforcing motor imitation by providing positive sensory feedback for the child's participation through one-to-many (robot-to-others) group imitation. This revised version of the teleoperated game of imitation promotes joint attention, motor imitation and turn-taking gameplay.

Revisions made to Phases IV, V and VI were significant. The "Imitate You, Imitate Me" and "Pass the Pose" games were excluded from the study procedure and replaced by a new, interactive game called "the Hat Game". The "Hat Game" begins when the child or co-present other places a hat on the robot. One of the researchers or child (with prompting or help, as needed) then asks, "CHARLIE, do you like your hat?", the robot responds by vigorously shaking its head and saying "Nooooo". This response typically elicits surprise and laughter from the child (and co-present others), thereby encouraging her/him to ask again or to try another hat.

To maintain the interest of the participants over the course of the study, the "Hat Game" was expanded to include an assortment of accessories including several kinds of hats, sunglasses, a scarf, a flower clip and wolf ears and an alternative, more positive response of, "Yeah!" was added. As

Fig. 4 A sample visual schedule used during sessions with the robot

the child becomes more able to ask the appropriate question, even if only in part, the game shifts from robot–child interactions to child–others interactions. When it is the child's turn, the child dons the hat, all others in the room direct their question to the child and wait for him/her to respond. When it is another person's turn, the child is encouraged to place the hat or other accessory on the researcher or caregiver, face that person and verbalize the question while pointing his/her finger at the person wearing the accessory.

The "Hat Game" is a core component of our intervention strategy and effectively encourages eye contact, directed attention, speech and social interaction by providing a positive sensory response to reinforce each child's efforts to communicate. Moreover, because the game is simple enough to be played in any setting, with almost any accessory, the increases in verbal utterances observed during game play in the clinic sometimes generalized to other settings. Anecdotally, several of the participants' caregivers reported that their child began initiating and/or playing the game at home and in the car during their participation in the study.

The last significant change to the original study procedure, was the addition of a visual schedule (Fig. 4) for several of the children participating in the study. By offering a visual menu of choices from which each participant could select an activity, the child was encouraged to actively participate in the direction of the session and to communicate his/her choice for the next exercise. While child-directed activity is part of Phase VI in the original study procedure, the improved version incorporates this aspect using the visual schedule from the very first few phases of the intervention.

6 Data analysis

Several participants that were recruited were ultimately excluded from study results for three primary reasons. First, to ensure reasonable consistency between participants, only three absences were allowable during the 6-week intervention. Participants 5 and 6 missed 4 and 5 sessions, respectively, due to family constraints and were consequently excluded from the study results. Second, an additional four participants completed the prescreening questionnaire and signed the informed consent, but did not continue in the study by attending the first diagnostic meeting at the Speech and Hearing Research Center. Several prospective participants were already receiving multiple therapies at the time of their invitation to the study and it is quite possible that their caregivers chose not to participate due to their already very demanding schedules (especially given that some of them also had young siblings.) Finally, Participant 1 completed the 12 sessions with the robot, but did not attend the final evaluation due to the child and his family traveling out of the country for 2 months before final session data could be collected. Since the final VABS-II and the final MLSUD were both missing, statistical analyses for this study did not included data collected from Participant 1. Eight participants met the study recruitment criteria, completed all evaluations and the entire 6-week robot intervention successfully. Three controls completed all requirements for participation as well.

Five evaluative tests were administered to each study group participant at the outset and at the completion of their participation in the study. These tests included the: (1) VABS-II, (2) MLSUD, (3) MIS, (4) UIA and (5) EVT2. A within-subject t test statistic was performed on scores obtained from composite scales and subscales and a discussion of results collected is provided in Sect. 6.1. Additional t test statistics were conducted to compare study group data with control group scores to assess relative improvements in five VABS-II communication and socialization scales.

6.1 Within-group data analyses

Raw data from the five evaluative tests administered during the field study are presented in Tables 1, 2, 3, 4, and 5. Pre- and post-communication and pre- and post-socialization

Table 1 Parent/caregiver reported Vineland Adaptive Behavior Scale results

Part	Pre-test comm.	Post-test comm.	Percent change (%)	Pre-test social	Post-test social	Percent change (%)
2	49	59	+20.41	61	61	0.0
3	42	40	−4.76	51	49	−3.92
4	79	79	0.0	72	79	+9.72
7	83	91	+9.64	83	95	+14.46
9	69	72	+4.35	74	97	+31.08
10	54	67	+24.07	57	86	+50.88
11	57	61	+7.02	61	81	+32.79
12	85	83	−2.35	74	81	+9.46
AVG	64.7	69.0	+6.65	66.6	78.6	+18.02

Participant identifier and beginning age are followed by (from left to right): (a) Composite communication pre-test score, (b) Composite communication post-test score, (c) Change in composite communication scores, (d) Composite social interaction pre-test score, (e) Composite social interaction post-test score, (f) Change in composite socialization scores

Table 2 Parent/caregiver reported VABS-II combined receptive and expressive communication v-scale results

Part.	Pre-test v-scale	Post-test v-scale	Percent change in pre/post-test V-scores (%)	Percent change in composite comm scores (%)
2	12	14	+16.67	+17.0
3	8	7	−12.5	−4.76
4	21	21	0.00	0.0
7	16	19	+18.75	+9.64
9	20	20	0.00	+4.35
10	13	17	+30.77	+24.07
11	14	16	+14.29	+7.02
12	18	20	+11.11	−2.35
AVG.	15.25	16.75	+9.84	+6.65

Participant identifier and beginning age are followed by (from left to right): (a) Total receptive and expressive communication v-scale pre-test score, (b) Total receptive and expressive communication v-scale post-test score, (c) Change in receptive and expressive communication pre-test/post-test v-scale scores and (d) Change in composite communication scores

composite scores are included in Tables 1 and 2. MLSUD scores recorded during the preliminary session, two intermediate sessions and the final session with participants are presented in Table 3 and results from the Unstructured Imitation Assessment—including data for social, requesting and joint attention behavior—are included in Table 4. Finally, results from the Expressive Vocabulary Test and Motor Imitation Scale are included in Table 5.

6.2 Between-group data analyses

Pre- and post-test VABS-II v-scale scores were collected from the control group and are included in Table 6. Five independent samples analyses were conducted, but only data collected for the four scales shown to be significant or with trends are included in the table. VABS-II v-scale scores collected for the study group are also represented in the table for comparison.

6.3 *T* test statistics

A paired-samples t test statistic was computed to evaluate the significance of increases reported for each of the nine, within-group evaluative categories: (1) VABS-II Communication Domain, (2) VABS-II Socialization Domain, (3) VABS-II Receptive and Expressive Communication v-scale scores, (4) MLSUD, (5) UIA social imitation, (6) UIA requesting, (7) UIA joint attention, (8) MIS and (9) EVT2. Six out of the nine t tests performed demonstrate that field tests resulted in statistically significant within-group increases. Additionally, an independent samples t test statistic between study and control groups was computed for five of the VABS-II scales including: (1) Expressive Language, (2) Receptive Language, (3) Interpersonal Skills, (4) Play and Leisure and, (5) Coping Skills. Results showed significant group differences or trends in four out of the five scales evaluated.

7 Results

A detailed discussion of data collected for study and control groups in addition to results obtained from within-subject and between-subject t tests performed are presented, by category, below.

7.1 Within-group results

VABS-II The VABS-II questionnaire was completed by a caregiver for each child in the study group. Composite com-

Table 3 Mean length of spontaneous utterance determination (MLSUD) results

Part.	Age	Prelim	Mid1	Mid2	Final	Greatest change (%)	Pre/post change (%)
2	4:10	0.37	1.17	1.52	1.08	+305.3	+188.0
3	4:1	0.0	0.0	0.0	0.0	0.0	0.0
4	3:5	2.76	2.90	3.26	3.77	+36.6	+36.6
7	3:10	1.75	3.24	3.98	1.86	+127.4	+6.3
9	6:2	2.38	3.13	3.21	3.18	+34.9	+33.6
10	6:5	1.55	1.72	1.97	1.83	+27.1	+18.1
11	6:5	1.41	2.42	2.95	2.5	+109.2	+77.3
12	3:3	1.80	2.92	2.25	2.05	+62.2	+13.9
AVG.	4:10	1.50	2.19	2.39	2.03	+59.3	+35.3

Participant identifier and beginning age are followed by (from left to right): (a) Preliminary evaluation MLSUD, (b) Session 4 MLSUD, (c) Session 8 MLSUD, (d) Final evaluation MLSUD, (e) Greatest increase in MLSUD, (f) Change in pre-test/post-test MLSUD

Mid1 and Mid2 changes include some rote speech learned for gameplay. Spontaneous speech increases are evaluated using only the Pre/Post percentage changes listed above

Table 4 Unstructured imitation assessment results

Part.	UIA-Soc(A) (%)	UIA-Soc(B) (%)	Change (%)	UIA-Req(A) (%)	UIA-Req(B) (%)	Change (%)	UIA-JA(A) (%)	UIA-JA(B) (%)	Change (%)
2	28.0	83.0	+196.4	20.8	45.8	+120.2	23.3	33.3	+42.9
3	5.0	22.2	+344.0	4.2	25.0	+495.2	3.3	23.3	+606.1
4	50.0	66.7	+33.4	33.3	33.3	0.0	30.0	46.7	+55.7
7	100	100	0.0	20.8	45.8	+120.2	13.3	30.0	+125.6
9	39.0	56.0	+43.6	20.8	71.0	+241.3	13.3	66.7	+401.5
10	33.3	28.0	−15.9	17.0	41.7	+145.3	20.0	40.0	+100.0
11	28.0	78.0	+178.6	29.0	45.8	+57.9	40.0	33.3	−16.8
12	61.1	100.0	+63.7	20.8	91.7	+340.9	43.3	90.0	+107.9
AVG	42.4	66.7	+57.3	20.8	50.0	+140.4	22.9	45.4	+98.3

Participant identifier followed by the percentage of successful imitations for: (1) Preliminary UIA-social, (2) Final UIA-social and, (3) Change between preliminary and final social scores; (4) Preliminary UIA-requesting, (5) Final UIA-requesting and, (6) Change between preliminary and final requesting scores; (7) Preliminary UIA-joint attention, (8) Final UIA-joint attention and, (9) Change between preliminary and final joint attention scores

Table 5 Expressive vocabulary test (EVT) and motor imitation scale (MIS) results

Participant	EVT(A)	EVT(B)	CHANGE	MIS(A)	MIS(B)	CHANGE
2	42	73	+73.8	44.0	44.0	0.0
3	42	42	0.0	0.0	0.0	0.0
4	114	121	+6.1	81.3	90.6	+11.4
7	97	114	+17.5	84.0	94.0	+11.9
9	111	105	−5.4	100.0	100.0	0.0
10	87	69	−20.7	100.0	88.0	−12.0
11	94	94	0.0	94.0	100.0	+6.4
12	114	119	4.4	87.5	96.9	+10.7
AVERAGE	87.6	92.1	+5.1	73.8	76.7	+3.9

Participant identifier followed by: (a) Preliminary EVT(A), (b) Final EVT(B), (c) Percent change in EVT score, (d) Preliminary MIS(A), (e) Final MIS(B), (f) Percent change in MIS score

Table 6 Parent/caregiver reported VABS-II between-subject V-scale results

Part	Receptive pre-	Receptive post-	Change	Inter pre-	Inter pre-	Change	Play pre-	Play post-	Change	Coping pre-	Coping post-	Change
2	7	8	+14.3	8	8	0	8	7	−12.5	8	9	+12.5
3	3	3	0	5	5	0	6	6	0	8	7	−12.5
4	9	9	0	8	11	+37.5	9	10	+11.1	13	13	0
7	7	9	+28.6	11	13	+18.2	8	11	+37.5	17	19	+11.8
9	12	11	−8.3	9	15	+66.7	9	13	+44.4	13	16	+23.1
10	7	10	+42.9	6	10	+66.7	7	11	+57.1	9	17	+88.9
11	7	9	+28.6	8	8	0	7	12	+71.4	9	15	+66.7
12	9	10	+11.1	9	10	+11.1	9	13	+44.4	13	12	−7.7
AVG	7.625	8.625	+13.1	8	10	+25.0	7.87	10.37	+31.8	11.25	13.5	+20.0
C1	5	6	+20.0	7	7	0	7	8	+14.3	9	11	+22.2
C2	5	5	0	7	8	+14.3	7	6	−14.3	9	10	+11.1
C3	3	3	0	5	4	−20.0	7	7	0	8	8	0
AVG	4.333	4.667	+7.7	6.333	6.333	0	7	7	0	8.667	9.667	+11.5

Participant identifiers are followed by (from left to right): (a) Receptive Language pre-test, (b) Receptive Language post-test, (c) Change Receptive Language pre- to post-test, (d) Interpersonal pre-test, (e) Interpersonal post-test, (f) Change in Interpersonal pre- to post-test, (g) Play pre-test, (h) Play post-test, (i) Change in Play pre- to post-test, (j)Coping Skills pre-test, (k) Coping Skills post-test, and (l) Change in Coping Skills pre- to post-test v-scale scores munication and socialization scores were extracted from each pre- and post-test VABS-II questionnaire to present within-study data for expressive and receptive communication and social skills results.

Field study results show a mean increase of 6.65 % in overall communication and a 18.02 % mean increase in social interaction skills as reported by caregivers on the Vineland-II Parent/Caregiver Rating Form (Table 1). The range of increases in composite communication scores collected was approximately 29 % and the range of composite socialization scores was approximately 55 %. Five out of the eight participants demonstrated an increase in the Communication Domain and six out of eight showed an increase in the Socialization Domain.

One possible cause for the lack of improvement in the VABS-II Communication Domain recorded for a few of the participants is that the communication composite scores include 3 areas of communication: (1) receptive, (2) expressive and (3) written. Lack of increases or relative decreases in the written portion of the communication composite may have also diminished actual gains in expressive and receptive communication. Because our study does not address written communication, results and analysis for the VABS-II v-scale scores representing the Receptive and Expressive Communication Domain scores are provided for comparison with the Communication Composite Scores (Table 2). As illustrated, average overall increases in composite communication scores were approximately 6.65 % while mean increases in the combined receptive and expressive communication domains were slightly higher at approximately 9.84 %.

Results from the within-subject t test show a mean improvement rating of $t(7) = +2.14$, where $p < 0.0699$ for the Communication Domain and a mean improvement rating of $t(7) = +3.06$, where $p < 0.0184$ for the Socialization Domain. These results demonstrate that while increases recorded for the VABS-II Communication Composite Domains do not indicate that the null hypothesis can probabilistically be ruled out, mean increases observed for the Socialization Domain are statistically significant. Additionally, t test results for the Receptive and Expressive Domain scores show a statistically significant mean gain of $t(7) = +2.51$, where $p < 0.0404$.

MLSUD A total of four separate MLSUD measures were collected; scores were computed at the first meeting with the child, during sessions 4 and 8 and at the very end of the 6-week intervention period (Table 3). Only MLSUD scores from the preliminary and final evaluation sessions are included and used to perform comparative assessments and t statistic analysis of increased speech since intermediate sessions featured a significant amount of rote (non-spontaneous) speech that was used to engage in games with the robot and with co-present others.

A mean increase of 35.3 % in spontaneous speech as calculated by the Mean Length Spontaneous Utterance Determination measure (Table 3) was observed after a 6-week intervention with the robot. The range of preliminary MLSUD scores was from 0.00 to 2.76 and final MLSUD scores ranged from 0.00 to 3.77. All participants but one demonstrated an increase in MLSUD score. The single participant whose MLSUD did not improve throughout the course of the intervention was nonverbal at the start of his 6-week intervention and did not demonstrate any acquired speech during the final evaluation. Although his scores for the VABS-II and EVT2 also did not improve, the child's scores did improve in each of the three categories included in the UIA evaluation. This might suggest that while the participant did not make any significant gains in communication or speech, some fundamental imitation and attention skills - key precursors to communication and speech - were improved throughout the course of the study.

The paired-samples t test showed a mean improvement of $t(7) = +3.56$, where $p < 0.0092$ for the MLSUD category. These t test statistics again confirmed that the increases observed in the collected raw data are significant for the population of children included in our study.

UIA The UIA is one of the two assessments we used to evaluate growth in imitation ability in three major areas: (1) social interaction, (2) requesting and (3) joint attention and consists of a total of 24 measures. Pecentages included in Table 4 reflect the ratio of the total number of points received by each child to the total number of points achievable on the assessment. Points provided for each measure reflect the number of examples successfully demonstrated by the child during the 60-min initial and final evaluations. The maximum number of points assigned for each of the 24 measures is "3", giving a total of 72 points possible. Analysis of the raw UIA data show that mean increases for each of the three UIA areas - social imitation, requesting and joint attention—were 57.3, 140.4 and 98.3 %, respectively.

Within-subject t test results show a significant increase in the social interaction domain ($t(7) = +3.02$, where $p < 0.0193$), the requesting domain ($t(7) = +3.79$, where $p < 0.0068$) and the joint attention domain ($t(7) = +3.23$, where $p < 0.0145$). Interestingly, these test results reveal that the most significant improvements were achieved in the requesting domain—a primary focus area of the intervention provided. The t test statistic confirms that the observed increases in UIA pre- and post-test scores for all three UIA domains were each statistically significant.

MIS The MIS is the only instrument used to assess motor imitation ability and was administered at the beginning of the study period and at the end, for each participant. The MIS evaluates motor imitation using a total of 16 measures which assess a child's ability to imitate meaningful and nonmeaningful actions and body movements. Percentages included in Table 5 reflect the ratio of the total number of points received by each child to the total number of points achievable. The maximum number of points assigned for each of the 16 measures is "2", giving a total of 32 points possible for the MIS. A "2" indicates a passing score, a "1" indicates an emerging skill and "0" indicates a failure for that particular skill.

Mean increases in scores for the MIS were marginal, improving by only 3.9 %. Since activities included in the intervention did not primarily focus on improving motor imitation skills as much as social interaction, communication and speech, these results were not unexpected. T test results showed a mean improvement of $t(7) = +1.07$, where $p < 0.3182$, which were not statistically significant for this population in this study.

EVT2 The Expressive Vocabulary test was administered primarily to provide additional data regarding each child's progress in word acquisition and retrieval as a secondary measure of communication skill. The EVT2 consists of a total of 190 items and is typically administered by a speech-language pathologist, psychologist or early childhood specialist. The test features a series of pictures depicting objects, people and situations and is administered by the examinee who prompts the child to name or describe a picture after being provided a stimulus question. Scores reported in Table 5 are based on the extracted Growth Scale Values (GSVs), a metric used for easily measuring each child's progress over time.

Improvements in the EVT2 were also slight and were not found to be statistically significant for this population of children using this intervention ($t(7) = +0.86$, where $p < 0.4166$). Again, given that the our primary study objectives did not include targeting the acquisition of new vocabulary or improving word retrieval ability, these statistical conclusions are not surprising. Instead, they do provide additional information about other mitigating factors that may contribute to an individual's performance on other tests administered during the study. For example, a child with apraxia of speech may improve marginally on the VABS-II Communication Domain, the MLSUD and the EVT2, but show greater improvements in the VABS-II social domain, the MIS and the UIA. Given that a few of the study participants had other known medical diagnoses, including some that limited the physical ability of the child, a future study with a larger study population and further analysis would shed light on the possible effects these complicating factors.

7.2 Between-group results

An independent samples *t* test between study and control groups showed significance in three of the five and a trend in one of the five VABS scales examined. Between-group increases in Receptive Language and Play and Leisure scales were found to be significant and scores collected for Interpersonal and Coping Skills indicated the presence of significant trends. Expressive language scores reported for both groups, however, did not show any inter-group significance. Participant pre- and post- v-scale scores for each of the four scales showing significance or trends and comparative gains are included in Table 6.

VABS-II receptive language Results from a between-subject t test showed significance of $t(8) = +1.97$, where $p < 0.0421$ for the Receptive Language Domain. Further, Cohen's effect size value (r=0.57) suggested a moderate to high practical significance. Mean increases in pre- and post-test receptive language v-scores show an overall average improvement of approximately 13.1 % for the study group and 7.7 % for the control group. Out of eight study group participants, overall increases were reported for five participants, no improvement in receptive language was reported for two study group participants and scores for one participant indicated a slight decrease. Receptive language scores for one out of the three controls showed a mean increase in receptive language, with the remaining two controls showing no change.

VABS-II play and leisure Scores recorded for each group demonstrate between-group significance in the Play and Leisure scale as well ($t(8) = +1.90$, where $p < 0.0469$) and were also found to have moderate practical significance with a Cohen's effect size value of r=0.56. Study group participants averaged an increase of approximately 31.8 % compared to a mean 0 % change in the study group over the 6-week study study period. Overall increases were reported for six study group participants, with one participant showing no change and one participant showing a slight decrease. Play and Leisure v-scale scores collected for the control group indicated a slight increase for one control, no change for one control and a slight decrease for the remaining control.

VABS-II interpersonal ability Trends were indicated on the VABS-II Interpersonal Scale subdomains ($t(8) = +1.72$ where $p < 0.0618$) with the study group increasing approximately 25 % and the control group showing no overall increase. Further, Cohen's effect size for this subdomain showed a moderate practical significance (r=0.52). Pre- and post- v-scale scores reported for five out of eight study group participants indicated an increase, with the remaining three study participants showing no improvement. Control group pre- and post-test v-scale scores indicate a slight improvement for one control, no change for one control and a slight decrease for the remaining control.

VABS-II coping skills Between group analyses also indicated a trend in the Coping Skills v-scale scores ($t(8) = +1.67$

where $p < 0.0667$) and a moderate effect size of r=0.51, with the study group showing overall improvement of approximately 20 % and the control group averaging approximately an 11.5 % increase. Overall improvement was reported for five study group participants, no change for one participant, and a slight decrease for the two remaining participants. Conversely, two out of three controls showed a mean increase in Coping with the remaining control showing no improvement overall.

8 Discussion

Results from within-group analyses indicated significant improvements in spontaneous utterances, social interaction, joint attention and requesting behaviors while overall communication scores, vocabulary and motor imitation did not reveal statistically significant increases. Scores increases for spontaneous utterances suggest an improvement in the study group's level of language proficiency. Additionally, statistically significant social interaction increases, reported by the UIA and the Socialization Domain of the VABS-II, revealed an improved ability to initiate or engage in social exchanges with others. Finally, joint attention and requesting also increased during the course of this study. Collectively, these improvements are consistent with the specific tasks included in the study activities. For instance, the component parts of the protocol require the child to (1) request that the robot play a song (requesting/joint attention), (2) imitate the robot during music/hand play and ask the robot if it liked the hat (joint attention) and, (3) turn to co-present others and initiate a social exchange during the "Hat Game"(social interaction).

Between-group analyses indicated meaningful differences between improvements reported for the study group and those reported for a control group that did not receive the additional robot-assisted intervention. These differences were most significant in two VABS-II domains, Receptive Language and Play and Leisure. Two other subscales, Interpersonal and Coping Skills, also revealed trends towards significance while differences in Expressive Language did not show any significance between groups. Receptive language reflects the child's ability to comprehend language and plays a critical role in effective social interaction. This result may have been attributed to the repeated practice of social exchanges but warrants further exploration of the current study's approach to determine which components, and to what extent, may have contributed to the promotion of receptive language ability. Resulting differences in Play and Leisure were more expected. These between-group differences are consistent with the play-based methodology employed in this study and may help to further inform how play in robot-assisted interventions can be used promote other key communication and socialization skills.

Field trials documented in this study describe the failures and subsequent successes of a novel robot-assisted intervention. Our initial approach focused on an autonomous robot interaction that targeted well-studied difficulties in imitation and turn-taking in populations of children with ASD. In practice, the initial interactive games were ineffective with this study group for a number of practical reasons. First, many of the participants in this study were very young and were moderately affected by ASD. Indeed, this study was designed for this population, since it was expected that they would potentially benefit the most from improved imitation and turn-taking skills. However, remaining in a stationary position long enough to engage the robot, maintaining an engagement for more than a brief period of time, and recognizing the causality of their own actions and the robot's movements were difficult for the children in this study to achieve. Fundamental changes were then made to the study protocol.

Taking inspiration from one of the study's nonverbal participants, who repeatedly initiated and seemed to enjoy a head-shaking imitation game, we redefined the robot's role and incorporated elements of this playful imitation into the existing clinical methodology. Additionally, we invited family to participate along with the child during each session. The resulting robot-assisted intervention yielded statistically significant improvements in key indicators of communication, interpersonal skills, play and socialization. Additionally, anecdotal evidence suggested that intervention-introduced activities were generalized, with participants initiating the same games in other contexts such as the car and in the home. The findings of this study detail a paradigm shift from many existing robot-assisted interventions. First, this study provides further evidence that a kinematically simple, low-cost robot can be used to effectively engage children with ASD and promote significant improvement of critical communication and socialization skills. Second, robot-assisted interventions benefit considerably from incorporating activities already enjoyed by the children they will serve as a central part of a scaffolded, clinical approach. Finally, inviting caregivers and siblings to participate in sessions with the robot and study personnel may have a significant impact on the generalization of activities from the clinic to other contexts. Collectively, these findings contribute important insights to inform the approach of future robot-assisted interventions.

9 Study limitations

This study measured increases in speech and socialization resulting from a new robot-assisted intervention paradigm with a small group of children with ASD. Given the limited sample size of both the study and control groups, results obtained are regarded as indicative of potential benefits from this approach and not necessarily representative of the target population. Although t tests were used to assess the significance of results for this study, we acknowledge that children who participated in this study may not be representative of the larger population. We conducted the Shapiro-Wilk test of normality for study and control groups. While normality test results indicated that the ASD study sample was normally distributed, some skewness in the control group Communication Domain post-test and Socialization Domain pre-test resulted, indicating that between-group results should be considered with caution. Future studies will include more rigorous statistical tests, and include larger sample sizes, to further validate the intervention methodology described herein. Finally, in order to assess the sustainability of improvements resulting from the robot-assisted intervention, a future study will retest study participants after an interval to test the retention of increased speech and socialization proficiency.

10 Conclusions

Our research resulted in the development, field testing and delivery of a new, interactive robot-assisted intervention for young children with ASD. This new technique resulted in relative increases in speech, communication and social interaction by scaffolding basic proficiencies through a series interactive games.

Our technique employs a robot as a confederate for children to practice fun, social exchanges before generalizing each task to co-present others. Two fundamental ideas make this approach valuable for use with other robots. First, multiple interaction modalities are presented to engender trust and to facilitate curiosity and engagement. These modalities include manual manipulation of the robot, teleoperation and autonomous imitation and turn-taking games. Second, context-specific imitation and turn-taking games are introduced through facilitator-robot demonstration and ultimately progress from child-robot exchanges to child–others exchanges as the child gains proficiency. Defining features of the interactive games include (1) a child-initiated verbal prompt, possibly delivered in the form of a question (and with help, as needed), (2) a contingent robot reaction, (3) an accompanying motor imitation component, and, (4) generalized child-robot exchanges to child–others interactions soon after an introductory period.

Robot tasks explored in this study can be automated with reasonable modifications. Software needed to implement the following two activities will leverage existing functionality to deliver a more autonomous robot. First, salutations can be automated using the previously-implemented face and hand detection to track the movement of the hand to trigger a response. Second, RFID tags can be added to the robot's head and hands and to each accessory to allow the robot to autonomously identify which accessory was being placed on its head and respond appropriately. This same modality can be implemented for initiating each coordinated movement with music. For instance, placing a an RFID-tagged card with a picture of a bus in the robot's hand might trigger the launch of the "Wheels on the Bus" activity.

This basic interaction framework can be easily replicated or adapted for use with other new and existing robots, using a wide range of social contexts to target specific therapeutic objectives. For example, to promote language, the interaction model might be based on a specific play scenario with a set of associated vocabulary as potential verbal prompts. The robot reaction may be contingent on the appropriate use of a verbal prompt and include a verbal response and accompanying movement that can be imitated to provide a nonverbal reinforcement of the feedback. Finally, after several practice trials with the robot, the social scenario is reenacted by others in the room and the child is encouraged to play the game with a human partner.

There is growing interest in increasing the autonomy of robots in robot-assisted autism interventions but questions still remain concerning which robot roles are appropriate for ultimately increasing human-to-human socialization and communication. Further, challenges in defining a set of useful, widely applicable autonomous robot tasks persist. For instance, rewarding a child appropriately during social interactions and after child-initiated speech is of critical importance to reinforcing socialization and communication. A human expert can more easily discern between a child who is having trouble articulating a request from a child who may not be making the effort and can determine when it is appropriate to trigger the robot's positive response. In this regard, human-in-the-loop and teleoperated approaches not only emphasize the importance of leveraging expertise from therapists, teachers and caregivers for developing robots in this domain but also provide insight as to what cues might be necessary for the robot to collect in order for it to be able to perform optimally during autonomous tasks.

While foundational skills such as motor imitation and turn-taking are key to facilitating communication and socialization, these skills must be practiced in a context that is conducive to interactions which are fun and generalize beyond child-robot exchanges. Our study demonstrates that robot-enabled interventions designed for children with an ASD can effectively promote the generalization of child-robot activities to child–other interactions.

References

Bekele, E. T., Lahiri, U., Swanson, A. R., Crittendon, J. A., Warren, Z. E., & Sarkar, N. (2013). A step towards developing adaptive robot-mediated intervention architecture (aria) for children with autism. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *21*(2), 289–299.

Boccanfuso, L., & O'Kane, J. M. (2011). CHARLIE: An adaptive robot design with hand and face tracking for use in autism therapy. *International Journal of Social Robotics*, pp. 1–11.

Brown, R. (1973). *A first language: The early stages*. Cambridge: Harvard University Press.

Cabibihan, J. J., Javed, H., Ang, M, Jr., & Aljunied, M. A. (2013). Why robots? a survey on the roles and benefits of social robots for the therapy of children with autism. *International Journal of Social Robotics*, *5*(4), 593–618.

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

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.

Feil-Seifer, D., & Matarić, M. (2008) Robot-assisted therapy for children with autism spectrum disorders. In *Proceedings of the 7th international conference on Interaction design and children*, IDC '08 (pp. 49–52). New York, NY, USA: ACM.

Grecez, J., Kaszubski, E., Atrash, A., & Matarić, M. J. (2014) Graded cueing feedback in robot-mediated imitation practice for children with autism spectrum disorders. In *The 23rd IEEE symposium on robot and human interaction communication (RO-MAN '14)* (pp. 561–566).

Gurley, J. (2011). Expressive vocabulary test. In S. Goldstein & J. Naglieri (Eds.), *Encyclopedia of child behavior and development* (2nd ed., pp. 622–623). New York: Springer.

Huskens, B., Verschuur, R., Gillesen, J., Didden, R., & Barakova, E. (2013). Promoting question-asking in school-aged children with autism spectrum disorders: Effectiveness of a robot intervention compared to a human-trainer intervention. *Developmental neurorehabilitation*, *16*(5), 345–356.

Ingersoll, B., & Lalonde, K. (2010a). The impact of object and gesture imitation training on language use in children with autism. *Journal of Speech, Language, and Hearing Research*, *53*, 1040–1051.

Ingersoll, B. R., & Lalonde, K. (2010b). The impact of object and gesture imitation training on language use in children with autism. *Journal of Speech, Language, and Hearing Research*, *53*, 1040–1051.

Kim, E. S., Paul, R., Shic, F., & Scassellati, B. (2012). Bridging the research gap: Making hri useful to individuals with autism. *Journal of Human-Robot Interaction*, *1*(1).

Kim, E. S., Berkovits, L. D., Bernier, E. P., Leyzberg, D., Shic, F., Paul, R., et al. (2013). Social robots as embedded reinforcers of social behavior in children with autism. *Journal of Autism and Developmental Disorders*, *43*(5), 1038–1049.

Lord, C., Rutter, M., Goode, S., Heemsbergen, J., Jordan, H., Mawhood, L., et al. (1989). Austism diagnostic observation schedule: A standardized observation of communicative and social behavior. *Journal of Autism and Developmental Disorders*, *19*(2), 185–212.

Project, A. (2011). Aurora project. http://www.aurora-project.com/.

Robins, B., & Dautenhahn, K. (2014). Tactile interactions with a humanoid robot: Novel play scenario implementations with children with autism. *International Journal of Social Robotics*, *6*(3), 397–415.

Scassellati, B., Admoni, H., & Mataríc, M. (2012). Robots for use in autism research. *Annual Review of Biomedical Engineering*, *14*, 275–294.

Services, W. P. (2013). Social communication questionnaire.

Stone, W. L., Ousley, O. Y., & Littleford, C. D. (1997a). Motor imitation in young children with autism: What's the object? *Journal of Abnormal Child Psychology*, *25*(6), 475–485.

Stone, W. L., Ousley, O. Y., & Littleford, C. D. (1997b). Motor imitation in young children with autism: What's the object? *Journal of Abnormal Child Psychology*, *25*(6), 475–485.

Vertue, F. (2007). *Vineland adaptive behavior scales*. New York: Springer.

Warren, Z., Zheng, Z., Das, S., Young, E., Swanson, A., Weitlauf, A., & Sarkar, N. (2014). Brief report: Development of a robotic intervention platform for young children with asd. *Journal of Autism and Developmental Disorders*, pp. 1–7.

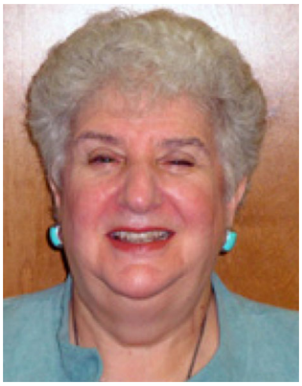
Welch, K. C., Lahiri, U., Warren, Z., & Sarkar, N. (2010). An approach to the design of socially acceptable robots for children with autism spectrum disorders. *International Journal of Social Robotics*, *2*, 391–403.

Williams, K. T. (1997). AGS Expressive Vocabulary Test, American Guidance Service. Expressive Vocabulary Test.

Yun, S. S., Choi, J., & Park, S. K. (2015). An interactive robot facilitating social skills for children. In *Proceedings of the tenth annual ACM/IEEE international conference on human-robot interaction extended abstracts, HRI'15 extended abstracts* (pp. 95–96).

Laura Boccanfuso is a Postdoctoral Associate in the Technology and Innovation Lab at Yale University. She received her Ph.D. in Computer Science from the University of South Carolina in 2014. Her research interests include socially assistive robots and the novel use of sensors and other technologies that advance the utility, safety and acceptance of robots working collaboratively with children with autism.

Sarah Scarborough is a Senior Clinical Instructor at the University of South Carolina Speech and Hearing Research Center. Sarah Scarborough received her M.A. from the University of Tennessee in Knoxville. She has worked in nursing homes, hospitals, outpatient clinics, schools and home health care. Since 2001, she has been a clinical instructor for the University of South Carolina. She has expertise in alternative/augmentative communication, adult neurogenics, autism and apraxia.



Ruth K. Abramson has a joint appointment in the Department of Neuropsychiatry and Behavioral Science in the School of Medicine and Communication Science and Disorders in the Arnold School of Public Health and is conducting research into genetic links to autism. Her research focus includes general autism topics, including language and speech in autistic children, the diagnosis of autism and the frequency of autism diagnoses.

Alicia V. Hall is a licensed Clinical Psychologist, recognized by the National Register of Health Services Providers in Psychology. She is a member of the clinical and research faculty at USC School of Medicine. Dr. Hall has worked with families with Autism Spectrum Disorders as a researcher and clinician since 2001.

Harry H. Wright is a pediatric and adolescent psychiatrist and Professor in the Department of Neuropsychiatry and Behavioral Science in the School of Medicine at the University of South Carolina. His research interests include psychiatric genetics, psychiatric epidemiology, multicultural issues, health services research, and infant and preschool mental health issues. His clinical practice is focused on young children and on children and adolescents with developmental disorders.

Jason M. O'Kane is an Associate Professor in Computer Science and Engineering and Director of the Center for Computational Robotics at the University of South Carolina. He holds the Ph.D. (2007) and M.S. (2005) degrees from the University of Illinois at Urbana-Champaign and the B.S. (2001) degree from Taylor University in Upland, Indiana, all in Computer Science. His research spans algorithmic robotics, planning under uncertainty, and computational geometry.