2018

Effectiveness of a robot social skills therapy for children with an autism spectrum disorder

Jenna Brooke Lebersfeld
University of Alabama at Birmingham

EFFECTIVENESS OF A ROBOT SOCIAL SKILLS THERAPY FOR CHILDREN WITH AN AUTISM SPECTRUM DISORDER

by

JENNA BROOKE LEBERSFELD

MARIA I. HOPKINS, CHAIR
FRED J. BIASINI
EDWIN W. COOK, III
SARAH E. O'KELLEY

A THESIS

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Master of Arts

BIRMINGHAM, ALABAMA

2018

EFFECTIVENESS OF A ROBOT SOCIAL SKILLS THERAPY FOR CHILDREN WITH AN AUTISM SPECTRUM DISORDER

JENNA BROOKE LEBERSFELD

MEDICAL/CLINICAL PSYCHOLOGY

ABSTRACT

This study examined the effectiveness of a robot-based social skills intervention for school-age children with Autism Spectrum Disorder (ASD) across a range of cognitive ability. Sixteen children were assigned to either the intervention group (*n* = 8), which received the robot intervention, or a control group (*n* = 8), which played non-emotion games with the robot. All participants reported high levels of enjoyment, motivation, and willingness to have future robot interaction sessions, indicating that the robot intervention is engaging and interesting for this population. However, there was no difference between groups on measures of emotion identification or generalized social skills. Exploratory analysis showed significant improvement within the intervention group on emotion identification accuracy and within both groups on parent-rated social skills. Overall, the results of this study indicate that robot-based interventions can be used as a method of teaching for children with ASD across a range of IQ levels and should be explored further. This intervention did not prove more effective for the intervention group compared with control participants. This may be due to the dose or content of the intervention, the outcome measures chosen, or the power of the study. Future research in this area should address study limitations and investigate which participant characteristics are predictive of social skills improvement following intervention.

Keywords: autism spectrum disorder, social robot, social skills, emotion recognition, robot-based intervention

ACKNOWLEDGMENTS

Many individuals contributed to the completion of this project. My research mentors, Drs. Maria Hopkins and Fred Biasini, along with the entire Social Technology for Autism Research (STAR) Lab, helped me take my initial interest in this area and develop it into a realizable project. I would like to thank Drs. Sarah O'Kelley and Edwin Cook for their continued support throughout the duration of this undertaking. Many thanks to Carl Stevens for his essential role in programming and building the robot. This research could not have been conducted without the involvement of the parents, children, and teachers who participated in the study and helped bring this project to fruition. This study was funded in part by the UAB Civitan International Research Center and the UAB Civitan/Sparks Clinics LEND program.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ABA	Applied Behavior Analysis
ADOS-2	Autism Diagnostic Observation Schedule, Second Edition
ANCOVA	analysis of covariance
ANOVA	analysis of variance
APA	American Psychiatric Association
ASD	Autism Spectrum Disorder
CDC	Centers for Disease Control and Prevention
CHANGE	Recognizing Change in Expression
DSM-5	Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition
ID	Intellectual Disability
IQ	Intelligence Quotient
INFER	Inferring Emotion
KBIT-2	Kaufman Brief Intelligence Test, Second Edition
M	Mean
MATCH-D	Expression Matching: Schematic Drawings
MATCH-F	Expression Matching: Face Photos
MDN	Median
MODEL	Modeling Emotional Expressions
PPVT-4	Peabody Picture Vocabulary Test, Fourth Edition
PRT	Pivotal Response Treatment

R-BESK	Robot-Based Emotion Skills Knowledge
SAM	Socially Animated Machine
SD	Standard Deviation
SE	Standard Error
SRS-2	Social Responsiveness Scale, Second Edition
SSIS	Social Skills Improvement System
STAR	Social Technology for Autism Research
T1	baseline assessment
T2	post-intervention assessment
TD	Typically Developing
UAB	University of Alabama at Birmingham

INTRODUCTION

Autism Spectrum Disorder

Autism Spectrum Disorder (ASD) is a neurodevelopmental disability involving deficits in social communication and restricted, repetitive patterns of behavior (American Psychiatric Association, 2013), and children with ASD often have additional difficulties with attention and motivation (L. K. Koegel, Singh, & Koegel, 2010; Narzisi, Muratori, Calderoni, Fabbro, & Urgesi, 2013). The most recent prevalence estimate from the Centers for Disease Control and Prevention (CDC, 2014) indicates that one in 68 children has ASD.

Outcomes

Outcomes for individuals with ASD vary greatly. Some are able to enter the workforce and live independently while others must receive constant support, living with family members or in group homes (Chamak & Bonniau, 2016). About half of those with ASD have below average IQ, with 23% in the borderline range (IQ = 71-85) and 31% qualifying for an intellectual disability (ID; IQ ≤ 70; CDC, 2014). The lifetime cost of care for an individual with ASD and ID is about $2.4 million, costing the United States an average of $236 billion per year (Buescher, Cidav, Knapp, & Mandell, 2014). Unfortunately, the burden of caring for this population often falls on families, resulting in stress, anxiety, and depression in parents of these individuals (Bitsika & Sharpley, 2004).

Additionally, cognitive ability and social skills are predictors of quality of life in individuals with ASD (Burgess & Gutstein, 2007; Eaves & Ho, 2008; McGovern & Sigman, 2005).

Emotion Identification

Children with ASD often struggle with emotion identification and recognition and have difficulty recognizing and interpreting facial expressions (Chasson & Jarosiewicz, 2014). These skills are foundational to understanding the mental and emotional states of others and are crucial for emotion regulation and social competence (Back, Ropar, & Mitchell, 2007; Grossmann & Johnson, 2007; Izard et al., 2001). These deficits do not improve naturally with development; in fact, as the social world becomes more complicated and children become more aware of their deficits, social skills difficulties are often exacerbated (White, Keonig, & Scahill, 2007).

Facial expressions contribute greatly to social development and the ability to interact with others, and the consequences of social skills deficits in this population can be tremendous (Amenta, Ferrari, & Balconi, 2014; White et al., 2007). Poor emotion recognition and regulation in school-age children has been linked to difficulties relating to peers, increasing the risk of peer rejection and social isolation (White et al., 2007). Poor social skills can contribute to underachievement in school and employment as well as the development of mood disorders such as anxiety disorder and depression (Howlin, 2000).

Given the high prevalence of ASD and its significant societal and familial impact, effective early interventions targeting the core deficits of this disorder are crucial. A multitude of therapies have been developed to address social skills deficits in individuals with ASD with varying degrees of success. The most common interventions are outlined briefly, below.

Social Skills Interventions

Traditional Therapy and Social Stories™

Social skills therapies have traditionally been led by a clinician and involve either directly teaching social skills to the individual with ASD or training peers or family members to interact with the child more appropriately (Bohlander, Orlich, & Varley, 2012). Therapy may be individual or in a group setting and usually takes place in a clinic or school (Bohlander et al., 2012). Group social skills training is facilitated by a teacher or therapist who directly teaches lessons to a group of children with ASD. These lessons are often manualized and focus on improving specific social skills such as initiation, joint attention, reading emotions, maintaining eye contact, or understanding humor (White et al., 2007). This type of training allows children with ASD to practice interacting with other children in a group setting and aims to promote generalization to situations outside the therapy session (Barry et al., 2003).

Social Stories™ (Gray & Garand, 1993) are stories that are written in first-person perspective by a parent or teacher to convey a specific social skill or behavior to the child with ASD, such as appropriate greetings or how to join an activity appropriately (Bohlander et al., 2012; Karkhaneh et al., 2010). Picture books can also be used in a similar way to teach social skills to this population (Bohlander et al., 2012).

These social skills interventions have promise for use with individuals with average or above average cognitive skills but are not appropriate for those with borderline or below average IQ. Group-based therapies have not been shown to be effective for individuals with lower cognitive ability, and the Social Stories™ intervention is only suitable for individuals with at least a first grade reading level (Karkhaneh et al., 2010).

Applied Behavior Analysis (ABA)

One of the most common types of therapy used for individuals with ASD and ID is Applied Behavior Analysis (ABA). ABA is based on the behavioral learning theory that individuals are more likely to repeat behaviors that are reinforced and less likely to repeat behaviors that are punished or ignored. Interventions based on principles of ABA have been shown to improve functional communication and socialization skills in children with ASD and decrease problematic behaviors such as aggression (Hanley, Iwata, & Thompson, 2001; Lovaas, 1987; Virués-Ortega, 2010). However, ABA requires intensive one-on-one intervention for 20 to 40 hours per week, and skills learned through ABA often fail to generalize to alternate settings (Virués-Ortega, 2010).

Peer and Family Training

Peer mentoring is another common technique used to teach social skills to individuals with ASD and below average IQ. During peer-training therapies, typically developing peers or siblings are taught how to interact more appropriately with children with ASD. For example, the peers or siblings may be taught how to “be a good buddy” by playing and talking with children with ASD, even if they do not respond (Bohlander et al., 2012; Laushey & Heflin, 2000). This therapy has proven to be one of the most effective strategies for increasing early social communication skills in preschool children with ASD (Bohlander et al., 2012).

Pivotal Response Treatment (PRT) is an intervention for children with ASD that involves teaching social reinforcement skills to parents and peers. Participants learn to interact with children with ASD, including increasing appropriate attention and following the lead of the child with ASD (R. L. Koegel & Koegel, 2012). Parents and peers are also taught to model appropriate social behaviors such as turn taking and conversation. This type of intervention requires extensive training, but it has shown to be effective in increasing social behaviors in children with ASD such as initiation and joint-attention (Bohlander et al., 2012; R. L. Koegel & Koegel, 2012).

In these types of therapies, the training is given directly to peers, siblings, and parents, not the child with ASD. Therefore, it is difficult to determine if behavior change occurs for the child with ASD, or if the apparent improvement in social skills is solely a function of interactions initiated by the peers or family members (Kroeger, Schultz, & Newsom, 2007). Though age-appropriate language is not necessary for this type of intervention, there is little research on those with profound ASD and a lack of evidence regarding generalization and maintenance of learned skills (Banda, Hart, & Liu-Gitz, 2010; Flynn & Healy, 2012).

Video Modeling

Video modeling social skills interventions involve children with ASD watching video demonstrations of themselves or others using appropriate social skills. These children are then taught to imitate the skills modeled in the video. This type of intervention can be used with individuals with ASD and below average cognitive skills, and social skills learned through this technique have been shown to generalize and maintain over time. However, individuals must be able to attend to the video and have developed the ability to imitate, a skill often delayed or limited in children with ASD, to benefit from this type of intervention (Bellini & Akullian, 2007; Flynn & Healy, 2012).

Technology-Based Interventions

Traditional social skills therapies for children with ASD require a human therapist to administer the intervention and involves interpersonal interactions between the therapist and patient. This requires extensive time and resources that are in high demand and are often not available. In addition, even in a therapeutic setting, human social interaction can be confusing and anxiety-provoking for children with ASD (Bellini, 2006). People with ASD feel more comfortable in predictable environments and often prefer interacting with technology over people (Putnam & Chong, 2008). Given the predictability of the environment and an inherent enjoyment of interacting with technology, technology-based therapies are more engaging and motivating for these individuals (Duquette, Michaud, & Mercier, 2008). Therefore technology-based interventions may be more effective than traditional social skills therapies.

Capitalizing on the inherent interest in technology in children with ASD, a variety of technology-based tools have been explored for use with children with ASD, including computer- and robot-based therapies. The advantage of computer-based avatars compared to human interactions is that avatars simulate social interactions in a predictable, non-threatening environment (Goldsmith & LeBlanc, 2004). FaceSay (Hopkins et al., 2011) is a computer-administered, avatar-based therapy that was created to teach social skills to individuals with ASD. It uses realistic, interactive avatars on a computer screen to promote social skills development, specifically facial- and emotion-recognition skills. FaceSay was successful in improving facial recognition, emotion recognition, and social interactions for individuals with ASD, and skills generalized to a naturalistic playground setting. However, individuals did not improve on all measures of social skills, and those with below average cognitive skills were less successful than those with average or above average IQ (Hopkins et al., 2011).

Based on the successes and limitations of computer-based social skills interventions, it was hypothesized that using a social robot as an interactive partner as opposed to a two-dimensional avatar would result in even greater improvement and generalization of social skills in this population. Previous robots have had success in teaching specific social skills to individuals with ASD and average IQ, including imitation and joint attention, and have been investigated as supportive tools for social skills therapy in this population (Boccanfuso & O'Kane, 2011; Boccanfuso et al., 2016; Kim et al., 2013).

Similar to avatar-based methods, the use of robots allows for a predictable and reliable environment (Robins, Dickerson, Stribling, & Dautenhahn, 2004). Children are most interested in robots that move predictably and contingently in response to the child's actions, and robots are perceived as being predictable and non-threatening by children with ASD since they can perform repetitive tasks with consistency and reliability (Boccanfuso & O'Kane, 2011; Feil-Seifer, 2008). Social robots allow children with ASD to acquire increasingly complex information about social interactions in an environment that is predictable and feels safe, whereas the social behavior of human interaction partners seems unpredictable (Robins & Dautenhahn, 2014; Robins et al., 2004). Because of this increased interest and predictability, people with ASD have heightened motivation to interact with a robot, which may further increase the effectiveness of the intervention. Given that the robot interaction takes place in the three-dimensional space, as opposed to on a two-dimensional screen in the computer-based therapies, the skills learned during robot therapy may be more likely to generalize to settings outside of the clinic or classroom setting.

Researchers investigating robots as tools for therapy in autism have reported increases in engagement, attention, and novel social behaviors when robots are involved in the social interaction (Diehl, Schmitt, Villano, & Crowell, 2012; Ricks & Colton, 2010; Scassellati, 2007). Children with ASD have increased interest in robots and attend more to robots compared to people and computers (Kim et al., 2013). In a study of 4- to 12-year-old children with ASD, children directed more vocalizations and attention to an adult confederate when interacting with a dinosaur robot compared to another adult or a computer game (Kim et al., 2013). In addition, children with ASD are more interested in interacting with robots compared to other non-robotic toys, as measured by level of interaction, gaze, and touch (Dautenhahn & Billard, 2002).

Administering social skills training through the use of social robots is promising, as social robots “occupy a special niche between inanimate toys and animate social beings” (Scassellati, Admoni, & Mataric, 2012). Social robots can engage children with ASD in social communication without eliciting the confusion and distress they feel while interacting with other people (Koch, 2017). Social robots may have the capacity to elicit desirable social behaviors in children with ASD that are not seen in other interaction settings (Ricks & Colton, 2010).

In addition, there is evidence that the use of child-preferred, intrinsic reinforcers leads to improvements in social engagement (Paul, 2008). Embedding social interaction into the delivery of a child’s preferred reinforcer elicits greater social initiation, increased orientation to face an interaction partner, and more positive affect (Kim et al., 2013). Children with ASD have been shown to use social behaviors to interact with robots, including initiating social interaction, turn-taking, imitation, joint attention, attention span, eye contact, and child-led speech (Dautenhahn & Werry, 2004; Duquette et al., 2008; Feil-Seifer & Matari’c, 2009; Kozima, Michalowski, & Nakagawa, 2009; Kozima, Nakagawa, & Yasuda, 2005; Pioggia et al., 2005; Robins, Dautenhahn, Te Boekhorst, & Billard, 2005). Studies have further demonstrated that children with ASD will interact with a parent, caregiver, or another human more often while engaged with a robot partner, for instance, by expressing excitement to a robot, and then turning to express this excitement to a parent (Feil-Seifer & Matari’c, 2009; Kozima et al., 2009; Robins et al., 2005). Interaction with a social robot elicits speech directed socially toward an adult confederate, in addition to the robot itself, and best facilitates interaction with another person, not just social interaction with objects (Kim et al., 2013).

Children with ASD often have difficulty generalizing skills learned during therapy sessions to more naturalistic settings outside of the clinic or classroom (Smith & Sung, 2014). Therefore, the therapy setting should mimic the outside setting as closely as possible. Given that robots exist and interact in three-dimensional space, they can generate an environment that resembles face-to-face social interaction more so than any other technology-based techniques (Koch, 2017). In addition, the use of humanoid robots, as opposed to non-humanoid robots, are necessary to teach social skills that will ultimately generalize to social interactions outside the therapeutic setting (Dautenhahn & Werry, 2004; Lord & Bishop, 2010; Ricks & Colton, 2010; Scassellati et al., 2012).

The social robots that have been created to date for use with individuals with ASD vary greatly in appearance, purpose, and effectiveness. The creation of robots for use with children with ASD began in the robotics and engineering fields, and initially the robots created for use with children with ASD looked much like traditional, non-humanoid robots (Ricks & Colton, 2010). Published literature in these areas focused on detailing aspects of programming and design. These studies set the stage for the development of different types of robots, ranging from non-humanoid to animal-like to android designs, but they did not adequately evaluate the clinical applications of these robots.

More recently, studies have been emerging addressing the effectiveness of these robots in improving social skills for children with ASD. A comprehensive review of studies addressing the clinical utility of social robots for children with ASD was published in 2016 (Pennisi et al.). A selection of these robots is summarized below.

Social Robot Interventions

Non-humanoid robots continue to be used for children with ASD, including a ball with touch sensors called the TouchPad (Lee, Takehashi, Nagai, & Obinata, 2012) and a robotic arm with an attached basketball hoop (Conn, Liu, Sarkar, Stone, & Warren, 2008). Other robots resemble cartoons. Keepon is a small, yellow robot that was developed to orient attention by moving its head and to express emotion, such as excitement, by bobbing up and down or rocking side to side (Kozima et al., 2009). Animal-like robots have also been created. Wada et al. developed a seal pet robot as an assistive tool in rehabilitation and robot-assisted activity (Wada, Shibata, Saito, Sakamoto, & Tanie, 2005). Pleo is a dinosaur robot with a multitude of sensors (Kim et al., 2013). Sony Aibo ERS-7 is a robotic dog that can recognize voice commands (François, Powell, & Dautenhahn, 2009). Humanoid robots look more like people, but remain predictable and repeatable. The NAO robot (Aldebaran Robotics Company) is made of plastic, has two arms and legs, and has luminescent eyes (Shamsuddin et al., 2012). KASPAR resembles a small boy with dark hair. It has a minimally expressive face made of a silicon-rubber mask (Wainer, Dautenhahn, Robins, & Amirabdollahian, 2014). Robota is a robot that looks like a small doll (Dautenhahn & Billard, 2002).

According to the Pennisi et al. (2016) paper, in all studies reviewed, children with ASD directed attention to the social robots. Children with ASD also directed social behaviors to robots more than a human interaction partner. Imitation improved when a robot was present in the interaction session, but joint attention did not. The use of a social robot can also increase social behaviors of the child with ASD toward a third interaction partner. Repetitive behaviors decreased and communication increased when interacting with a robot compared to a human partner. The authors of this review caution that given the small number of subjects included in these studies, these results should not considered conclusive but should instead be used as hypotheses for future studies (Pennisi et al., 2016).

Only two previous studies investigated the use of social robots for individuals with ASD and impaired IQ. One study found that children with ASD and moderately impaired IQ (40-54) were receptive to interacting with social robots (Shamsuddin et al., 2012). Another study comparing individuals with ASD and three levels of IQ (moderately impaired, mildly impaired, and borderline) did not find differences between these groups in their robot interactions (Shamsuddin, Yussof, Mohamed, Hanapiah, & Ismail, 2013). Individuals with ASD and below average cognitive skills are often excluded from participating in therapeutic studies due to difficulty with recruitment and retention, and researchers have called for more research on effective interventions for this population (Kim et al., 2013).

Regarding methods, most of the published social robot literature uses a case study or case series design. For those studies which include a larger sample size, there is often only a single group comprised of all children with ASD with no comparison group. For the few studies which have utilized a control group, this control group was often typically developing (TD) individuals. Many outcomes measured in previous research involved socially-relevant aspects of the child-robot interaction (e.g., gaze, imitation, communication), but outcome measures were often subjective, qualitative descriptions of participant behavior during the robot interaction as opposed to utilizing psychometrically sound, objective outcome measures. Generalizability of these skills to contexts outside the therapeutic setting was rarely measured (Begum, Serna, & Yanco, 2016). In addition, the effect of individual participant characteristics including age and cognitive ability has not been examined (Pennisi et al., 2016).

THE ROBOT SAM

Past research shows that children with ASD prefer to interact with cartoon-like or animalistic robots; however, these robots typically offer only a limited range of facial expressions that do not generalize well to the human face (Ricks & Colton, 2010). Anthropomorphic robots have the greatest potential for generalization and have been effective in teaching imitation skills to children with ASD (Ricks & Colton, 2010). However, individuals with ASD exhibit anxiety when interacting with humans, particularly human faces, and often prefer interacting with animals instead (Kleinhans et al., 2010; Prothmann, Ettrich, & Prothmann, 2009). To bridge this gap, this research team created a robot, Socially Animated Machine (SAM).

Appearance

The team hypothesized that modeling SAM after a monkey would have the highest likelihood of generalization compared to any other animal. The monkey design would avoid producing anxiety associated with human faces, but the facial expressions would closely resemble those created with human faces (Ricks & Colton, 2010). SAM has an approachable, animal-like appearance with features of a human face (see Figure 1). SAM's eyes are cartoon-like and displayed on a screen, allowing for the inclusion of details not seen in other robots, such as eye lids and skin folds, which are involved in emotional facial expressions. SAM has eyebrows and a mouth which also move, and SAM displays six facial expressions: happiness, sadness, anger, fear, surprise, and disgust (see Figure 2). Typically developing children were able to identify SAM’s emotions with high accuracy, at or above the level of other social robots (Koch et al., 2017).

Figure 1. SAM
(Socially Animated Machine).

People respond more positively to robots that have appropriate facial expressions and emotions coordinated with speech during an interaction (Leite, Pereira, Martinho, & Paiva, 2008; Rabbitt, Kazdin, & Scassellati, 2015); therefore, SAM was designed to talk and coordinate speech with emotional facial expressions. SAM can move its head and arms which enables the production of gestures including head nod, head shake, wave, and point. The legs do not have movement capabilities and, therefore, remain stationary.

Much of the previous robot literature describes robots that must be controlled by an outside observer, such as a researcher, therapist, or teacher. SAM, on the other hand, is completely autonomous, and the intervention runs without researcher input. Implementing SAM in a clinical or therapeutic setting would allow therapists and teachers to serve more individuals using fewer resources. For additional details regarding the design and development of the SAM robot monkey, please refer to previously published work on this topic (Koch et al., 2017).

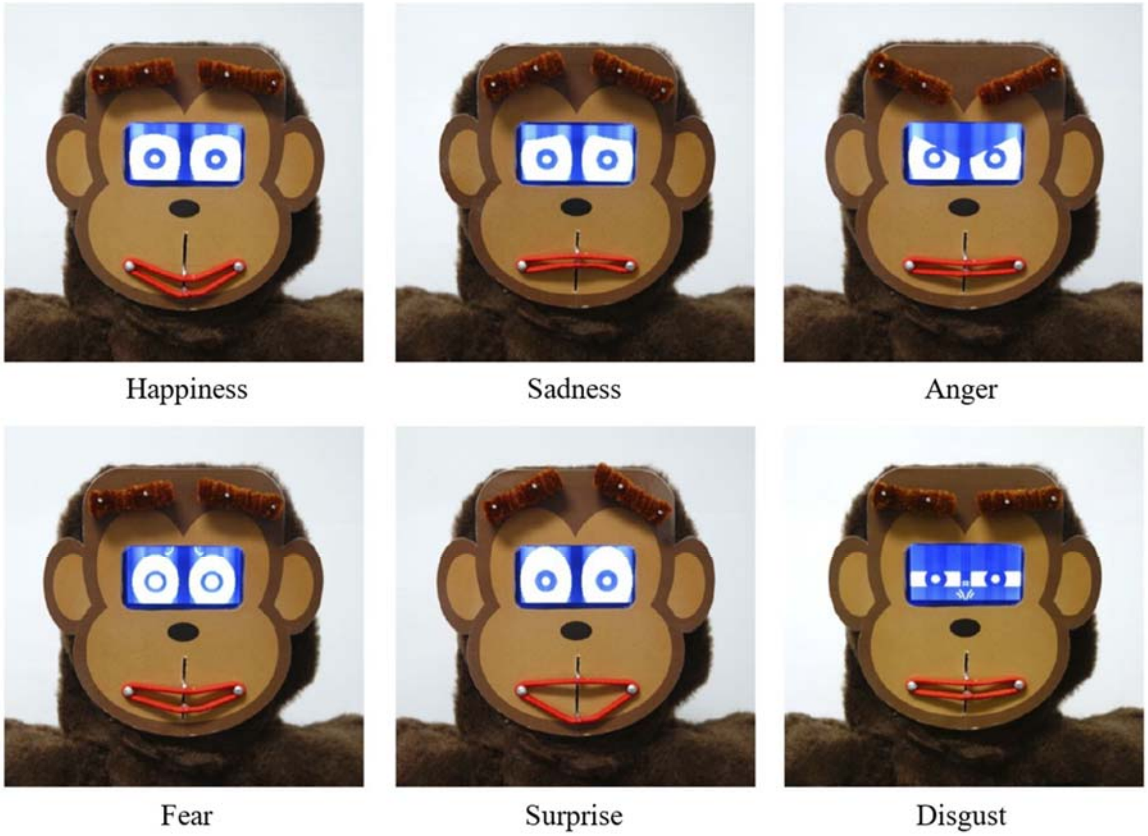


Figure 2. SAM's emotional facial expressions.
Note: Adapted from "A Feasibility Study Evaluating the Emotionally Expressive Robot SAM" by S. A. Koch, C. E. Stevens, C. D. Clesi, J. B. Lebersfeld, A. G. Sellers, M. E. McNew, F. J. Biasini and M. I. Hopkins, 2017, *International Journal of Social Robotics, 9(4)*, p. 601-613. Copyright 2017 by Springer International Publishing. Adapted with permission.

Intervention Program

During the intervention sessions, the child is seated at a table facing the SAM robot. SAM sits behind a touchscreen tablet and talks directly to the child while presenting response options on the tablet. Participants indicate their choices by touching the word or picture on the tablet. Notably, participation in this intervention does not require verbal expressive language.

The SAM robot intervention consists of five mini-games that target emotion recognition and identification:

- *Modeling Emotional Expressions (MODEL);*
- *Expression Matching: Schematic Drawings (MATCH-D);*
- *Expression Matching: Face Photos (MATCH-F);*
- *Recognizing Change in Expression (CHANGE);*
- *Inferring Emotion (INFER).*

During *MODEL*, SAM introduces the target emotional facial expressions by modeling and labeling each emotion. During *MATCH-D* trials, SAM models and labels an emotion, and the child is asked to choose the matching emotional facial expression from an array of six schematic drawings on the tablet. These schematic drawings of emotional facial expressions are black and white and were included to highlight the essential features of the human face in a simplified manner (MacDonald, Kirkpatrick, & Sullivan, 1996). The *MATCH-F* game mimics *MATCH-D*, but the presented stimuli are black and white photos of human emotional facial expressions from Ekman & Friesen (1975).

The *CHANGE* and *INFER* games involve a more complex understanding of emotions. During *CHANGE*, SAM introduces a brief social scenario designed to elicit a specific emotion. These social scenarios were developed by the STAR Lab and were validated in typically developing adults and children with high accuracy (Koch, 2017). SAM models and labels the target emotion following the vignette, and after a five-second delay, SAM models and labels another emotion. The child is asked to choose whether the emotion has changed. In the *INFER* game, SAM recites the same scenarios and does not model or label the corresponding emotion. The child is then asked to identify the correct emotion that applies to the story. These games are introduced through a series of eight, weekly lessons which last about 15 to 25 minutes each. See Table 1 for session content.

Table 1

Intervention Session Content

Session	**Content**	**Estimated Length**
T1	MODEL (6 trials) MATCH-D (12 trials) MATCH-F (12 trials)	15 minutes
Session 1	MODEL (6 trials) MATCH-D (24 trials)	15 minutes
Session 2	MODEL (6 trials) MATCH-F (24 trials)	15 minutes
Session 3	MODEL (6 trials) CHANGE (12 trials) INFER (12 trials)	25 minutes
Session 4	MODEL (6 trials) CHANGE (24 trials)	25 minutes
Session 5	MODEL (6 trials) INFER (24 trials)	25 minutes
Session 6	MODEL (6 trials) MATCH-D (6 trials) MATCH-F (6 trials) CHANGE (6 trials) INFER (6 trials)	20 minutes
T2	MODEL (6 trials) MATCH-D (12 trials) MATCH-F (12 trials)	15 minutes

Note: Table adapted with permission from Koch (2017).

Previous research with SAM showed that the robot was engaging for children with ASD and average IQ, and that these children were happy and comfortable when interacting with the robot (Koch et al., 2017). Participants spent a greater percentage of time looking at the face during an interaction with SAM compared to a similar human interaction (Koch et al., 2017). The intervention group improved more than the control group on an emotion matching task, but there was a ceiling effect (Koch, 2017). Given that the Koch (2017) study was conducted with individuals with ASD and average to above average IQ, it was hypothesized that this intervention may be even more effective for individuals with ASD and below average cognitive ability.

Control Robot Interaction

During the control interaction sessions, the child interacts with SAM in a similar manner as the intervention session, but the content of the sessions differs. During the control sessions, the child is presented with an array of six dance moves to choose from. After the child makes the selection on the tablet, SAM performs the dance move. The control interaction sessions were designed to be equally as engaging as the intervention sessions but free of emotion-based content. Therefore, SAM does not speak or move any facial features during these sessions. The amount of time spent playing this game is comparable to the intervention sessions.

OBJECTIVES

Given the success of the SAM robot intervention for children with ASD and average to above average IQ, the primary goal of this research study was to determine whether the SAM robot intervention is effective for children with ASD across all levels of cognitive ability. Specific aims and hypotheses are outlined below.

Specific Aims

Aim 1

To investigate whether the SAM robot is enjoyable and motivating for children with ASD. It was hypothesized that families would attend all eight sessions of the study and that parent and child ratings of the SAM interactions would indicate high levels of motivation and enjoyment for all participants, regardless of group membership (Koch, 2017).

Aim 2

To determine if the SAM robot intervention results in improvements in emotion recognition immediately following the intervention. It was expected that emotion-identification accuracy would increase immediately following the intervention as measured by the tablet for those in the intervention group compared to controls (Hopkins et al., 2011; Koch, 2017).

Aim 3

To determine whether children with ASD who participated in the SAM intervention show improvements in generalized social-emotional skills. It was predicted that the intervention group would show greater improvements in global social-emotional skills as rated by parents and teachers. Additionally, teacher-rated, classroom-based social skills were predicted to improve more for those in the intervention group (Hopkins et al., 2011).

Aim 4

To discover the effectiveness of this intervention in improving facial recognition in children with ASD immediately following completion of the program. It was hypothesized that the intervention group would make greater gains on a clinician-administered facial recognition measure than the control group (Hopkins et al., 2011).

METHODS

Participants

Sample Size Considerations

A priori power analysis using the software G*Power 3 (Faul, Erdfelder, Lang, & Buchner, 2007) was performed to determine the sample size required to find meaningful group differences. Effect size for a parent-rated social skills questionnaire from previous work with an avatar-based social skills intervention for this population (Hopkins et al., 2011) was large (Cohen's $d = 1.01$). Power analysis was conducted for fixed effects, main effects, and interactions of analysis of covariance (ANCOVA) using the aforementioned effect size. It was determined that a total sample size of 34 participants (17 per group) would be adequate to achieve power of .80. Due to recruitment difficulties, it was not feasible to enroll this number of participants. Given that the majority of literature published in this area has utilized a case study or case series design with fewer than eight participants (Pennisi et al., 2016), including only 16 participants for the current study was still expected to make a meaningful contribution to this field of research.

Participant and Parent Characteristics

Participant characteristics are reported in Table 2. Seventeen children were initially enrolled in the study. One child withdrew prior to study completion due to anxiety experienced during interactions with the robot. All other enrolled participants in both groups were able to complete the robot study. Sixteen children with ASD ages five to 13 years completed the study. There were 15 male participants. Twelve participants were White (and not of Hispanic origin) and four were Black (and not of Hispanic origin).

Participants were recruited from local and community centers in Birmingham, Alabama that serve families of children with ASD. Children were required to be between the ages of four and 14 and have a diagnosis of ASD to participate. Eligibility was confirmed by researcher administration of the *Autism Diagnostic Observation Schedule, Second Edition* (*ADOS-2*; Lord et al., 2012). The experimenter also completed a rating of diagnostic certainty on a scale from one to five, with four or five indicating a high level of confidence in the ASD diagnosis. All participants met classification for *autism* (n = 14) or *autism spectrum* (n = 2) on the *ADOS-2* and received a diagnostic certainty rating of four (n = 5) or five (n = 11). Additionally, participants with uncorrected vision or hearing problems were excluded from the study due to possible interference with viewing and hearing the stimuli required for study participation. Cognitive ability was assessed using the *Kaufman Brief Intelligence Test, Second Edition* (*KBIT-2*; Kaufman & Kaufman, 1990), and ability level ranged from severely impaired to high average. Receptive language skills were measured with the *Peabody Picture Vocabulary Test, Fourth Edition* (*PPVT-4*; Dunn & Dunn, 2007) and ranged from severely impaired to above average. For parent characteristics, refer to Table 3. Parent education was varied and ranged from high school to advanced degree. Participants were assigned to either the intervention (n = 8) or control group (n = 8). Fisher's exact tests and independent samples *t*-tests were conducted to evaluate the presence of between-group differences. There were no differences.

Table 2

Participant Characteristics

	Intervention (n = 8)	Control (n = 8)	Overall (N = 16)	Group Differences
	M (SD)	*M (SD)*	*M (SD)*	*p*
Age (years)	9.74 (1.65)	10.02 (2.77)	9.88 (2.21)	.809
ADOS-2 Comparison Score	8.0 (1.2)	7.1 (1.8)	7.6 (1.6)	.273
KBIT-2				
Verbal IQ	71.3 (24.9)	73.6 (14.1)	72.4 (19.6)	.818
Nonverbal IQ	86.8 (30.4)	93.1 (13.1)	89.9 (22.9)	.595
IQ Composite	77.5 (28.8)	80.9 (13.5)	79.2 (21.8)	.771
PPVT-4				
Standard Score	76.1 (29.8)	87.8 (23.3)	81.9 (26.5)	.400
	Freq. (%)	Freq. (%)	Freq. (%)	
Gender				1.00
Male	7 (87.5)	8 (100.0)	15 (93.8)	
Female	1 (12.5)	0 (0.0)	1 (6.2)	
Ethnicity				1.00
White	6 (75.0)	6 (75.0)	12 (75.0)	
Black	2 (25.0)	2 (25.0)	4 (25.0)	
ADOS-2 Classification				.467
Autism	8 (100.0)	6 (75.0)	14 (87.5)	
Autism Spectrum	0 (0.0)	2 (25.0)	2 (12.5)	

Table 3

Parent Characteristics

	Intervention (n = 8)	Control (n = 8)	Overall (N = 16)	Group Differences
	M (SD)	*M (SD)*	*M (SD)*	*p*
Education (years)				
Mother	15.6 (2.8)	16.0 (2.1)	15.8 (2.4)	.769
Father	15.4 (2.8)	15.3 (2.7)	15.4 (2.6)	.951
Household Income (thousands)	122.7 (88.0)	159.4 (77.7)	140.3 (82.8)	.407

Procedure

The protocol for this study was approved by the Institutional Review Board at the University of Alabama at Birmingham (UAB). Participants were recruited from organizations in Birmingham, Alabama that work with children with ASD (e.g., Autism Society of Alabama), flyers were posted at ASD-focused events and community centers including local parks and libraries, and 146 sites were contacted directly, including schools, therapy centers, and community centers. The primary investigator conducted a brief phone screener with 39 interested families to determine initial eligibility. The research protocol was explained in detail over the phone and any questions were addressed. Thirty-four participants were eligible to participate based on the phone screener, and 23 attended the initial session. Five families did not participate due to ineligibility or living a substantial distance from the study site. Others were lost to follow-up. At the first session, parents provided written consent, and participant assent was obtained for individuals ages seven and older who were cognitively able to assent.

The assent form also served as a visual schedule to set expectations for the participants. Following eligibility testing at T1, two participants did not qualify for an ASD diagnosis based on the *ADOS-2*. One participant discontinued due to anxiety regarding the robot interaction. Twenty participants were enrolled, and 16 participants have completed the study.

Testing procedures at the initial meeting (T1) included completion of demographic, descriptive, and outcome measures by the parent and child followed by the T1 SAM robot session. The T1 session lasted between 90 to 120 minutes. Teacher questionnaires were provided to the parent in a stamped, addressed envelope, and parents gave these questionnaires to their child's teacher to complete and mail back. For children who were in a homeschool program or who were not currently attending school (e.g., over the summer), an appropriate alternative individual was identified to complete the questionnaires (e.g., therapist, camp counselor).

Participants were then assigned to group based on an algorithm designed to match groups on IQ while maintaining an equal number of participants per group (Appendix A). The first four participants were randomized to group. All future participants were assigned based on overall IQ composite score. The IQ mean of each group was calculated. When a new participant was tested, IQ was measured, and this IQ was determined to be (a) higher than both group IQ means, (b) lower than both group IQ means, or (c) between group IQ means. If participant IQ was (a) higher than mean IQ for both groups, the participant was placed in whichever group had the *lower* IQ mean. If the new participant IQ was (b) lower than both group IQ means, this participant was assigned to the group with *higher* group IQ. This would decrease the difference between group means. If participant IQ was (c) between the mean IQ levels for both groups, the participant was assigned to the group with smaller *n* to attempt to equate sample size per group. If there were equal *n* per group, the participant group assignment was randomized.

Parents, teachers, and participants were blind to group membership. Children in both groups interacted with the SAM robot eight times, once per week. Both groups played emotion games with SAM at T1 to evaluate emotion identification accuracy at baseline. The content of the next six sessions differed between groups, with the intervention group completing the SAM emotion intervention and the control group playing a dance game with SAM. During the robot sessions, children were provided with visual supports (e.g., timer, picture symbols) as well as reinforcers (e.g., breaks, preferred food items) when needed to decrease anxiety and maintain engagement. At T2, parents and participants completed outcome measures and teacher questionnaires were provided to the parents. Participants in both groups played emotion games with SAM for the final session. See Table 4 for timing of specific measures. Most participants were able to complete the study within eight weeks. Due to scheduling difficulties, two participants had three-month delays between the T1 session and regularly-scheduled weekly sessions. Other participants encountered routine absences due to family schedules, holiday breaks, or illness, and completed the study in an average of 9.5 weeks.

Following study completion, parents were debriefed and informed of group assignment, and control participants were given the option to complete the SAM intervention. Enrollment in the SAM intervention for control participants following study completion was completely optional, and additional data were not collected following study completion for these participants.

Table 4

Length and Duration of Measures

Measure	T1	T2
Parent		
Consent Procedures	20 min	
Demographics Questionnaire	10 min	
SRS-2	20 min	20 min
Enjoyment Questionnaire		5 min
Total Duration	50 min	25 min
Child		
Assent (if able)	10 min	
ADOS-2	60 min	
KBIT-2	20 min	
PPVT-4	15 min	
Benton	10 min	10 min
SAM Robot Interaction	15 min	15 min
Enjoyment Questionnaire (if able)		10 min
Total Duration	130 min	35 min
Teacher		
SRS-2	20 min	20 min
SSIS	25 min	25 min
Total Duration (minutes)	65 min	65 min

Measures

Demographic and Descriptive Measures

Demographics questionnaire. Parents completed a brief questionnaire about child and family characteristics. Child information included age, gender, and ethnic and racial identity. Family information included household area (e.g., rural, urban, etc.,), household income, parental education, marital status, and employment status.

Autism Diagnostic Observation Schedule, Second Edition (*ADOS-2*; Lord et al., 2012). The *ADOS-2* is an instrument used to aid in the diagnosis of ASD. It is a semi-structured observation interaction session with an administrator and the child, and it is appropriate for use within the designated age range. Based on observed behaviors, the measure evaluates skills in the areas of social communication and restricted and repetitive behaviors and yields a diagnostic classification of autism, autism spectrum disorder, or non-spectrum. A comparison score indicates the level of ASD symptoms and ranges from one to 10, with higher scores indicating greater levels of ASD symptomatology.

Diagnostic certainty rating. The diagnostic certainty rating is a brief rating of the researcher's certainty of the ASD diagnostic accuracy on a scale from one to five. A rating of four or five indicates moderate to high levels of certainty.

Kaufman Brief Intelligence Test, Second Edition (*KBIT-2*; Kaufman & Kaufman, 2004). The *KBIT-2* assesses general cognitive abilities and was used as an estimate of intelligence. The *KBIT-2* generates verbal, nonverbal, and composite domain scores along with verbal and nonverbal age equivalents. Domain and composite scores are standard scores with a mean of 100 and a standard deviation of 15, with higher scores indicating more developed skills. This measure has strong psychometric characteristics, with a mean internal consistency reliability coefficient of .92 and a mean test–retest reliability coefficient of .90.

Peabody Picture Vocabulary Test, Fourth Edition (*PPVT-4*; Dunn & Dunn, 2007). The *PPVT-4* is a measure of receptive language. Individuals are presented with four color pictures as response options on a page. For each item, the examiner says a word, and the examinee responds by pointing to the picture that best illustrates the meaning of the word. Overall scores are standard scores with a mean of 100 and a standard deviation of 15, with higher scores indicating more-developed abilities. Psychometrics for this measure are high, with an internal consistency reliability coefficient alpha of .97 and an adjusted test-retest reliability coefficient of .93.

Outcome Measures

Robot-Based Emotion Skills Knowledge (R-BESK). The *R-BESK* is the measure of emotion identification accuracy as measured by participant responses to robot prompts during the robot games at T1 and T2. Participants in both groups completed this measure. Children were asked to match SAM's emotional facial expressions to the emotions displayed in schematic face drawings (MacDonald, Kirkpatrick, and Sullivan, 1996; *MATCH-D)* and photos of human faces (Ekman and Friesen, 1975; *MATCH-F)*. Each of the six target emotions (happiness, sadness, anger, fear, surprise, and disgust) is presented four times, twice with schematic drawings and twice with photos of human faces. Accuracy for matching emotions was recorded via responses made on the touchscreen tablet, with possible scores ranging from zero correct to a maximum total correct score of

24. A previous study showed strong test-retest reliability ($r = .79$) with a slight upward shift from T1 to T2 (Koch, 2017).

Enjoyment Questionnaire (Child). The enjoyment questionnaire completed by the children in the study consisted of two questions, worded as follows: (1) “How much did you like talking to the robot?” and (2) “How much would you like to talk with the robot again?” Each question was rated on a scale from zero to 10, with higher ratings indicating increased enjoyment and motivation to return. A visual of a thermometer was used to aid children in understanding the questions and completing the questionnaire. Scores were combined for a maximum score of 20. For this study, ratings between 16 to 20 were indicative of high favorability, ratings between 10 to 15 were indicative of moderate favorability, and ratings between zero to nine were indicative of low favorability.

Enjoyment Questionnaire (Parent). The parent enjoyment questionnaire was used to measure the parent interpretation of their child’s enjoyment of the SAM robot interactions. Three questions were rated on a scale from zero to 10 yielding a maximum score of 30 on this measure. The questions were worded as follows: (1) “My child enjoyed interacting with the robot,” (2) “My child was motivated to come to the robot sessions,” and (3) “My child would like to interact with the robot again in the future.” Ratings between 24 to 30 were indicative of high favorability, ratings between 15 to 23 were indicative of moderate favorability, and ratings between zero to 14 were indicative of low favorability.

Social Responsiveness Scale, Second Edition (SRS-2), Parent and Teacher Rating Scales (Constantino & Gruber, 2012). The *SRS-2* is 65-item rating scale that focuses on the severity of social impairments common to individuals with ASD. This measure yields scores measuring social awareness, social information processing, reciprocal social communication, social anxiety and avoidance, and ASD-specific preoccupations and traits, as well as a composite total score. Standardized *t*-scores are provided, with a mean of 50 and a standard deviation of 10, and higher scores indicate greater impairment. Scores of 60-69 indicate mild impairment, and scores ≥ 70 indicate moderate to severe difficulties. Correlations for internal consistency ranged from .94 to .96.

Social Skills Improvement System (SSIS) Rating Scales, Teacher Form (Gresham & Elliott, 2008). The *SSIS* teacher rating scale provides information regarding student social skills, problem behaviors, and academic competence, and generates standard scores in these areas, with a mean of 100 and a standard deviation of 15. Higher scores in the social skills and academic competence domains indicate better abilities, whereas higher scores on the problem behaviors domain indicate more frequent and severe behavior problems. Additional subscale scores are available for a variety of individual social skills (e.g., communication, cooperation, empathy) and problem behaviors (e.g., externalizing, internalizing, hyperactivity). Internal consistency reliability coefficient alpha ranges from .96 to .97 for the targeted age range, and test-retest reliability was strong (adjusted r = .83).

Benton Facial Recognition Test (Benton), Short Form (Benton, Sivan, Hamsher, Varney, & Spreen, 1983). The *Benton* is a clinician-administered measure of facial recognition. It involves the presentation of a target face, and the participant must choose the correct match from an array of six photos. It was originally developed for use with individuals with traumatic brain injury, but it has also been used in recent research with individuals with ASD (Weigelt, Koldewyn, & Kanwisher, 2012). This study used the 27-item, short form version of this measure given the age range and expected attention span of the participants. Short form raw scores range from zero to 27. Severe impairment is defined as a raw score ≤ 17, 18 correct indicates moderate impairment, 19 items correct is borderline impaired, and scores ≥ 20 are in the normal range. There is currently no evidence on the reliability and validity of the Benton Facial Recognition test in this population; however, this measure has been used to assess facial recognition abilities in children with ASD in previous studies (Tang et al., 2015).

Data Analysis

Missing Data

One child enrolled in the intervention group was unable to complete the enjoyment questionnaire due to cognitive ability. One set of pre-test parent questionnaires was not returned. For teacher raters, one set of pre-test questionnaires was not collected, and three sets of post-test questionnaires have not been returned. One standard score could not be calculated on the *SSIS* due to too many missing items on the questionnaires. All missing data were excluded listwise from individual analyses.

Analytic Plan

The primary tests of therapy effect were conducted using one-way ANCOVAs with post-test scores as the dependent variable and pre-test scores as the covariate. Exploratory analyses followed with a one-way analysis of variance (ANOVA) with Time (2 levels: T1 and T2) as the independent variable and test scores as the dependent variable, and a within-group paired sample *t*-test for each group to investigate change over time.

Assumptions

All assumptions were tested using the following methods. Extreme outliers were defined as any value greater than ±3 *SD* from the mean. Normal distribution of the residuals was measured using the Shapiro-Wilk *W* statistic. The Brown-Forsythe modification of the Levene test was used to test for homogeneity of variances, and the Durbin-Watson test was used to evaluate error independence. For ANCOVA analyses, the absence of group differences on pre-test scores was evaluated using an independent samples *t*-test. A linear relationship between scores at pre-test (covariate) and post-test (dependent variable) was determined using visual inspection of a scatter plot. Assumptions violations for child enjoyment and emotion identification accuracy analyses are noted below. Assumptions were met for all other tests.

Participant enjoyment. For child-rated enjoyment, residuals were not normally distributed ($p < .05$), skew was severe (-1.265), and the assumption of homogeneity of variances was violated ($p < .05$). Within the intervention group ($n = 7$), all but one child gave the highest possible rating, with one child rating moderate enjoyment. In the control group (n = 8), four children gave the maximum enjoyment rating and four indicated moderate enjoyment. Given these assumption violations and score distributions, the Mann-Whitney U test was conducted. Group distributions were compared through visual inspection of a population pyramid. The asymptomatic significance level was used due to many ties in the data (i.e., many participants in both groups gave the highest rating of 20). Parent-rated participant enjoyment residuals were not normally distributed ($p < .05$). However, the *t*-test is robust to this violation when skew is not severe (Stonehouse & Forrester, 1998; severe skew defined as < -1 or > 1; skew = -.847). Therefore, an independent samples *t*-test was conducted between groups with parent-rated child enjoyment scores as the dependent variable.

Emotion identification accuracy. One child in the intervention group responded randomly to the *R-BESK* task and scored much lower than other children. This participant was an outlier and was excluded from this analysis. The data showed a ceiling effect, and the within-group residuals were not normally distributed, as measured by Shapiro-Wilk's W statistic (p < .05). However, ANCOVA is robust to this violation and the analysis proceeded as planned.

RESULTS

Enjoyment and Motivation

Children and parents in both groups were expected to rate participant enjoyment as high. As shown in Table 5, as anticipated, child-rated enjoyment was high for both groups, and median enjoyment scores were not statistically significantly different between groups. Additionally, parent-rated participant enjoyment was high for both groups, and there was no difference between groups on this measure.

Emotion Identification Accuracy

The intervention group was expected to improve on emotion identification skills more than the control group. Table 6 shows that the analysis revealed no significant difference in *R-BESK* accuracy between groups. An ANOVA conducted across all participants in both groups revealed a significant main effect of time. Comparing change scores within groups, the intervention group showed a significant increase in accuracy over time ($t(6) = 2.521$, $p < .05$), whereas the control group did not ($t(7) = 2.040$, $p > .05$).

Generalized Social-Emotional Skills

For measures of social skills, it was predicted that scores would improve to a greater extent for the intervention group compared to the control group. Note that the *SRS-2* is a measure of impairment, therefore a decrease in scores indicates improvement,

Table 5

Participant Enjoyment Ratings

Participant Enjoyment	Intervention	Control	Combined Sample	Group Effect	
Participant Ratings (max = 20)	*n* = 8	*n* = 7	*n* = 15	*U*	*p*
M (SE)	19.3 (.7)	16.5 (1.4)	17.8 (.9)		
Mdn	20.0	17.5	20.0	39	.128
Range	15-20	10-20	10-20		
Parent Ratings (max = 30)	*n* = 8	*n* = 8	*n* = 16	*t* (1,14)	*p*
M (SE)	25.4 (5.3)	27.4 (4.1)	26.4 (4.7)	.845	.412
*Mdn (*Range)	30	26.5	30		
Range	19-30	17-30	17-30		

Table 6

Comparison for Outcome Measures across T1 and T2

	Intervention	Control	Group Effect		Combined Sample	Time Effect	
Variable	*M* (*SE*)	*M* (*SE*)	*F*	*p*	*M* (*SE*)	*F*	*p*
R-BESK	*n* = 7	*n* = 8			*n* = 15		
T1	19.4 (1.3)	18.1 (1.1)			18.7 (.8)		
T2	22.7 (1.0)	20.7 (1.0)	2.313	.154	21.6 (.7)		
T2-T1	3.4 (1.4)	2.4 (1.2)			2.9 (.9)	7.057	.013*
Social-Emotional Skills							
SRS-2 Parent	*n* = 8	*n* = 7			*n* = 15		
T1	81.1 (2.0)	78.1 (2.6)			79.7 (1.6)		
T2	75.5 (2.4)	71.3 (2.5)	1.434	.254	73.5 (2.4)		
T2-T1	-4.3 (1.8)	-8.4 (2.9)			-6.2 (1.7)	4.616	.040*
SRS-2 Teacher	*n* = 6	*n* = 6			*n* = 12		
T1	82.0 (4.4)	68.5 (4.6)			75.3 (3.7)		
T2	72.3 (4.8)	69.5 (4.8)	.148	.710	70.9 (2.9)		
T2-T1	-8.3 (3.4)	-.3 (6.9)			-4.3 (3.9)	.854	.366

*p < .05

Note: Adjusted means and standard errors are provided for T2 scores. A decrease in *SRS-2* scores indicates improvement.

Table 6 (continued)

Comparison for Outcome Measures across T1 and T2

	Intervention	Control	Group Effect		Combined Sample	Time Effect	
Variable	*M* (*SE*)	*M* (*SE*)	*F*	*p*	*M* (*SE*)	*F*	*p*
SSIS Teacher	*n* = 5	*n* = 6			*n* = 11		
T1	78.2 (6.7)	86.8 (4.2)			82.9 (3.9)		
T2	84.7 (4.9)	86.4 (4.5)	.067	.802	85.1 (3.7)		
T2-T1	3.2 (2.5)	2.3 (5.5)			2.7 (3.1)	.166	.688
Benton	*n* = 8	*n* = 8			*n* = 16		
T1	16.6 (1.6)	19.6 (1.0)			18.1 (1.0)		
T2	18.5 (1.0)	18.5 (1.0)	.002	.962	18.5 (.8)		
T2-T1	1.0 (1.1)	-.3 (1.1)			.4 (.8)	.080	.779

$^{*}p < .05$
Note: Adjusted means and standard errors are provided for T2 scores.

whereas for the *SSIS*, higher scores signify more developed skills. As indicated in Table 6, the analysis revealed no significant difference in parent-rated social skills between groups, and an exploratory, one-way ANOVA revealed a significant main effect of time across all participants. Both the intervention group and control group showed significant improvement over time within each group (Intervention: $t(7) = -2.387$, $p < .05$; Control: $t(6) = -2.946$, $p < .05$). For teacher-rated overall social skills, analyses revealed no significant difference between groups. Exploratory analyses revealed no significant main effect of time. Within-group *t*-tests also showed no significant improvement in scores for either group, although the intervention group trended toward significance (Intervention: $t(5) = -2.421$, $p = .06$; Control: $t(5) = -0.48$, $p > .05$).

Teacher ratings of classroom-based social skills did not differ between groups. Exploratory analyses revealed no main effect of time. *T*-tests were conducted to explore the change within each group, and neither group showed a significant improvement in social skills at school (Intervention: $t(4) = 1.291$, $p > .05$; Control: $t(5) = .422$, $p > .05$).

Facial Recognition

The intervention group was predicted to improve on facial recognition abilities moreso than the control group. Data for this analysis are presented in Table 6. Analyses revealed no significant difference in facial recognition accuracy between groups, and exploratory analyses revealed no main effect of time. Facial recognition accuracy did not improve within either group from pre-test to post-test (Intervention: $t(7) = .907$, $p > .05$; Control: $t(7) = -.218$, $p > .05$).

DISCUSSION

The purpose of this study was to investigate the effectiveness of the SAM robot intervention for school-age children with ASD across a wide range of cognitive ability. The initial analyses focused on the child's enjoyment and motivation to participate in the intervention. It was noted that one child who enrolled in the study experienced anxiety while interacting with the robot and discontinued participation following the T1 session. The comorbidity of anxiety disorder and ASD is high, and current estimates range from 42% to 79% (Kerns et al., 2017). Although precautions were taken during the study to decrease participant anxiety as much as possible, the SAM robot intervention is contraindicated for children who experience significant anxiety while interacting with the robot. Given that the other children enrolled in the study were able to attend all sessions and complete the study, in general, the structure of this intervention can be completed with this population.

One child in the intervention group did not understand the nature of the task and was unable to focus on the intervention. This participant selected answer at random throughout the duration of the intervention. Future participants should be screened for their ability to sit at a table and attend to a task and should be excluded from participating if they are unable to do so. Other therapeutic techniques to address attention and behavior difficulties, such as ABA, would be appropriate for these children prior to participating in this intervention or any similar robot- or technology-based interventions.

As hypothesized, both parent and participant ratings indicated high levels of child enjoyment, motivation, and willingness to interact with the robot again across all participants. There was no difference between groups on this measure. This is of particular importance when considering intervention methods for children with ASD. Given that engagement and motivation can be difficult in this population, and that children who are intrinsically motivated by the learning process will be more likely to benefit from it, continuing to pursue the methodology of robot-based interventions with this population is a worthwhile endeavor.

Secondly, this study aimed to determine whether emotion recognition and identification abilities were improved following the SAM robot intervention compared to controls. Contrary to expectations, accuracy on this measure at post-test did not differ significantly between groups when controlling for pre-test scores. However, overall, all participants improved from pre-test to post-test, and this change was significant for the intervention group but not the control group. It is possible that the small sample size reduced the power of the ANCOVA to find meaningful differences between groups. These results should be interpreted cautiously given that the use of *t*-tests to evaluate change introduces increased error that is accounted for in the ANCOVA analysis.

Improvement on the emotion identification accuracy on the *R-BESK* in both groups, as well as a main effect of time across the combined sample, is likely the result of practice effects. The T1 and T2 *R-BESK* sessions contain randomized, identical stimuli and consist of a relatively simple emotion matching task of schematic faces and face photos. It is likely that participants learned the appropriate emotion matching responses during the T1 session, and these skills carried over to the T2 session, seven sessions later, even for control children who had not been exposed to the stimuli in the intervening weeks.

The third set of analyses considered the effect of the robot intervention on generalized social-emotional skills, as rated by parents and teachers. Unexpectedly, no differences between groups were seen on ASD-specific social skills by parents or teachers or on teacher-rated classroom social skills. However, the overall combined sample improved over time for parent-rated social skills, and analysis of change scores showed significant improvement over time within both the intervention group and the control group separately for these ratings. This indicates that parents rated global social skills improvement for the overall sample, including for those in the control group. This is certainly a curious finding and may imply a possible placebo effect. Although parents were blind to group membership, given the time and effort required for study completion, parents may have been hopeful that the participant had been enrolled in the intervention group. This may have resulted in inadvertently better ratings following study completion. Additionally, information was not collected from parents regarding any concurrent social skills interventions in which children may have been participating, such as school or outpatient therapies, that may have resulted in social skills improvement outside of the study. Finally, the effect of the SAM robot intervention on improving facial recognition abilities was examined, and the the intervention group did not show improvements.

The lack of group differences in skill acquisition and generalization may be explained by a variety of factors. It is possible that although using a robot-based approach for teaching skills to this population is enjoyable and motivating, the current model of the SAM robot intervention is not effective for this population in acquiring directly taught social-emotional knowledge or generalizing skills to home, school, and community settings. It could be that the “dose” of the intervention, including session frequency, duration, and/or total number of sessions, needs to be increased to see meaningful improvement. Additionally, the content of the intervention may need to be improved to be accessible for this lower ability population. For example, the receptive language required for the sessions involving brief emotion-based scenarios may have been too complex for some participants to fully understand.

The measures chosen in this study may not have been adequately sensitive to find specific skill improvements gained from this relatively short intervention, particularly given the small sample size and decreased power of this study. Another possibility is that the control group did not serve as a true control, since they were exposed to an emotion-related activity at pre-test. Additionally, although the control activity with SAM was designed to be completely “emotion-free” with no head or face movement and no speech, it is possible, albeit unlikely, that this activity had inadvertent positive effects on social-emotional processing and understanding.

Future Directions

One major flaw with this study was the inadequate power to find significant group differences due to small sample size. However, the majority of the robot intervention literature with this population has utilized case study or case series designs of eight or fewer participants. Therefore, inclusion of 16 participants in a controlled trial is expected to contribute meaningfully to the field. Additionally, this sample only enrolled one female participant and parents tended to be well-educated and upper-middle class. Future research should attempt to reflect the gender ratio of males to females typically seen in ASD (4:1) and should recruit a diverse sample that is more representative of the ASD population. This would allow for more accurate generalization to the global ASD population. The design of this study is certainly a strength compared to other robot intervention studies for this population. The use of a control group that had equal exposure to the robot compared to the intervention group allowed for adequate blinding of families and teachers, whereas other similar studies have utilized a wait-list controlled trial in which parents, teachers, and participants were not blind to group membership. Future research should consider employing an A/B study design where control participants are not exposed to intervention stimuli at baseline. This design would allow all participants to do the emotion intervention as well as the dance activity with the robot, and would increase power, as there would be twice as many observations and participants could be compared to their own performance at baseline.

Children with ASD and below average cognitive ability frequently are excluded from research investigating effective social skills interventions. Future research studies in this area should strive to include individuals with ASD across all levels of cognitive skills, with a particular focus in including individuals with below average IQ in intervention research studies. Although recruitment of these families was certainly a challenge, inclusion of these individuals in this study provides evidence that conducting research with this population is possible with adequate time and effort and when proper supports are in place.

Researchers in this area should consider conducting additional randomized controlled trials in which participants and researchers are blind to group membership and should include different combinations of intervention dose and content to identify the critical elements of social robot interventions for this population. Future research with the SAM robot will need to enroll a larger sample size to increase power and investigate which participant characteristics are predictive of social skills improvement.

LIST OF REFERENCES

Amenta, S., Ferrari, C. & Balconi, M. (2014). Facial expression decoding in autistic and asperger children. In V. B. Patel, V. R. Preedy & C. R. Martin (Eds.), *Comprehensive guide to autism* (pp. 1885–1904). New York: Springer.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: Author.

Back, E., Ropar, D. & Mitchell, P. (2007). Do the eyes have it? Inferring mental states from animated faces in autism. *Child Development*, *78*(2), 397–411.

Banda, D. R., Hart, S. L. & Liu-Gitz, L. (2010). Impact of training peers and children with autism on social skills during center time activities in inclusive classrooms. *Research in Autism Spectrum Disorders*, *4*(4), 619–625.

Barry, T. D., Klinger, L. G., Lee, J. M., Palardy, N., Gilmore, T. & Bodin, S. D. (2003). Examining the effectiveness of an outpatient clinic-based social skills group for high-functioning children with autism. *Journal of Autism and Developmental Disorders*, *33*(6), 685–701.

Begum, M., Serna, R. W. & Yanco, H. A. (2016). Are robots ready to deliver autism interventions? A comprehensive review. *International Journal of Social Robotics*, *8*(2), 157–181.

Bellini, S. (2004). Social skill deficits and anxiety in high-functioning adolescents with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, *19*(2), 78–86.

Bellini, S. (2006). The development of social anxiety in adolescents with autism spectrum disorders. *Focus on Autism and Other Developmental Disabilities*, *21*(3), 138–145.

Bellini, S. & Akullian, J. (2007). A meta-analysis of video modeling and video self-modeling interventions for children and adolescents with autism spectrum disorders. *Exceptional Children*, *73*(3), 264–287.

Benton, A. L., Sivan, A. B., Hamsher, K., Varney, N. R., & Spreen O. (1983). *Benton facial recognition test*. Lutz, FL: Psychological Assessment Resources (PAR), Incorporated.

Bitsika, V. & Sharpley, C. F. (2004). Stress, anxiety and depression among parents of children with autism spectrum disorder. *Australian Journal of Guidance and Counselling*, *14*(2), 151–161.

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*, *3*(4), 337–347.

Boccanfuso, L., Scarborough, S., Abramson, R. K., Hall, A. V., Wright, H. H. & O'Kane, J. M. (2016). A low-cost socially assistive robot and robot-assisted intervention for children with autism spectrum disorder: Field trials and lessons learned. *Autonomous Robots*, *41*(3), 637–655.

Bohlander, A. J., Orlich, F. & Varley, C. K. (2012). Social skills training for children with autism. *Pediatric Clinics of North America*, *59*(1), 165–74. doi:10.1016/j.pcl.2011.10.001

Buescher, A. V., Cidav, Z., Knapp, M. & Mandell, D. S. (2014). Costs of autism spectrum disorders in the United Kingdom and the United States. *JAMA Pediatrics*, *168*(8), 721–728.

Burgess, A. F. & Gutstein, S. E. (2007). Quality of life for people with autism: Raising the standard for evaluating successful outcomes. *Child and Adolescent Mental Health*, *12*(2), 80–86.

Centers for Disease Control and Prevention. (2014). Prevalence of autism spectrum disorder among children aged 8 years. *Surveillance Summaries*, *63*(2), 1–21.

Chamak, B. & Bonniau, B. (2016). Trajectories, long-term outcomes and family experiences of 76 adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, *46*(3), 1084–1095.

Chasson, G. & Jarosiewicz, S. R. (2014). Social Competence Impairments in Autism Spectrum Disorders. In V. B. Patel, V. R. Preedy & C. R. Martin (Eds.), *Comprehensive guide to autism* (pp. 1099–1118). New York: Springer.

Conn, K., Liu, C., Sarkar, N., Stone, W. & Warren, Z. (2008). Affect-sensitive assistive intervention technologies for children with autism: An individual-specific approach. In *Robot and human interactive communication* (pp. 442–447).

Constantino, J. N. & Gruber, C.P. (2012). *Social responsiveness scale* (2nd ed). Los Angeles, CA: Western Psychological Services.

Dautenhahn, K. & Billard, A. (2002). Games children with autism can play with Robota, a humanoid robotic doll. In S. Keates, P. J. Clarkson, P. M. Langdon, & P. Robinson (Eds.), *Cambridge workshop on universal access and assistive technology* (pp. 179–190). London: Springer.

Dautenhahn, K. & Werry, I. (2004). Towards interactive robots in autism therapy: Background, motivation and challenges. *Pragmatics & Cognition*, *12*(1), 1–35.

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.

Dunn, L. & Dunn, L. (2007). *Peabody picture vocabulary test* (4th ed.). Bloomington, MN: Pearson Assessments.

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.

Eaves, L. C. & Ho, H. H. (2008). Young adult outcome of autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *38*(4), 739–47.

Ekman, P. & Friesen, W. V. (1975). *Unmasking the face: A guide to recognizing emotions from facial clues*. Englewood Cliffs, NJ: Prentice-Hall.

Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G* Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191.

Feil-Seifer. (2008). *Socially assistive robot-based intervention for children with autism spectrum disorder*. In *Proceedings of NEWHRI Workshop*, Pasadena, CA

Feil-Seifer, D. & Matari'c, M. J. (2009). Toward socially assistive robotics for augmenting interventions for children with autism spectrum disorders. *Experimental Robotics: The Eleventh International Symposium*, *54*, 201–210.

Flynn, L. & Healy, O. (2012). A review of treatments for deficits in social skills and self-help skills in autism spectrum disorder. *Research in Autism Spectrum Disorders*, *6*(1), 431–441.

François, D., Powell, S. & Dautenhahn, K. (2009). A long-term study of children with autism playing with a robotic pet: Taking inspirations from non-directive play therapy to encourage children's proactivity and initiative-taking. *Interaction Studies*, *10*(3), 324–373.

Goldsmith, T. R. & LeBlanc, L. A. (2004). Use of technology in interventions for children with autism. *Journal of Early and Intensive Behavior Intervention*, *1*(2), 166.

Gray, C. A. & Garand, J. D. (1993). Social stories: Improving responses of students with autism with accurate social information. *Focus on Autistic Behavior*, *8*(1), 1–10.

Gresham, F., & Elliott, S. (2008). *Social skills improvement system*. Minneapolis, MN: Pearson Assessments.

Grossmann, T. & Johnson, M. H. (2007). The development of the social brain in human infancy. *The European Journal of Neuroscience*, *25*(4), 909–19.

Hanley, G. P., Iwata, B. A. & Thompson, R. H. (2001). Reinforcement schedule thinning following treatment with functional communication training. *Journal of Applied Behavior Analysis*, *34*(1), 17–38.

Hopkins, I. M., Gower, M. W., Perez, T. A., Smith, D. S., Amthor, F. R., Wimsatt, F. C. & Biasini, F. J. (2011). Avatar assistant: improving social skills in students with an ASD through a computer-based intervention. *Journal of Autism and Developmental Disorders*, *41*(11), 1543–1555.

Howlin, P. (2000). Outcome in adult life for more able individuals with autism or Asperger syndrome. *Autism*, *4*(1), 63–83.

Izard, C., Fine, S., Schultz, D., Mostow, A., Ackerman, B. & Youngstrom, E. (2001). Emotion knowledge as a predictor of social behavior and academic competence in children at risk. *Psychological Science*, *12*(1), 18–23.

Karkhaneh, M., Clark, B., Ospina, M. B., Seida, J. C., Smith, V. & Hartling, L. (2010). Social Stories^{TM} to improve social skills in children with autism spectrum disorder: A systematic review. *Autism*, 641–662.

Kaufman, A.S. & Kaufman, N.L. *Kaufman brief intelligence test* (2nd ed.). Circle Pines, MN: American Guidance Services.

Kerns, C. M., Renno, P., Storch, E. A., Kendall, P. C., & Wood, J. J. (2017). *Anxiety in Children and Adolescents with Autism Spectrum Disorder: Evidence-Based Assessment and Treatment*. London: Academic Press.

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

Kleinhans, N. M., Richards, T., Weaver, K., Johnson, L. C., Greenson, J., Dawson, G. & Aylward, E. (2010). Association between amygdala response to emotional faces and social anxiety in autism spectrum disorders. *Neuropsychologia*, *48*(12), 3665–3670.

Koch, S. A. (2017). *Effectiveness and Acceptability of a Robot-Based Social Skills Intervention for Children with Autism Spectrum Disorder* (Doctoral dissertation).

Koch, S. A., Stevens, C. E., Clesi, C. D., Lebersfeld, J. B., Sellers, A. G., McNew, M. E., Biasini, F. J. & Hopkins, M. I. (2017). A Feasibility Study Evaluating the Emotionally Expressive Robot SAM. *International Journal of Social Robotics*, *9*(4), 601–613.

Koegel, L. K., Singh, A. K. & Koegel, R. L. (2010). Improving motivation for academics in children with autism. *Journal of Autism and Developmental Disorders*, *40*(9), 1057–66. doi:10.1007/s10803-010-0962-6

Koegel, R. L. & Koegel, L. K. (2012). *The PRT Pocket Guide: Pivotal Response Treatment for Autism Spectrum Disorders*. Baltimore, MD: Brookes Publishing.

Kozima, H., Michalowski, M. P. & Nakagawa, C. (2009). Keepon. *International Journal of Social Robotics*, *1*(1), 3–18.

Kozima, H., Nakagawa, C. & Yasuda, Y. (2005). Interactive robots for communication-care: A case-study in autism therapy. In *Institute of Electrical and Electronics Engineers (IEEE) international workshop on robots and human interactive communication* (pp. 341–346).

Kroeger, K., Schultz, J. R. & Newsom, C. (2007). A comparison of two group-delivered social skills programs for young children with autism. *Journal of Autism and Developmental Disorders*, *37*(5), 808–817.

Laugeson, E. A., Frankel, F., Gantman, A., Dillon, A. R. & Mogil, C. (2012). Evidence-based social skills training for adolescents with autism spectrum disorders: The UCLA PEERS program. *Journal of Autism and Developmental Disorders*, *42*(6), 1025–36. doi:10.1007/s10803-011-1339-1

Laushey, K. M. & Heflin, L. J. (2000). Enhancing social skills of kindergarten children with autism through the training of multiple peers as tutors. *Journal of Autism and Developmental Disorders*, *30*(3), 183–193.

Lee, J., Takehashi, H., Nagai, C. & Obinata, G. (2012). Design of a therapeutic robot for interacting with autistic children through interpersonal touch. In *IEEE RO-MAN: The 21st IEEE international symposium on robot and human interactive communication* (pp. 712–717).

Leite, I., Pereira, A., Martinho, C. & Paiva, A. (2008). Are emotional robots more fun to play with? In *The 17th IEEE international symposium on robot and human interactive communication* (pp. 77–82).

Lord, C. & Bishop, S. L. (2010). Autism spectrum disorders: Diagnosis, prevalence, and services for children and families. *Society for Research in Child Development*, *24*(2).

Lord, C., Rutter, M., DiLavore, P., Risi, S., Gotham, K., & Bishop, S. (2012). *Autism diagnostic observation schedule* (2nd ed.). Torrance, CA: Western Psychological Services.

Lovaas, O. I. (1987). Behavioral treatment and normal educational and intellectual functioning in young autistic children. *Journal of Consulting and Clinical Psychology*, *55*(1), 3–9.

MacDonald, P. M., Kirkpatrick, S. W. & Sullivan, L. A. (1996). Schematic drawings of facial expressions for emotion recognition and interpretation by preschool-aged children. *Genetic, Social, and General Psychology Monographs*, *122*(4), 373–388.

McGovern, C. W. & Sigman, M. (2005). Continuity and change from early childhood to adolescence in autism. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, *46*(4), 401–8.

Narzisi, A., Muratori, F., Calderoni, S., Fabbro, F. & Urgesi, C. (2013). Neuropsychological profile in high functioning autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *43*(8), 1895–1909.

Paul, R. (2008). Interventions to improve communication in autism. *Child and Adolescent Psychiatric Clinics of North America*, *17*(4), 835–856.

Pennisi, P., Tonacci, A., Tartarisco, G., Billeci, L., Ruta, L., Gangemi, S. & Pioggia, G. (2016). Autism and social robotics: A systematic review. *Autism Research*, *9*(2), 165–183.

Pioggia, G., Igliozzi, R., Ferro, M., Ahluwalia, A., Muratori, F. & De Rossi, D. (2005). An android for enhancing social skills and emotion recognition in people with autism. In *IEEE transactions on neural systems and rehabilitation engineering*, *13*(4), 507–15.

Prothmann, A., Ettrich, C. & Prothmann, S. (2009). Preference for, and responsiveness to, people, dogs and objects in children with autism. *Anthrozoös*, *22*(2), 161–171.

Putnam, C. & Chong, L. (2008). Software and technologies designed for people with autism: What do users want? In *Proceedings of the 10th international ACM SIGACCESS conference on computers and accessibility* (pp. 3–10).

Rabbitt, S. M., Kazdin, A. E. & Scassellati, B. (2015). Integrating socially assistive robotics into mental healthcare interventions: Applications and recommendations for expanded use. *Clinical Psychology Review*, *35*, 35–46.

Ricks, D. J. & Colton, M. B. (2010). Trends and considerations in robot-assisted autism therapy. In *Robotics and Automation (ICRA), 2010 IEEE International Conference on* (pp. 4354–4359).

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.

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.

Robins, B., Dickerson, P., Stribling, P. & Dautenhahn, K. (2004). Robot-mediated joint attention in children with autism: A case study in robot-human interaction. *Interaction Studies*, *5*(2), 161–198.

Scassellati, B. (2007). How social robots will help us to diagnose, treat, and understand autism. In S. Thrun, R. Brooks & H. Durrant-Whyte (Eds.) *Robotics research: Results of the 12th international symposium ISRR* (pp. 552–563). New York: Springer.

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

Shamsuddin, S., Yussof, H., Ismail, L. I., Mohamed, S., Hanapiah, F. A. & Zahari, N. I. (2012). Humanoid robot NAO interacting with autistic children of moderately impaired intelligence to augment communication skills. *Procedia Engineering*, *41*, 1533–1538.

Shamsuddin, S., Yussof, H., Mohamed, S., Hanapiah, F. A. & Ismail, L. I. (2013). Stereotyped behavior of autistic children with lower IQ level in HRI with a humanoid robot. In *Advanced Robotics and its Social Impacts (ARSO), IEEE Workshop on* (pp. 175–180).

Smith, V. & Sung, A. (2014). Computer Interventions for ASD. In V. B. Patel, V. R. Preedy & C. R. Martin (Eds.), *Comprehensive guide to autism* (pp. 2173–2189). New York: Springer.

Stonehouse, J. M. & Forrester, G. J. (1998). Robustness of the t and U tests under combined assumption violations. *Journal of Applied Statistics*, *25*(1), 63–74.

Tang, J., Falkmer, M., Horlin, C., Tan, T., Vaz, S. & Falkmer, T. (2015). Face recognition and visual search strategies in autism spectrum disorders: Amending and extending a recent review by Weigelt et al. *PloS One*, *10*(8), e0134439.

Virués-Ortega, J. (2010). Applied behavior analytic intervention for autism in early childhood: Meta-analysis, meta-regression and dose-response meta-analysis of multiple outcomes. *Clinical Psychology Review*, *30*(4), 387–399.

Wada, K., Shibata, T., Saito, T., Sakamoto, K. & Tanie, K. (2005). Psychological and social effects of one year robot assisted activity on elderly people at a health service facility for the aged. In *Robotics and automation: Proceedings of the 2005 IEEE international conference on* (pp. 2785–2790).

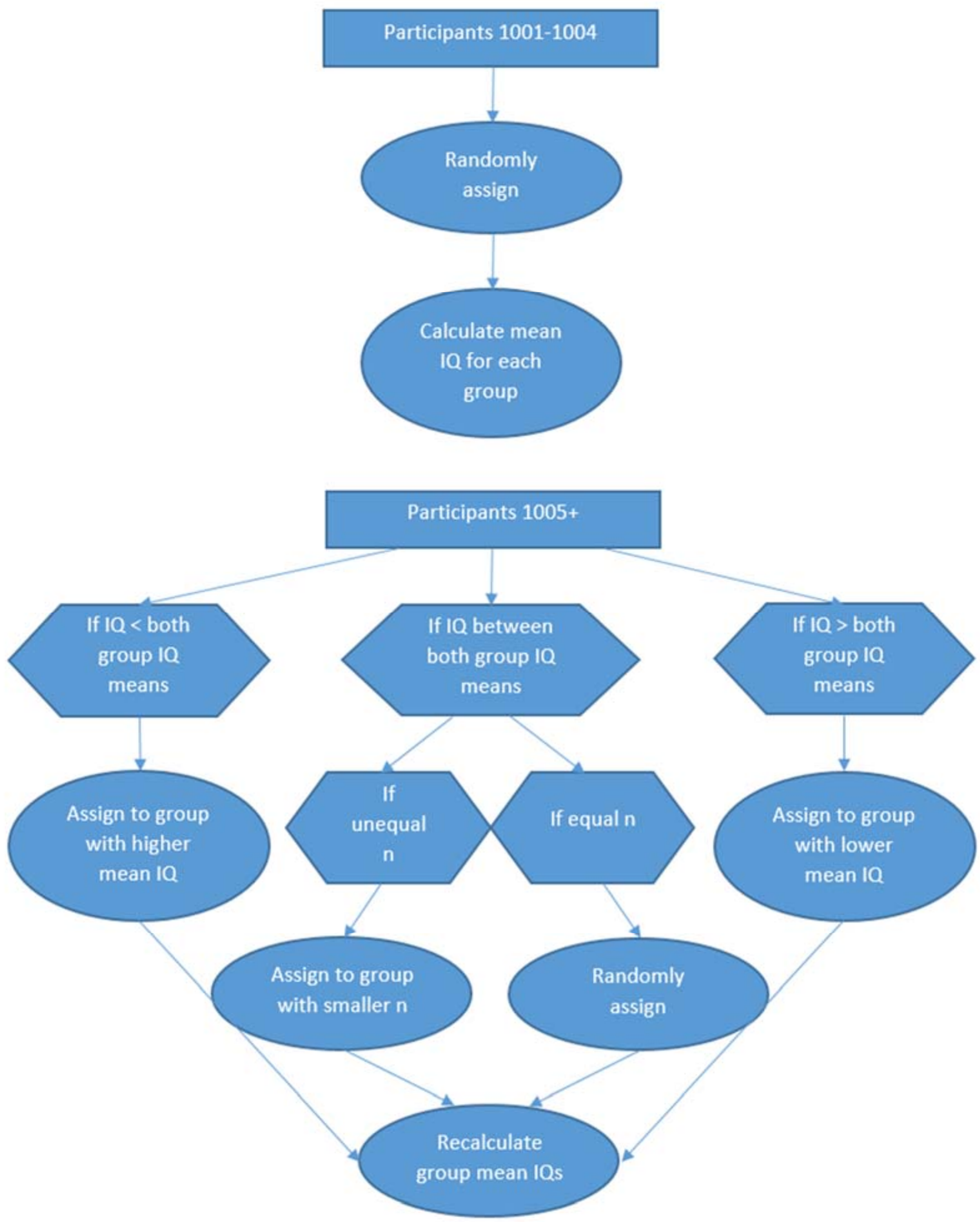
Wainer, J., Dautenhahn, K., Robins, B. & Amirabdollahian, F. (2014). A pilot study with a novel setup for collaborative play of the humanoid robot KASPAR with children with autism. *International Journal of Social Robotics*, *6*(1), 45–65.

Weigelt, S., Koldewyn, K. & Kanwisher, N. (2012). Face identity recognition in autism spectrum disorders: a review of behavioral studies. *Neuroscience \& Biobehavioral Reviews*, *36*(3), 1060–1084.

White, S. W., Keonig, K. & Scahill, L. (2007). Social skills development in children with autism spectrum disorders: a review of the intervention research. *Journal of Autism and Developmental Disorders*, *37*(10), 1858–68.

APPENDIX A

GROUP ASSIGNMENT ALGORITHM


Participants 1001-1004
Randomly assign
Calculate mean IQ for each group
Participants 1005+
If IQ < both group IQ means
If IQ between both group IQ means
If IQ > both group IQ means
Assign to group with higher mean IQ
If unequal n
If equal n
Assign to group with lower mean IQ
Assign to group with smaller n
Randomly assign
Recalculate group mean IQs

APPENDIX B

ETHICAL APPROVAL

THE UNIVERSITY OF ALABAMA AT BIRMINGHAM

Institutional Review Board for Human Use

Form 4: IRB Approval Form
Identification and Certification of Research
Projects Involving Human Subjects

UAB's Institutional Review Boards for Human Use (IRBs) have an approved Federalwide Assurance with the Office for Human Research Protections (OHRP). The Assurance number is FWA00005960 and it expires on November 8, 2021. The UAB IRBs are also in compliance with 21 CFR Parts 50 and 56.

Principal Investigator:	LEBERSFELD, JENNA
Co-Investigator(s):	BIASINI, FRED J O'KELLEY, SARAH ELIZABETH
Protocol Number:	**X151007002**
Protocol Title:	*Effectiveness of a Robot Social Skills Therapy for Children with Autism Spectrum Disorder*

The IRB reviewed and approved the above named project on 12/22/16. The review was conducted in accordance with UAB's Assurance of Compliance approved by the Department of Health and Human Services. This Project will be subject to Annual continuing review as provided in that Assurance.

This project received EXPEDITED review.

IRB Approval Date: 12/22/16

Date IRB Approval Issued: 12/22/16

IRB Approval No Longer Valid On: 12/22/17

Expedited Reviewer
Member - Institutional Review Board
for Human Use (IRB)

Investigators please note:

The IRB approved consent form used in the study must contain the IRB approval date and expiration date.

IRB approval is given for one year unless otherwise noted. For projects subject to annual review research activities may not continue past the one year anniversary of the IRB approval date.

Any modifications in the study methodology, protocol and/or consent form must be submitted for review and approval to the IRB prior to implementation.

Adverse Events and/or unanticipated risks to subjects or others at UAB or other participating institutions must be reported promptly to the IRB.

470 Administration Building
701 20th Street South
205.934.3789
Fax 205.934.1301
irb@uab.edu

The University of
Alabama at Birmingham
Mailing Address:
AB 470
1720 2ND AVE S
BIRMINGHAM AL 35294-0104