Research Article **Open Access**

Wing-Yue Geoffrey Louie*, Jessica Korneder, Ibrahim Abbas, and Cristyn Pawluk

A Study on an Applied Behavior Analysis Based Robot-Mediated Listening Comprehension Intervention for ASD

Abstract: Autism spectrum disorder (ASD) is a life-long developmental condition that affects an individual's ability to communicate and relate to others. Despite such challenges, early intervention during childhood development has shown to have positive long-term benefits for individuals with ASD. Namely, early childhood development of communicative speech skills has shown to improve future literacy and academic achievement. However, the delivery of such interventions is often time-consuming. Socially assistive robots (SARs) are a potential strategic technology that could help support intervention delivery for children with ASD and increase the number of individuals that healthcare professionals can positively affect. For SARs to be effectively integrated in real-world treatment for individuals with ASD they should follow current evidence-based practices used by therapists such as Applied Behavior Analysis (ABA). In this work, we present a study that investigates the efficacy of applying well-known ABA techniques to a robot-mediated listening comprehension intervention delivered to children with ASD at a university-based ABA clinic. The interventions were delivered in place of human therapists to teach study participants a new skill as a part of their overall treatment plan. All the children participating in the intervention improved in the skill being taught by the robot and enjoyed interacting with the robot, as evident by high occurrences of positive affect as well as engagement during the sessions. One of the three participants has also reached mastery of the skill via the robot-mediated interventions.

Keywords: Socially Assistive Robots, Autism Spectrum Disorder, Applied Behavior Analysis, Long-Term Human-Robot Interaction Study, Discrete Trial Training

***Corresponding Author: Wing-Yue Geoffrey Louie, Ibrahim Abbas:** Intelligent Robotics Laboratory at Oakland University, Rochester, MI 48307 USA, E-mail: louie@oakland.edu

Jessica Korneder, Cristyn Pawluk: Applied Behavior Analysis Clinic at Oakland University, Rochester, MI 48307 USA, E-mail: korneder@oakland.edu

1 Introduction

The Centers for Disease Control estimates that 1 in 54 children have autism spectrum disorder (ASD) in the United States [1, 2]. ASD is a life-long developmental condition that affects an individual's capability to communicate and relate to others. The field of Applied Behavior Analysis (ABA) has evidence-based treatments for addressing communication and social deficits as well as reducing challenging behaviors in these populations [3]. These methods are utilized to teach communication, social, play, group instruction, and academic behaviors as well as activities of daily living. In order to provide effective treatment outcomes, such methods should begin during early childhood (i.e. 1.5 - 6 years old) and it is recommended that children receive 20-40 hours a week of one-on-one instruction from a healthcare professional with training in behavior analysis [4]. Such intensive one-on-one treatment limits the number of individuals that healthcare professionals can positively affect.

Socially assistive robots (SARs) could be a supportive technology to these healthcare professionals as they provide intervention to individuals with ASD. Numerous SARs have already been developed and utilized for interventions targeted towards individuals with ASD. Some applications of these robot-mediated interventions for individuals with ASD have included: 1) imitation therapy [5, 6], 2) improving social skills (e.g. turn taking, joint attention, eye gaze, greetings/goodbyes) [7, 8], 3) encouraging self-initiated social interactions [5, 9, 10], 4) reducing challenging behaviors [11, 12], and 5) improving emotion recognition [12, 13]. In general, these robot-mediated interventions have had positive outcomes with behavior-based changes for individuals with ASD. However, it remains difficult to integrate this technology in real-world treatment plans because

current robot-mediated interventions do not follow ABA standards used for the treatment of individuals with ASD [14].

ABA focuses on addressing the unique deficits and behavioral excesses in individuals with ASD. Since these deficits and/or excesses are unique to each individual the following standard procedure is used in real world ABA clinical practice: 1) assessment of deficits and excesses in an individual, 2) developing an intervention plan to address an individual's unique needs, 3) and implementing the intervention with on-going frequent reassessment [15]. Namely, the first step in ABA clinical practice is to perform assessments to determine the appropriate goals for each individual using tools such as the Verbal Behavior – Milestones Assessment and Placement Program (VB-MAPP) [16] and Vineland Adaptive Behavior Scale [17]. This is followed by the development of a clear intervention plan which includes clearly defined procedures for instruction, error correction, prompt fading, reinforcement and performance data collection. Finally, the intervention is implemented as a part of an individual's overall treatment plan and his/her performance is frequently reassessed to modify the level of prompting or intervention procedures as necessary. Previous robot-mediated ABA-based intervention research has primarily focused on intervention development and implementation independent of an individual with ASD's treatment program. Furthermore, they have not focused on teaching new skills to individuals with ASD.

The long-term goal of our research team is to integrate autonomous SARs into clinical practice in order to extend the capabilities of therapists for providing accessible therapies to individuals with ASD. Our current research efforts have been towards investigating the efficacy of SARs in delivering interventions as a part of a curriculum for children with ASD to meet developmental milestones.

In this work, we present a study which evaluates the efficacy of an ABA-based robot-mediated listening comprehension intervention for children 3-8 years old with ASD at an ABA clinic. The intervention was developed in close collaboration with a Board-Certified Behavior Analyst (second author) to follow standard ABA procedures and replicate the behaviors of human therapists. To the authors' knowledge this is the first robot-mediated intervention that has a robot deliver a ABA-based intervention to teach individuals with ASD a completely new skill as a part of their overall treatment plan.

2 Background

ABA therapy structures instruction using an Antecedent, Behavior, and Consequence (ABC) model [3]. Antecedent, also known as a discriminative stimulus (S^D), refers to the environmental and social factors that occur before a desired target behavior. Behavior is an individual's response to the antecedent. Consequence refers to what occurs immediately after a behavior, which can include methods such as verbal praise, rewards, warnings, or error corrections. As an illustrative example, an antecedent could be a greeting (e.g. saying "hi" and waving) delivered by a SAR to Mary, a child with ASD. The behavior is then whether Mary greets the robot back according to the presented learning opportunity. Consequence would then be the robot providing verbal praise such as "Great job!" or rewards such as a healthy snack to the child for a correct behavior (i.e. greeting the robot).

There are several types of ABA instruction typically utilized by therapists in clinical settings such as Discrete Trial Training (DTT), Pivotal Response Training (PRT), and Verbal Behavior Approach (VBA) [18]-[20]. These types of ABA instruction vary from more structured (i.e. DTT) to more naturalistic (i.e. PRT) forms of teaching. Herein, we focus on DTT as it is a structured form of instruction that can be independently delivered by robots. DTT is an application of the principles of ABA in a structured teaching environment to teach behaviors (e.g. social, language, imitation) [18]. Namely, DTT involves providing clear instruction of a targeted behavior or set of behaviors. This includes breaking down a larger set of skills (e.g. cleaning ones room or navigating a social interaction) and systematically teaching individual behaviors (e.g. sorting toys or waving hello) that can be combined to reach the larger skill via positive reinforcement for correct responses or error corrections for incorrect responses. Prompting and careful selection of materials are also a part of the process.

To date, there has only been a handful of studies evaluating the use of SARs in DTT-based therapies [21]-[25]. In [21], the character-like SAR iRobi and CARO were used in an intervention to teach eye contact and facial expression recognition to children with ASD. The intervention consisted of observational learning where the robot first modeled the correct behavior with a human therapist and the robot then followed up by requesting the child perform the behavior. Correct responses were reinforced by providing access to social

stimuli (e.g. music). Encouragement and breaks were provided for incorrect responses. The overall efficacy of the robot-mediated intervention was evaluated over eight sessions with eight children with ASD that had difficulty distinguishing between emotions and typically made minimal eye contact. Overall, the participants decreased in eye contact and increased in their abilities to recognize the robot's emotions. However, the individual learning outcomes of the children were not reported.

In [22], the humanoid robot Nao was used to teach adolescents with ASD appropriate greeting behaviors. The intervention utilized a DTT format where the robot first greeted the individual with ASD and provided him/her an opportunity to respond correctly. Correct responses were reinforced with social praise and incorrect or no responses were followed with a least-to-most prompting protocol. Namely, a model prompt was first delivered by a therapist (i.e. demonstrating a greeting) and in the case that the individual still failed to emit the greeting behavior a verbal instruction prompt was delivered by the robot (i.e. explicitly requesting individual with ASD to provide a greeting). The efficacy of the robot-mediated intervention was evaluated with three adolescent participants with ASD that already had the ability to initiate and respond to social greetings but generally did not demonstrate the behaviors. Each child at least participated in seventeen intervention sessions but only one of the three participants were able to acquire the response to greeting in the absence of prompts.

In [23], the humanoid robot Nao delivered interventions, inspired by DTT and Floortime, to teach children with ASD imitation behaviors. The intervention consisted of the robot requesting the child to imitate its arm movements and maintaining his/her engagement during the activity. Correct responses were reinforced with social praise. Incorrect responses or no response was followed by a verbal prompt either indicating the mistake the child made or verbally instructing the child on the arm gesture he/she should execute. A study with two children with ASD demonstrated the robot could facilitate an imitation therapy with the participants but the efficacy of the intervention was not evaluated.

In [24], the humanoid Zeno robot delivered a DTT-based emotion recognition intervention to children and adolescents with ASD. The robot played three different emotion recognition games: guessing the emotion of the robot, selecting a picture of a human expressing a specific emotion, and identifying a human's emotion from a picture. Correct responses were reinforced with social praise. A least-to-most prompting protocol was used for incorrect responses and consisted of increasing the number of verbal hints provided to the individual. A study with two children with ASD over three intervention sessions did not show significant improvements in the children's ability to recognize emotions in comparison to their baseline performance.

In [25], the Nao robot facilitated gesture imitation sessions with children with ASD. The sessions were in a DTT format where the robot first presented a gesture to the child and requested him/her to copy it. The child was then reinforced with praise if he/she imitated the robot's gesture correctly. If he/she did not imitate the gesture correctly the robot would demonstrate the correct gesture and request the child to try again. A study with eight children with ASD was conducted to evaluate whether there were differences in children's engagement and performance during robot-mediated sessions in comparison to human-mediated sessions. Each child participated in a single robot-mediated intervention session and a single human-mediated intervention session. Overall, the children with ASD were more engaged and performed better in the robot-mediated sessions.

In general, current research in robot-mediated DTT-based therapies for individuals with ASD have focused on developing interventions that increase an individual's performance on an existing social skill (i.e. imitation, emotion recognition, greeting, or eye contact) within his/her repertoire using least-to-most prompting protocols [21]-[25]. To date, there has not been research on teaching a skill that is not already within an individual with ASD's existing repertoire and integrated as a part of an individual's treatment plan. Teaching a completely new skill to an individual with ASD requires the application of a set protocol of DTT techniques to be effective and efficient [18, 29]. The efficacy of such robot-mediated DTT-based interventions have also been inconclusive with both positive improvements in participant skill execution [21, 22] no change in participants' abilities [24]. With the exception of [22], most studies have also not investigated the individual improvements in participant skill execution after the interventions and/or had a limited number of intervention sessions.

The research presented in this paper extends the existing literature on robots delivering DTT-based interventions to individuals with ASD by: 1) developing a language acquisition intervention; 2) investigating the efficacy of a robot-mediated intervention that teaches a new skill not already in an individual's existing repertoire; and 3) integrating the intervention, in place of a human therapist, as a part of an individual's overall treatment plan at an ABA clinic.

Fig. 1. The humanoid Nao robot

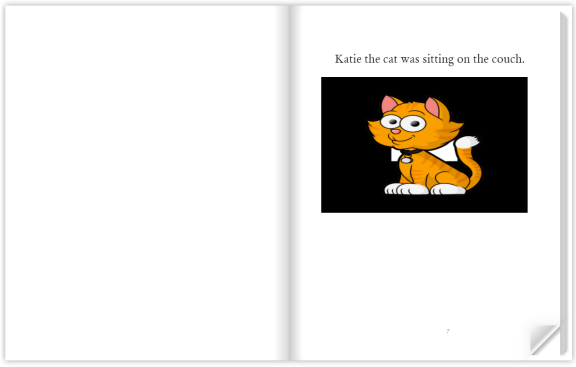

Fig. 2. Sample page from the e-book

3 Robot-Mediated Listening Comprehension Intervention

We have developed a DTT-based robot-mediated listening comprehension intervention to evaluate whether a robot can utilize evidence-based practices in ABA to teach answering wh-questions. We have chosen a listening comprehension intervention because early childhood education on communicative speech skills has been shown to be one of the most important factors for positive long-term quality of life for individuals with ASD [26, 27]. Specifically, the ability to answer wh-questions is one of the most important communicative skills in early childhood education [28]. The robot-mediated intervention consists of a robot mediator reading passages from electronic books (e-books) and asking the child wh-questions with the aim of fading prompts until the child can independently as well as correctly answer wh-questions.

3.1 Humanoid Nao Robot

In this study we utilize the humanoid Nao robot, Figure 1, which stands 58cm in height with two degrees of freedom (DOF) in the head, five DOF in each arm, one DOF in each hand, and six DOF in each leg to enable the robot to produce physical movements. Nao can speak verbally through speakers. Furthermore, it has seven touch sensors, four bidirectional microphones, and two 2D cameras allowing the robot to sense the environment around itself. The robot can also be controlled remotely via a Wi-Fi connection.

3.2 E-Book

Our team has designed an e-book for the robot mediator to use during our interventions. It has been designed specifically for teaching answering wh-questions for children. The e-book is accessible via a web browser and has a page turning animation like a physical book. The pages can be turned by either pressing the left or right arrow keys on a keyboard or clicking on the e-book using a mouse. During the robot-mediated intervention, the robot turns the page on the e-book by sending a command that mimics pressing the left or right arrow keys on a keyboard. The motivation for this is to mimic as closely as possible the human-based intervention. On each page of the book is a picture of an animal with a sentence consisting of a main subject performing an action in a location. For example, Katie the cat was sitting on the couch. The illustration is a digital image of the main subject within a neutral background without any animations. An example page in the book is shown in Figure 2.

3.3 Intervention Design

The listening comprehension intervention, Figure 3, in this paper was designed by a Board-Certified Behavior Analyst to closely replicate existing ABA therapies and DTT teaching procedures implemented by behavior technicians in the clinic. In this intervention, the antecedent was the discriminative stimulus (S^D) of a robot reading a passage from the e-book and asking a wh-question. The behavior was a child's response to the wh-question. The consequence was then either: 1) a positive reinforcement via social praise and dancing from the robot for a correct response or 2) an error correction for an incorrect response from the child.

Each intervention session began with the robot-mediator introducing itself and notifying the child that they will be reading together during the session. The robot then read through the entire book (3 pages) while requesting the child to answer wh-questions after each

page. Namely, after the robot read over one page and displayed the illustration to the child he/she was asked a wh-question. Wh-questions included only who, what, and where questions. Questions were structured as follows: 1) "Who was (action)?", 2) "What is (noun) doing?", and 3) "Where did (noun perform an action)?". See Table 1 for the questions and correct responses used during the intervention. If the child answered the wh-question correctly, the robot would provide reinforcement in the form of both verbal praise (e.g. Great Job!) and a fun dance. When the child incorrectly responded to the robot's question, the robot verbally provided the answer to the question for the child to model and repeat. After three attempts at evoking a correct response from a child, the robot would move on to a new question.

During the intervention, prompts were also used when a child did not respond to a question within a certain period of time. These prompts were gradually faded until a child could independently answer the wh-questions. Prompts are commonly used in ABA therapy as supplementary stimuli to increase the probability of correct responses to an antecedent stimulus (e.g. wh-question) and structured to gradually transfer responses from prompts to the naturally occurring antecedent stimulus [29]. Example forms of prompt structuring include most-to-least prompts, graduated guidance, least-to-most prompts, and time delay prompts. In this intervention we used a most-to-least prompting protocol via time delay based on a child's prior performance to wh-questions. There are three levels of prompts before a child achieves skill independence/mastery. These levels include: 1) an immediate prompt (i.e. a 0-second delay immediately following the SD), 2) a 2-second delay before a prompt, or 3) no prompt. A child always begins at a level 1 prompt and the opportunity to independently respond gradually increases with each prompt level. The prompt level increases by one level if they answer 8 out of the 9 wh-questions correctly for two consecutive intervention sessions. A child has achieved mastery of the entire skill only when they can answer 8 out of the 9 wh-questions correctly for two consecutive intervention sessions with no prompts (i.e. during level 3).

The time-delays for the prompt levels and the way in which new prompt levels are introduced are well-established in standard ABA clinical practice as well as outlined by the Autism Curriculum Encyclopedia (ACE) from the New England Center for Children (NECC) [30]. The time-delays have been shown to provide children with ASD sufficient opportunity to respond correctly in time. Furthermore, the reason new levels are not gradually introduced (e.g. having some L2 prompts with L3 prompts in a session) is to reduce the probability of prompt dependency by fading away prompts as soon as possible [29]. Namely, prompt dependency refers to the phenomenon where an individual only responds correctly when a prompt is presented [31]. This form of time delay prompt fading has been shown in human delivered interventions to minimize the number of error responses and effectively as well as efficiently teach new skills to individuals with ASD [29].

4 Study Design

We conducted a study to investigate the efficacy of the robot-mediated intervention for teaching wh-questions to three children with ASD at a university-based ABA clinic. Furthermore, we also evaluated the children's engagement, communication, and affect during the sessions.

4.1 Participants

The inclusion criteria for these participants were: 1) 3-8 years old, 2) DSM-5 diagnosis for ASD, 3) currently receiving ABA therapy 4) have not mastered wh-questions, and 5) can follow 1-step gross motor instructions. Prior to initiating the intervention written informed consent was obtained from the participants' parental guardians and the director of the clinic evaluated whether the child already possessed the skill by using the Verbal Behavior Milestones Assessment Placement Program (VB-MAPP) [16]. VB-MAPP is a standard assessment tool used in ABA to specifically evaluate whether individuals with an ASD diagnosis possess a language or social skill. We also conducted a baseline session where the questions were presented to the participant by the robot without prompts or reinforcement. Participants were included in the study only if the director of the clinic indicated that the participant did not possess the skill according to the VB-MAPP assessment and he/she could not answer any of the questions during the baseline session.

A total of three children from the university-based ABA clinic met the inclusion criteria and participated in our study. Participant one was five years old and female. Participant two was four years old and male. Participant three was five years old and male. Each of the participants in this study were also evaluated prior to

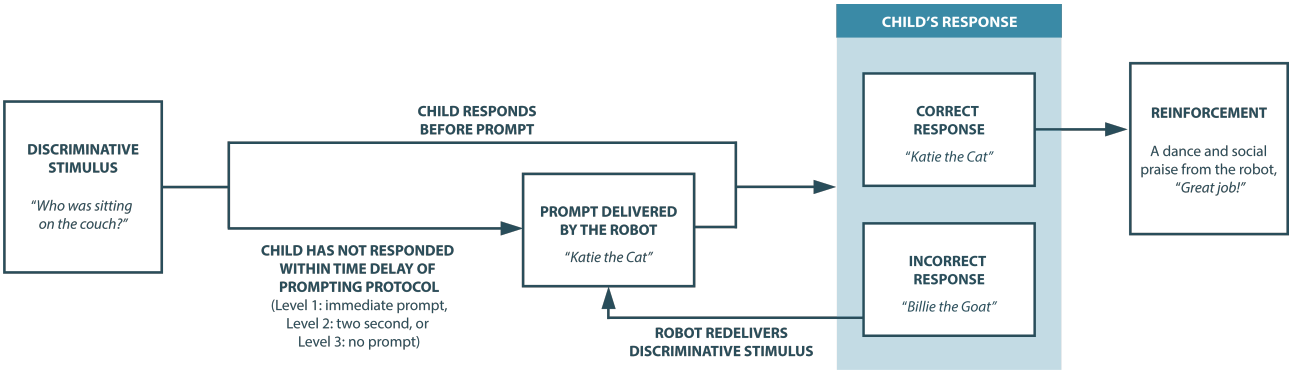

Fig. 3. Listening comprehension intervention session

Table 1. Wh-question bank used for the interventions

Question Type	Sample Question	Prompts/Correct Response
Sub-skill (who)	Who was sitting on the couch?	Katie the cat
	Who was chasing their tail?	Doug the dog
	Who was slithering on the ground?	Sam the snake
Sub-skill (what)	What is Katie the cat doing?	Sitting on the couch
	What is Doug the dog doing?	Chasing his tail
	What is Sam the snake doing?	Slithering on the ground
Sub-skill (where)	Where did Katie the cat sit?	On the couch
	Where did Doug the dog chase his tail?	By the pond
	Where did Sam the snake slither?	On the ground

Table 2. Vineland scores for study participants

Participant	1	2	3
Adaptive Behavior Composite Score	72	69	75
Communication Score	74	83	90
Socialization Score	68	48	74
Daily Living Skills Score	79	81	80

the study using the Vineland Adaptive Behavior Scales - Third Edition [17]. Vineland-3 is a standardized instrument used to measure adaptive behavior and support the diagnosis of intellectual and developmental disabilities such as ASD. This is the standard instrument used at the ABA clinic where we conducted our study. The scores for each participant are presented in Table 2. Namely, the Adaptive Behavior Composite (ABC) score is a composite score which evaluates an individual's ability in three adaptive behavior sub-domains: communication, socialization, and daily living. Communication refers to an individual's ability to listen, understand, and express him/herself through speech as well as read and write. Socialization refers to an individual's ability to function in social situations. Finally, daily living skills refers to an individual's ability to perform everyday tasks of daily living.

4.2 Intervention Setting

The intervention took place at a university-based ABA clinic within a private therapy room approximately 8' x 10' in size with carpeted floor, Figure 4. The room has a table and four chairs. The robot mediator sits or stands across from the child to conduct the intervention. A computer monitor on the table presents the e-book to the child and the robot mediator refers to the e-book throughout the intervention. A behavior technician was always present in the room to keep the child from physically interacting with the robot, to help keep the child seated in a chair in front of the robot, and to collect data. A video camera was also placed in the room to record the interaction for post-study analysis. During the robot-mediated interventions a researcher was controlling the actions of the robot during the intervention and collected performance data to ensure interobserver reliability. In order to successfully control the robot's actions during the intervention, the researcher controlling the robot was always able to see and hear the child.

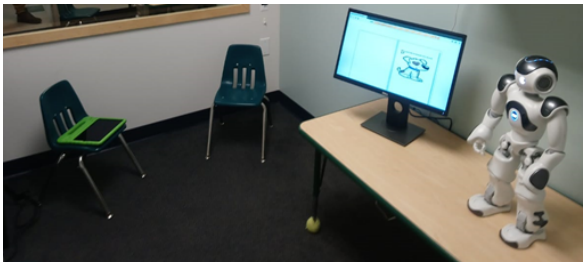

Fig. 4. Intervention setting

4.3 Robot Control Interface

The researcher controlling the robot utilized a customized Wizard-of-Oz GUI to set up the robot prior to the start of an intervention session and to control the actions of the robot during the intervention. Prior to the start of each session, the researcher chose the level of prompts to be utilized for the therapy session that day. The choice in the level was based on the performance of the child during his/her previous sessions. During an intervention session, the researcher utilized five push buttons to control the actions of the robot. These actions included: 1) greeting the child, 2) reading a page in the book and asking a wh-question, 3) reinforcing a correct response, 4) correcting an incorrect response, and 5) moving onto the next wh-question. Namely, the researcher used the greeting action (Button 1) to have the robot greet the child at the beginning of the session. The researcher would then control the robot to go through 27 wh-questions. Herein, each wh-question will be referred to as a trial. A trial consisted of the researcher controlling the robot to read a page in the book and ask the child a wh-question (Button 2). A time-delayed prompt would be automatically provided by the robot based on the prompt level selected by the researcher prior to the start of the intervention session. This ensured that the robot delivered the prompts at the correct time and did not rely on the researcher's reaction time. If the child responds correctly the researcher would control the robot to reinforce the child with both social praise and a dance (Button 3). In the case that a child responds correctly before the prompt, the reinforcement will prevent the prompt from occurring. If the child responds incorrectly the researcher would control the robot to correct the child and request the child to model and repeat the answer (Button 4). The researcher will move onto the next wh-question (Button 5) once the child has responded correctly or the robot cannot evoke a correct response from the child after three attempts. Button 5 selects at random without replacement the next wh-question and changes the phrases the robot will use for button 2, button 3, and button 4 to be within the context of the new wh-question.

4.4 Procedure

The intervention was conducted over three months with each child participating in an intervention session 3-5 times a week according to schedule availability and only once per a day. Each session was 30-45 minutes depending on the performance and preferences of the participants. Participants could request to end the session at any time. A typical session included presenting a maximum of 27 wh-questions (who, what, where) with short breaks after the presentation of 9 questions. Herein, we will refer to each set of 9 wh-questions as a sub-session. There were a total of 9 unique wh-questions used by the robot in this study. The exact questions are presented in Table 1. Each wh-question in the bank was presented 3 times during a session for the total of 27 questions.

4.5 Data Collection

Data was collected on the participants' performance on answering wh-questions during the robot-mediated intervention. A plus ("+") was recorded for a correct response without a prompt, a plus p ("+p") for a correct response after a prompt, a minus ("-") for an incorrect response without a prompt, and a minus p ("-p") for an incorrect response after a prompt. When a participant did not provide a response it was counted as an incorrect response after a prompt ("-p") at level 1 and 2. At level 3 and during mastery, a participant not responding was counted as an incorrect response without a prompt ("-").

We also conducted a post-study analysis of the video recorded sessions to investigate the children's engagement, communication, and affect during the robot-mediated interventions sessions. The following definitions were utilized by the coders:

Engagement was defined by the percentage of intervals that included the participant looking at any part of the robot or computer screen for at least 2 consecutive seconds during the interval. Namely, each session was discretized into 10-second intervals and the participants were classified as looking at the robot or computer screen if their gaze was directed towards the robot or computer screen at least once for 2 seconds during the interval. If the participant looked twice at the robot

Table 3. Interobserver agreement scores

Data	Average IOA Scores (Range)
Performance	**92% (87-96%)**
Engagement	**95% (83-100%)**
Communication	**95% (81-100%)**
Affect	**96% (81-100%)**

and/or computer screen within the interval it was still only counted once during that interval. The percentage of intervals engaged was then calculated by the number of intervals the participant was looking at the robot and/or computer screen divided by the total number of intervals in a session.

Communication was defined as the frequency during a session that a child either independently initiated verbal interaction with the robot (e.g. “Hi Robot!”) or engaged in verbal interaction with the robot as a part of the intervention.

Affect was defined as the percentage of intervals during a session when the participant demonstrated positive, negative, or neutral affect. A percentage of intervals was calculated for each of the three affective states. Each session was discretized into 10-second intervals and the participants affective state during the interval was categorized into positive, negative, and/or neutral. During an interval a participant could be categorized into more than one affective state (e.g. positive and negative). Hence, the affect percentages represent the percentage of the session where a participant was in a particular affective state. A positive affective state was identified by smiles, positive comments/movements to the robot, laughing, and touching/attempts at touching the robot. Negative affect was identified by crying, vocal protests (e.g. "No!"), attempts to leave, whining, frowning, or whimpering. Neutral affect was when neither positive nor negative affect was displayed by the participants behaviors.

The performance, engagement, communication, and affect data were coded by two independent behavior technicians to ensure interobserver agreement (IOA). IOA was scored using the trial-by-trial method which measures the agreement between the coders divided by the total number of trials. The IOA scores are presented in Table 3.

5 Results & Discussion

Figure 5, Table 4, and Figure 6, summarizes the results of this study. Namely, Figure 5 presents the participants’ percent of correct responses with and without prompts over the course of the intervention sessions. As previously mentioned, all participants first participated in a baseline session where they did not respond correctly to any of the wh-questions after being read a text. Table 4 presents the Pearson correlations between participants’ performance and their engagement, communication, and affect during the sessions. For significant correlations, we also present in Figure 6 participant behavioral data (i.e. engagement, communication, or affect) alongside their performance data during the sub-sessions.

5.1 Participant 1

Performance - In the initial intervention sessions, participant one was initially not responding to the robot. The director of the clinic suggested having the therapist they are typically paired with provide prompts along with the robot to familiarize the participant with the intervention and the robot because both were new experiences to her. Similar practices are held within the clinic when a child’s treatment is transferred to a new therapist he/she is unfamiliar with so that the child can become comfortable with the new individual as well as set the expectations of the interactions. The human prompts were removed once participant one answered 89% of the wh-questions correctly over two sessions with the assistance of the human prompts at level one. The child only moved onto level two once the human prompts were removed and the child answered 89% of the wh-questions over two sessions at level one with the robot. All prompting after level one came only from the robot.

Participant one required five sessions before the human prompts were removed and sixteen sessions total to complete prompt level one. She took two sessions to complete prompt level two. When level three was introduced, participant one responded correctly on only 3% of opportunities presented. As a result, level two was reintroduced according to the recommendations of the director of the clinic (i.e. the BCBA). Participant one required two sessions to complete prompt level two again so level three was reintroduced. After four sessions at level three, the intervention was discontinued due to lack of progress. During DTT, it is a common practice

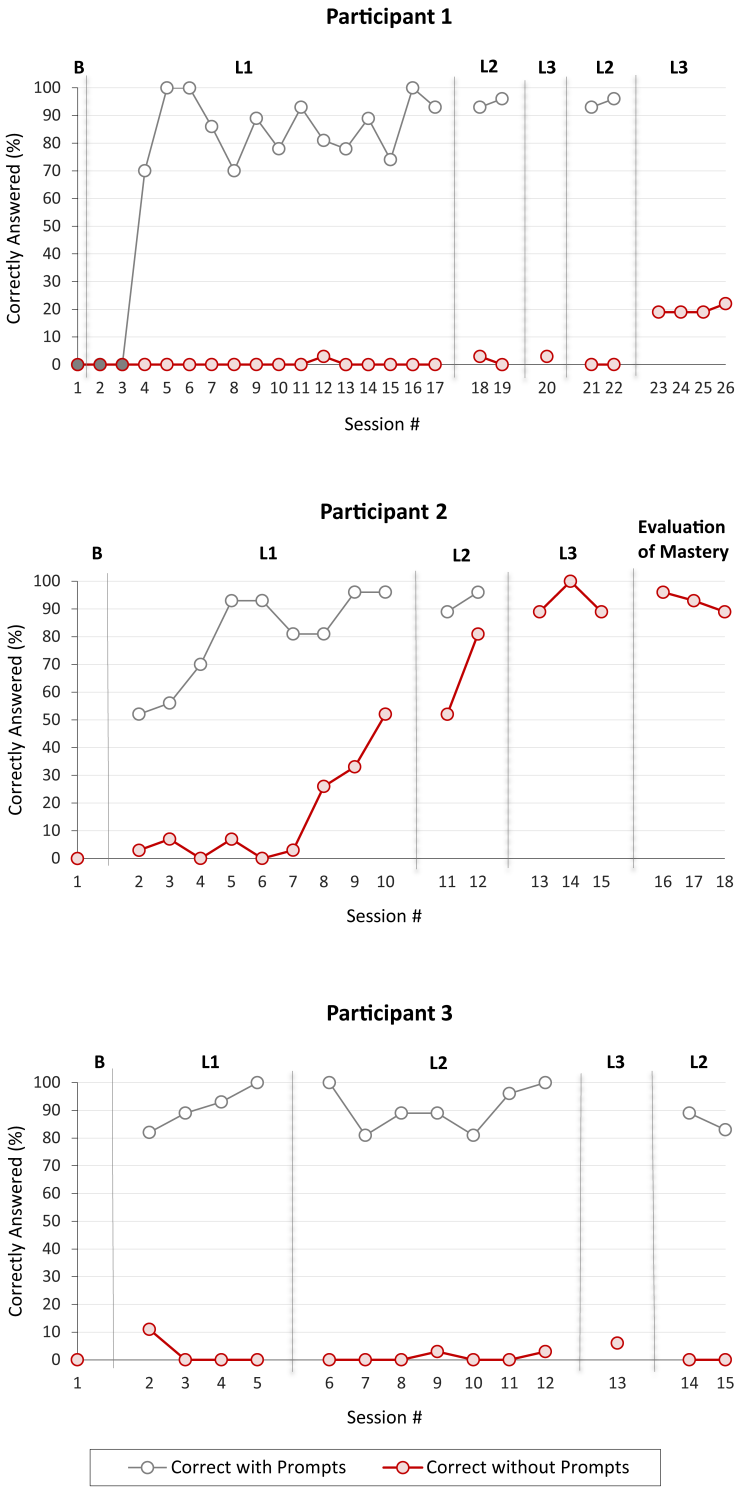

Fig. 5. Participant one, two, and three's performance during the intervention sessions. B indicates the baseline session. The sessions for each participant are also grouped according to the prompt level used by the robot during the sessions (L1: Level one; L2: Level two; L3: Level three).

Fig. 6. Participant one, two, and three's engagement, communication, and affect alongside their performance during the intervention sub-sessions. The sessions for each participant are also grouped according to the prompt level used by the robot during the sessions (L1: Level one; L2: Level two; L3: Level three). Note: Only significant correlations are presented.

Table 4. Pearson correlation coefficients between participant performance and their engagement, communication, and affect during the intervention

	P1		P2		P3	
	Correct With Prompts	Correct Without Prompts	Correct With Prompts	Correct Without Prompts	Correct With Prompts	Correct Without Prompts
Engagement	.295*	-.259*	.309*	-.257	-.930	.157
Communication	-.155	-.512*	-.691*	-.463*	.260	.045
Positive Affect	.085	-.103	.105	-.167	.490*	.106
Neutral Affect	.002	.104	-.084	.143	.100	-.105
Negative Affect	-.193	-.027	-.139	.184	.867*	.025
*Pearson correlation coefficient with $p<0.05$						

to evaluate current interventions to be able to effectively determine learner progress and potential programming changes [32].

For this participant, a customization to the protocol was made to improve responding during the intervention because initially participant one was not responding to the robot's immediate prompts. After session three, the behavior technicians that were familiar with participant one suggested that a human-like voice may improve the participant's response to the robot's prompts. Hence, a modification was made to instead use a pre-recorded familiar human voice on the robot. The results demonstrated that the child immediately started to respond correctly to prompts after the modification to deliver the prompts using a pre-recorded human voice. We hypothesize that participant one preferred a human voice to a synthetic voice because of familiarity with human voices and perceived it as a reinforcer. Similar preferences towards more human-like voices over synthetic voices have been observed with different populations of users [33]-[35].

As observed in the graph (Figure 5), with the modifications participant one responded correctly with prompts but was still unable to complete level three (i.e. answer wh-questions without prompts) or beat the prompt. Namely, the participant never provided a correct response before the prompt even when increasingly longer time delays were used before prompting. It is hypothesized that the prerequisite skills (i.e. rotating questions about an item) for this intervention were not met prior to the start of the intervention. To illustrate, participant one at the beginning of the study had not fully mastered the prerequisite skill of answering rotating questions (i.e. "What is it?", "What color is it?" and "What shape is it?") in relation to an item (e.g. tennis ball). Assessing prerequisite skills and identifying the appropriate skills to teach is a common challenge in treatment for ASD and not unique to only robot-mediated interventions [36].

Engagement, Communication, and Affect - On average, participant one was engaged during 90.5% of the sessions. Her engagement was negatively correlated with her ability to answer questions without prompts ($r = -0.259$, $p < 0.05$) and positively correlated with her ability to answer questions with prompts ($r = 0.295$, $p < 0.05$). The negative correlation between engagement and answering correctly without prompts suggests the participant was waiting for the robot to provide a prompt. With this participant, a time delay prompting protocol may have lead to prompt dependency because of the difficulty of the task for her [29].

She would also attempt to independently communicate verbally with the robot on average nine times during a session. Her communication with the robot was negatively correlated with her ability to answer questions without prompts ($r = -0.512$, $p < 0.05$). This is expected because when she is asked to respond without prompts her error rate increases and as a result is required to increase communication with the robot because of the error correction procedure of the intervention.

During the sessions participant one also displayed positive affect during 22.5% of the session, negative affect during 3.7% of the session, and neutral affect during 74.5% of the session. Her affect however did not have any significant correlation with her performance. Per the expert staff within the clinic, participant one has a history of being compliant irrespective of her affect during typical ABA instruction.

5.2 Participant 2

Performance - Similar to participant one, participant two initially required human prompts at level one. Par-

ticipant two required five sessions before the human prompts were removed and nine sessions total to complete prompt level one. Participant two then took three sessions to complete prompt level two and two sessions to complete level three. In total it took fourteen sessions before participant two achieved complete mastery of the wh-question answering skill related to the text designed for this study. Furthermore, we had three additional sessions with participant two that showed that he retained the mastery of the skill.

Engagement, Communication, and Affect - On average, participant two was engaged 89.2% during the sessions. His engagement was positively correlated with his ability to answer questions with prompts ($r = 0.309$, $p < 0.05$). He would also attempt to independently communicate verbally with the robot on average eighteen times during a session. His communication with the robot was negatively correlated with his ability to answer questions without prompts ($r = -0.463$, $p < 0.05$) and with prompts ($r = -0.691$, $p < 0.05$). For participant two, increases in communication were a result of the high rate of echoing many parts of the robot intervention and the participant reading the book out loud (e.g. "Katie the cat was sitting on the couch"). These types of communication interfered with his performance with the primary task (i.e. answering the robot questions) because they were distractors from the task. This was further supported because when participant two was engaged with the primary task he performed better with the prompts.

During the sessions participant two also displayed positive affect during 28.7% of the session, negative affect during 0.6% of the session, and neutral affect during 70.3% of the session. His affect however did not have any significant correlation with his performance. Due to appropriate selection of the intervention, participant two successfully went through the intervention which likely contributed to his high positive, high neutral, and low negative affect during the sessions.

5.3 Participant 3

Performance - Participant three required four sessions to complete prompt level one, and seven sessions to complete prompt level two. When level three was introduced participant three responded correctly on only 6% of opportunities presented so level three was discontinued and level two was reintroduced. Participant three required two sessions to complete prompt level two but the intervention was discontinued due to requests from the child to no longer participate in the study. Similar to participant one it was hypothesized that the participant did not possess the prerequisite skills before the start of the intervention. Hence, the task was hypothesized to be too difficult and frustrating for participant three. This was further emphasized by the higher negative affect during the intervention as compared to the other participants.

Engagement, Communication, and Affect - On average, participant three was engaged during 76.5% of the sessions. He would also attempt to independently verbally communicate with the robot on average ten times during a session. However, neither his engagement nor communication was significantly correlated with his performance. In contrast to participant one, engagement and communication for participant three occurred due to interest in the robot and related less to compliance with directions. Hence, the actions associated with participant three's engagement and communication were often unrelated to the particular task (e.g. pretending to feed the robot, statements and actions of affection).

During the sessions participant three also displayed positive affect during 46.3% of the session, negative affect during 9.4% of the session, and neutral affect during 44.6% of the session. His positive affect was positively correlated with his ability to answer questions correctly with prompts ($r = 0.490$, $p < 0.05$). His negative affect was also positively correlated with his ability to answer questions correctly with prompts ($r = 0.867$, $p < 0.05$). The staff suggested that participant three likely responded correctly to prompts when negative affect was displayed because the participant was aware that answering questions would lead to a break from the session. This was further emphasized by the child frequently requesting the sessions to be ended.

5.4 Summary of Interventions

To summarize only one of the three participants were able to reach mastery on the robot-mediated listening comprehension intervention. All participants responded correctly to prompts and enjoyed interacting with the robot, as evident by high occurrences of positive affect as well as engagement. For participants one and three, prompts were not faded completely due to the skill difficulty. The experts hypothesize that these participants did not have the prerequisite skills to master this type of intervention. The participants demonstrated before the intervention the ability to answer wh-questions but not rotating questions on a particular topic. Typical

programming for children with a diagnosis of ASD requires many modifications to reach mastery [37]. Due to the rigor of research, modifications were not appropriate and as such modifications could not be made to the robot-mediated intervention to have all the children reach mastery. In typical interventions, delivery modifications would have been implemented to help the children reach mastery. For future robot-mediated interventions, the research team will work with the expert to precisely select skills to target for intervention. Namely, the skills should be in line with the participants' current repertoire.

It is important to ensure skills taught by a robot to a child with ASD are transferable to human interactions and can be retained long-term. To validate the skill taught by the robot, we had a human therapist present the nine wh-questions from the listening comprehension intervention without prompts to participant two one year after the child interacted with the robot. We only validated participant two because he was the only child to reach mastery of the skill. Participant two was able to answer seven out of nine of the questions the first time. The human therapist re-presented the two incorrectly answered questions after asking two to three other questions and participant two could answer the questions correctly the second time. It is important to note when the therapist presented the questions the first or second time, he/she did not provide prompts, reinforcements, corrections, or any sort of feedback. These results demonstrate that the robot-taught skills were transferable to human interactions. Furthermore, this intervention was last presented to the child a year after he/she interacted with the robot. This suggests that robot-taught skills can be retained long-term by a child with ASD.

Engagement, communication, and affect results were unique to the individual participants. This is expected because ASD is idiosyncratic in how the behaviors manifest in each child [38]. This is further emphasized by our differing results from those found in [22] where the children's engagement, communication, and affect did not have any significant correlation with the children's performance during the social greetings interventions. While these parameters presented differently in each of the participants, the parameters could potentially be utilized as additional feedback for: 1) the robot to autonomously personalize its behaviors during an intervention or 2) engineers and therapists to personalize the design of the intervention to improve the child's performance. Hence, in the future it could be beneficial to develop approaches which identify what human-robot interaction parameters correlate with a child's performance and focus on intentionally influencing these parameters.

5.5 Summary of Robot Intervention Behaviors

The robot-mediated listening comprehension intervention presented in this work focused on having a SAR replicate the core behaviors of human therapists during DTT-based therapies with children with ASD. These behaviors included the delivery of the: discriminative stimulus, error corrections, prompts, and reinforcements. During the intervention, the discriminative stimulus was used to evoke a response from the child. Error corrections, prompts, and reinforcements were all then used to increase the probability of a child correctly responding to a discriminative stimulus in the future. Although all the children improved on their ability to answer wh-questions, only one child reached mastery during this study.

Discussions with the clinicians suggested that the main areas for improvement in our robot-mediated interventions included moving beyond the core therapist DTT behaviors and developing a larger range of reinforcements, high-probability requests [3], and pairing procedures [39] for the robot. Namely, a current limitation with the existing intervention is that the robot only has five phrases it uses for social praise which are paired with dances. Despite the participants initially enjoying these reinforcers during the interventions, we noticed they became monotonous over longer periods of time. In the future interventions could be improved by increasing the variety of reinforcers delivered by the robot such as including reinforcers typically used by human therapists such as: edibles, preferred items (e.g. stickers), or other forms of social reinforcers (e.g. jokes, praise, and songs). A robot should also be capable of either adapting the type of reinforcement delivered according to a child's performance or inquiring from them directly what may be their preferred reinforcer. Furthermore, it should be noted that the goal is to have the children respond to social reinforcers (e.g. "nice job") and limit other types of reinforcement because social reinforcement is what naturally occurs in human-human interactions.

As previously mentioned, at the beginning of the interventions participant one and two were not responding to the robot during the intervention. The clinicians hypothesized there were two potential explanations: 1) the participants were unfamiliar with how to interact or

respond to the robot during an intervention; and 2) the robot was new to the participants so they had yet developed a relationship with it. To address these two issues the director of the clinic suggested having the therapist they are typically paired with provide prompts along with the robot to familiarize the participants with the intervention and the robot because both were new experiences to the participants. While we weren't able to systematically evaluate this process we observed that initial human prompts helped the children understand how to interact with the robot. The human prompts were then faded away to enable the robot to independently deliver the entire intervention.

To improve the robot-mediated interventions, the robot should have behaviors to independently manage the process when it is first introduced to the children. The clinicians have suggested that the robot could deliver high-probability requests to familiarize the children with how to interact and respond to the robot. High-probability requests in ABA refers to making requests that are already easy for an individual prior to making requests that may be initially difficult for an individual such as a new skill they are learning [3]. Hence, a robot teaching a new skill to a child should have a set of requests that the child has already mastered so that they can be used during the initial interactions with a child to build a child's experience with interacting with the robot. Furthermore, the high-probability requests can also be used to increase the motivation of children during intervention sessions they are struggling with by interspersing the requests a child has mastered with the skills they are learning during the intervention.

The clinicians also recommended the robot have a pairing procedure to build a positive relationship with the children because having the robot immediately deliver an intervention may be too large of a step for the robot. Pairing procedures refer to building rapport with a child and offering reinforcers (e.g. preferred items, edibles, and social reinforcers) without making requests or demands [39]. Such a process builds the association for a child that the robot is fun and will lead to more fun elements in the future. Pairing procedures are also commonly used by human ABA therapists in their first interactions with children so that they can build a positive relationship with the children [39].

6 Conclusions

Overall, the participants all improved in their capability to answer wh-questions over the sessions with prompts from the robot but only one participant achieved mastery to date. Such differences in success are expected because children with ASD often require customized intervention plans due to their differences in capabilities and preferences.

With the insights gained from this study, we intend on modifying our intervention to address the aforementioned challenges that have arisen during our study. Namely, when selecting interventions to be delivered by the robot, it is imperative to identify targets at the participants' skill level. One such method to identify these targets is via tools such as the Verbal Behavior – Milestones Assessment and Placement Program (VB-MAPP) supporting skills protocol [16]. The present study also has the robot only utilize the core ABA techniques for implementing DTT. However, for future studies we plan to explore motivational ABA techniques utilized during human facilitated DTT sessions. These motivational techniques could include: high-probability requests (e.g. interspersing mastered tasks with acquisition tasks), pairing procedures, and varying the robot reinforcers. We will also explore a broader set of socially valid targets for the robot to deliver efficient and effective ABA-based interventions. Finally, we also intend on investigating changes in children's behavior during long-term interactions with the SAR and changes in performance as well as behavior over multiple different interventions.

7 Acknowledgements

This work was supported by the National Science Foundation CRII Award (#1948224). We would also like to thank Foong Ling Chen, Rachel Prescott, Julie Kopp, and Rachel Lai all the staff at the Oakland University Applied Behavior Analysis clinic for their help with the research.

References

[1] Centers for Disease Control and Prevention, Key Findings From The ADDM Network, 2016. Accessed on: July 10, 2020. [Online]. Available at: https://www.cdc.gov/ncbddd/autism/addm-community-

report/key-findings.html

[2] B. Zablotsky, L. I. Black, M. J. Maenner, L. A. Schieve, and S. J. Blumberg, "Estimated prevalence of autism and other developmental disabilities following questionnaire changes in the 2014 national health interview survey", National Health Statistics Report, vol. 87, pp. 1-19, 2015.

[3] J. O. Cooper, T. E. Heron, and W. L. Heward, Applied behavior analysis, New York: Pearson, 2007.

[4] Y. X. Zhang and J. R. Cummings, "Supply of certified applied behavior analysts in the united states: Implications for service delivery for children with autism", Psychiatric Services, pp. 1-4, 2019.

[5] B. Robins, K. Dautenhahn, R. Te Boekhorst, and A. Billard, "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, vol. 4, pp. 105-120, 2005.

[6] J. Greczek, E. Kaszubski, A. Atrash, and M. Mataric, "Graded cueing feedback in robot-mediated imitation practice for children with autism spectrum disorders", Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication, Edinburgh, Scotland, pp. 561-566, 2014.

[7] E. Bekele, J. A. Crittendon, A. Swanson, N. Sarkar, and Z. E. Warren, "Pilot clinical application of an adaptive robotic system for young children with autism", Autism, vol. 18, no. 5, pp. 598-608, 2014.

[8] C. A. Pop et al., "Social robots vs. computer display: Does the way social stories are delivered make a difference for their effectiveness on ASD children?", Journal of Educational Computing Research, vol. 49, no. 3, pp. 381-401, 2013.

[9] D. Feil-Seifer and M. Mataric, "Robot-assisted therapy for children with autism spectrum disorders", Proceedings of International Conference on Interaction Design and Children, Chicago, Illinois, pp. 49-52, 2008.

[10] K. Jordan, S. Hellersteth, and H. Mulligan, "Robotic Technology for Teaching Adolescents with Autism Spectrum Disorder: A Feasibility Study", Proceedings of International Conference on Rehabilitation Engineering & Assistive Technology, Tampines, Singapore, pp. 1-4, 2012.

[11] H. Andreae, P. Andreae, J. Low, and D. Brown, "A study of auti: A socially assistive robotic toy", Proceedings of International Conference on Interaction Design and Children, Aarhus, Denmark, pp. 245-248, 2014.

[12] B. A. English, A. Coates, and A. Howard, "Recognition of gestural behaviors expressed by humanoid robotic platforms for teaching affect recognition to children with autism - A healthy subjects pilot study, Proceedings of International Conference on Social Robotics", Tsukuba, Japan, pp. 567-576, 2017.

[13] G. Pioggia, R. Igliozzi, M. Ferro, A. Ahluwalia, F. Muratori, and D. De Rossi, "An android for enhancing social skills and emotion recognition in people with autism", IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 13, no. 4, pp. 507-515, 2005.

[14] M. Begum, R. W. Serna, and H. A. Yanco, "Are robots ready to deliver autism interventions? A comprehensive review", International Journal of Social Robotics, vol. 8, no. 2, pp. 157-181, 2016.

[15] Behavior Analyst Certification Board Inc., Applied Behavior Analysis Treatment of Autism Spectrum Disorder: Practice Guidelines for Healthcare Funders and Managers, 2014. Accessed on: July 29, 2020. [Online]. Available at: https://casproviders.org/asd-guidelines/

[16] M. L. Sundberg, VB-MAPP: Verbal Behavior Milestones Assessment and Placement Program, California: AVB Press, 2008.

[17] S. S. Sparrow, D. V. Cicchetti, and C. A. Saulnier, Vineland Adaptive Behavior Scales–Third Edition, Texas: Pearson 2016.

[18] T. Smith, "Discrete Trial Training in the Treatment of Autism, Focus on Autism and Other Developmental Disabilities", vol. 16, no. 2, pp. 86-92, 2001.

[19] R. L. Koegel and L. K. Koegel, "The PRT Pocket guide: Pivotal response treatment for Autism Spectrum Disorders", Maryland: Brookes Publishing, 2012.

[20] M. L. Barbera, The verbal behavior approach: How to teach children with autism and related disorders, London: Jessica Kingsley Publishers, 2007.

[21] S. S. Yun, H. Kim, J. Choi, and S. K. Park, "A robot-assisted behavioral intervention system for children with autism spectrum disorders", Robotics and Autonomous Systems, vol. 76, pp. 58-67, 2016.

[22] M. Begum, R. W. Serna, D. Kontak, J. Allspaw, J. Kuczynski, and H. A. Yanco, "Measuring the efficacy of robots in autism therapy: How informative are standard HRI metrics?", Proceedings in ACM/IEEE International Conference on Human-Robot Interaction, Portland, United States, pp. 335-342, 2015.

[23] Y. Feng, Q. Jia, and W. Wei, "A control architecture of robot-assisted intervention for children with autism spectrum disorders", Journal of Robotics, pp. 1-12, 2018.

[24] M. Salvador, A. S. Marsh, A. Gutierrez, and M. H. Mahoor, "Development of an ABA autism intervention delivered by a humanoid robot", Proceeding in International Conference on Social Robotics, Kansas City, United States, pp. 551-560, 2016.

[25] Z. Warren, Z. Zheng, S. Das, E. M. Young, A. Swanson, A. Weitlauf, and N. Sarkar, "Brief Report: Development of a Robotic Intervention Platform for Young Children with ASD", Journal of Autism and Developmental Disorders, vol. 45, pp. 3870-3876, 2015.

[26] C. Gillberg and S. Steffenburg, "Outcome and prognostic factors in infantile autism and similar conditions: A population-based study of 46 cases followed through puberty", Journal of Autism and Developmental Disorders, vol. 17, no. 2, pp. 273–287, 1987.

[27] A. Levy and A. Perry, "Outcomes in adolescents and adults with autism: A review of the literature", Research in Autism Spectrum Disorders, vol. 5, no. 4, pp. 1271-1282, 2011.

[28] E. J. Sanders and K. A. Erickson, "Wh-Question answering in children with intellectual disability", vol. 76, pp. 79-90, 2018.

[29] G. S. MacDuff, P. J. Krantz, and L. E. McClannahan, "Prompts and prompt-fading strategies for people with autism", in Making a Difference: Behavioral Intervention for Autism, C. Maurice, G. Green, & R. M. Foxx, Ed. Texas: PRO-ED, 2001, pp. 37-50.

[30] Autism Curriculum Encyclopedia (ACE) from the New England Center for Children (NECC), 2020. Accessed on: July 10, 2020. [Online]. Available at: https://www.acenecc.org/about-the-ace/

[31] K. M. Clark and G. Green, "Comparison of two procedures for teaching dictated-word/symbol relations to learners with autism", Journal of Applied Behavior Analysis, vol. 37, no. 4, pp. 503-507, 2004.

[32] D. C. Lerman, A. L. Valentino, and L. A. LeBlanc, Discrete trial training, in Early Intervention for Young Children with Autism Spectrum Disorder, R. Lang, T. Hancock, N. Singh, Ed. Switzerland: Springer International Publishing, 2016, pp. 47-84.

[33] F. Eyssel, S. Bobinger, and L. De Ruiter, "'If you sound like me, you must be more human': On the interplay of robot and user features on human-robot acceptance and anthropomorphism", Proceedings of ACM/IEEE International Conference on Human-Robot Interaction, Massachusetts, United States, pp. 125-126, 2012.

[34] R. Tamagawa, C. I. Watson, I. H. Kuo, B. A. Macdonald, and E. Broadbent, "The effects of synthesized voice accents on user perceptions of robots", International Journal of Social Robotics, vol. 3, no. 3, pp. 253-262, 2011.

[35] A. Baird, E. Parada-Cabaleiro, S. Hantke, F. Burkhardt, N. Cummins, and B. Schuller, "The perception and analysis of the likeability and human likeness of synthesized speech", Proceedings of 19th Annual Conference of the International Speech Communication Association, Hyderabad, India, pp. 2863-2867, 2018.

[36] A. Persicke et al., "Evaluation of the concurrent validity of a skills assessment for autism treatment", Research in Autism Spectrum Disorders, vol. 8, no. 3, pp. 281-285, 2014.

[37] C. Shepley and J. Grisham-Brown, "Applied behavior analysis in early childhood education: An overview of policies, research, blended practices, and the curriculum framework", Behavior Analysis in Practice, vol. 12, no. 1, pp. 235-246, 2019.

[38] R. L. Koegel and L. K. Koegel, Teaching children with autism, Maryland: Brookes Publishing, 1995.

[39] B. Esch, J. E. Carr, and L. L. Grow, "Evaluation of an enhanced stimulus-stimulus pairing procedure to increase early vocalizations of children with autism", Journal of Applied Behavior Analysis, vol. 42, no. 2, pp. 225-451, 2009.