Robot-mediated Group Instruction for Children with ASD: A Pilot Study

Madeline Trombly*[1], Pourya Shahverdi*[1], Nathan Huang*[1], Qinghua Chen*[1], Jessica Korneder[2]
Wing-Yue Geoffrey Louie[1], *IEEE Member*

Abstract— **Children diagnosed with autism spectrum disorder (ASD) typically work towards acquiring skills to participate in a regular classroom setting such as attending and appropriately responding to an instructor's requests. Social robots have the potential to support children with ASD in learning group-interaction skills. However, the majority of studies that target children with ASD's interactions with social robots have been limited to one-on-one interactions. Group interaction sessions present unique challenges such as the unpredictable behaviors of the other children participating in the group intervention session and shared attention from the instructor. We present the design of a robot-mediated group interaction intervention for children with ASD to enable them to practice the skills required to participate in a classroom. We also present a study investigating differences in children's learning behaviors during robot-led and human-led group interventions over multiple intervention sessions. Results of this study suggests that children with ASD's learning behaviors are similar during human and robot instruction. Furthermore, preliminary results of this study suggest that a novelty effect was not observed when children interacted with the robot over multiple sessions.**

I. Introduction

Social robots have the potential to become a prevalent part of everyday society, with uses ranging from healthcare to education. Many current studies have explored the implementation of a social robot within the classroom setting [1]–[3]. Research exploring the use of social robots in secondary education and beyond have largely focused on the students' perceptions of the robot's effectiveness, credibility, and ability to successfully integrate into the learning space as a teacher and/or a teaching aide [4]–[6]. Whereas research conducted in primary school and early childhood education classrooms have attained a broader collection of data that focuses more on actual behaviors that indicate children's ability to learn from or alongside a robot [7]–[10]. Research on using social robots for childhood education has specifically focused on the behaviors and learning of children with autism spectrum disorder (ASD). This is because children with ASD may have deficits regarding human social interactions and may be more comfortable interacting with a robot to initially practice social behavior [11]. Therefore, it is an open research question whether an ASD therapy can be considered more effective if delivered by a social robot rather than by a human.

The majority of studies that observe children with ASD's interactions with social robots have been limited to one-on-one interactions where a child is the only one to directly engage with the robot [7]–[10]. One-on-one interactions present a child with limited choice as to where they direct their focus; it is either towards the teaching agent or to any other objects in the room. There are limited spontaneous distractions to occupy the child's attention. Hence, these one-on-one interventions do not teach children with ASD the vital group interaction skills they require to participate in traditional classroom settings. Namely, it is more common in school settings for instructors to deliver group instructions to maximize their ability to teach many individuals simultaneously [12].

Group interactions and classroom settings are unique because they may impact the learning behaviors of children [13]. In a group setting, a child has to contend with all of the potential interactions of a one-on-one session as well as the unpredictable behaviors of the other children in the session. The inconsistency of the accompanying children's behaviors during a group instruction session also has the potential to affect children's learning. Whereas children participating in a one-on-one instruction session would not face the same distractions. As a result, group or classroom behaviors typically need to be taught both explicitly and in practice for children with ASD. Namely, children with ASD often work towards developing skills necessary for learning in a classroom environment such as attending and appropriately responding to an instructor's requests [14].

This research aims to investigate the efficacy of a robot-mediated group interaction intervention for children with ASD. This addresses an existing gap in the literature as current studies primarily focus on one-on-one interactions and do not investigate robot interactions with groups of children with ASD. Towards this goal, we have been conducting a long-term human-robot interaction (HRI) study to investigate the learning behaviors of children with ASD during robot-mediated group instruction.

In this paper, we present preliminary results of this HRI study which compares robot facilitated and human facilitated group instruction sessions with children with ASD. During these sessions we measured children's learning behaviors over multiple group instruction sessions. The human-led and robot-led conditions were then compared to identify whether there were differences between the conditions as well as to determine if the conditions lead to maintained or improved learning behaviors over time.

II. Literature Review

Current applications of social robots focus on one-on-one interactions between a robot and a child. To date, there have

This work was supported by the National Science Foundation grant #1948224

[1]Intelligent Robotics Laboratory, Oakland University, Michigan, USA (e-mail: louie@oakland.edu)

[2]Applied Behavior Analysis Clinic, Oakland University, Michigan, USA

* Co-first authors that contributed equally to this work

been a wide range of different activities and delivery methods utilized during these one-on-one interactions. Robot and child interactions have focused on teaching a new skill, promoting the retention of skills and/or the maintenance of a child's engagement within the interaction [7]–[10].

In [10], a study was conducted to explore the impact of the NAO humanoid robot displaying adaptive emotional feedback on the vocabulary retention and social engagement of children during one-on-one vocabulary teaching sessions. Over the course of two weeks, twenty-four children participated in four one-on-one vocabulary teaching sessions with the robot. During a session, the robot introduced itself, conducted a vocabulary pre-test, played a "Snakes and Ladders" game with the child while interspersing instances of vocabulary exercises, and concluded with a vocabulary post-test. An emotion and memory model was also implemented to allow the robot to learn the children's personalities and provide appropriate emotional feedback during the sessions. Overall, results of the study with the children demonstrated that the emotional and memory model supported maintenance of children's social engagement and improved their vocabulary retention.

In [9], a study was conducted with children diagnosed with ASD to investigate how children's attention and engagement change over time during child-robot game interactions. These game interactions included playing cards, building structures, or completing a puzzle. Six children with ASD participated in twenty one-on-one child-robot game sessions over six months as a part of their Pivotal Response Treatment based therapy. To assess the quality of each interaction, data on the child's behavioral cues and affect were collected to determine the child's duration of attention and engagement during a game session. The study results demonstrated no significant change in the attention and engagement that the children displayed towards the robot in the game but a notable increase in the attention and engagement that the child directed towards their parent present in the room.

In [8], children with ASD's engagement over time was investigated during individualized robot therapy. Individualized goals such as joint attention, imitation, and turn-taking were targeted based off individual need. Eleven children diagnosed with ASD participated in the study and took part in seven to ten 15-minute sessions over six months. Sessions consisted of a variety of activities common to ASD therapy. During each session, children interacted with a Wizard-of-Oz controlled NAO robot that facilitated the therapy activities. Results demonstrated that engagement with the robot across multiple sessions could be maintained and children's engagement was significantly improved during therapy activities.

In [7], a study was conducted to investigate the social attention and verbalization patterns of children during a robot-mediated imitation game. Fifteen typically-developing children and two children with ASD participated in eight 30-minute imitation game sessions with the Isobot humanoid robot over the course of six weeks. Analysis of the data indicated that typically-developing children and children with ASD both directed their attention towards the robot during a session more often than they did elsewhere, and both engaged in spontaneous communication with the trainer. Results also showed that children became less interested in the robot over time due to their increasing familiarity with the activities.

Fig. 1: (a) Robot instructor condition (b) Human instructor condition

Current research involving interactions between social robots and children diagnosed with ASD in an instructional setting generally focuses on one-on-one child-robot interactions. These studies tend to investigate the children's attention [7]–[10] and acquisition as well as retention of skills [10] over time. To the best of our knowledge, there has been minimal exploration on the learning behaviors of children with ASD when interacting with a social robot within a group setting. There has also been some discrepancies in one-on-one studies as to how a child's familiarity with the robot and activities impacts their attention, with some claiming that familiarity improves attention [8] and others asserting that familiarity negatively impacts attention [7]. The study presented in this paper aims to investigate the effects of robot-led group interactions on the attention and learning of children with ASD as well as these effects over time.

III. Group Interaction Intervention

We developed a robot-facilitated group interaction intervention to enable children with ASD to practice the skills required to participate in a group, Figure 1(a). We chose a group interaction intervention because group social interaction is a common challenge for children with ASD and they typically work towards gaining skills that are necessary for participation in a regular classroom setting such as attending and appropriately responding to an instructor's requests [14]. The group interaction intervention, herein called "Circle Time", was designed to replicate the traditional human-led sessions held at the university-based applied behavior analysis (ABA) autism clinic. The group instruction sessions adhered to existing ABA protocols and teaching methods typically utilized by the clinic's therapists during therapies with children with ASD. All protocols used during the session were verified by a board certified behavior analyst.

A. Intervention Setting

Circle time intervention sessions took place in a university-based ABA autism clinic within a classroom approximately 25' x 30'. The classroom has a projector located on the ceiling and projects onto a screen found at the front of the room. A podium is next to the screen at the front of the room and the remaining portion of the room has open floor space.

The Pepper humanoid robot instructor stood at the front of the room under the screen, facing away from the screen and towards the group of children. Three to five children per session were seated on round cushions in a semi-circle formation in the open space of the room and faced the instructor. Each child was about six feet away from the robot. Following circle time procedures at the clinic, children's therapists sat behind the children (one for each child) and at a sufficient distance (1ft) to minimize distractions during the group intervention. Researchers were located at back of the room to observe the session and control the robot remotely.

B. Intervention Design

A single robot-led intervention session consists of learning opportunities targeting group interaction skills interspersed with fun group activities. This is a common approach for organizing an ABA-based therapy for teaching children with ASD group interaction skills that includes following group instructions, following individual instructions directed toward a single child in the group, and refraining from following individual instructions that are directed towards another child. The group activities included taking attendance, a sing/dance along, a "Find the Object" game, a "Copy Me" game, a "Labeling Animal" game, "Yoga", and a clean-up activity. The descriptions for each group activity as well as the group interaction skill teaching activities are described below.

1) Attendance Activity: At the beginning of every session, the robot instructor introduced itself by saying, "Hi kids! My name is Pepper and I will be your Circle Time leader today. Are you ready to get started?" The robot then began taking attendance immediately after its introduction by saying, "First let's see who is here today." The robot would then take attendance by addressing each child by name and asking them to perform a one-step instruction (e.g., "clap your hands," "touch your nose," and "raise your arms").

2) Sing/Dance Along: During a sing/dance along activity, the children were asked by the robot to perform an action (e.g., "clap your hands") to indicate their desire to choose a song. The robot selected a child from those that performed the action to come to the front of the room and select a song on the screen. Six songs were displayed on the screen at a time. Children selected a song by picking up a long pointer stick located on the floor next to the robot and used it to indicate their choice. After selecting the song they returned the pointer stick to its place on the floor. The song that the child pointed to was then played.

3) "Find the Object" Game: The "Find the Object" game consisted of a variety of objects (e.g., colors, animals, shapes) scattered in the open space of the room behind the children. During the game, an image of a single object was shown on the projector screen while the robot requested that the children find the indicated object and drop it in the bin located on the floor next to the robot. This game was repeated for three different objects.

4) "Copy Me" Game: The robot led the children through the "Copy Me" game by demonstrating an action (e.g., guitar, golf, gorilla) and requesting the children to replicate the action. Children were allowed several seconds to complete the action. After all or the majority of the children attempted to replicate the desired action, the robot returned to a neutral position and praised the children (e.g.,"You are doing great!"). Four different movements were completed in the game.

5) "Labeling Animals" Game: The "Labeling Animals" game consisted of projecting an image of an animal (e.g., panda, giraffe, lion) on the projector screen, with the robot instructor pointing to the image of the animal and asking the group, "What animal is this?" The children then responded by calling out the name of the animal. If none of the children responded after a couple seconds, the robot would then repeat the question. If the majority of the children correctly labeled the animal, the robot instructor praised the group as a whole, saying, "You're right, it is a [animal]!" The game consisted of the children labeling a total of four different animals.

6) Yoga Activity: The robot instructor led the group of children in a yoga activity which consisted of a yoga pose (e.g., cat, tree, chair) being projected on the screen and the instructor saying, "Try the [pose] with me!" or "Let's try the [pose] together!" The instructor would then perform the pose and a corresponding sound or music would play while the robot instructor remained in the pose for several seconds. After the robot returned to a standing position, the screen displayed a new pose and the entire process was repeated. The activity consisted of three different yoga poses.

7) Clean-up Activity: Every session then ended with a clean-up activity. The robot instructor indicated the end of the session by saying, "Thank you all for playing with me today, I had so much fun! Please pick up your cushions and put them in the bin, it's time for snack!" After the children had placed their cushions in the bin and lined up, their therapists led them out of the room and to the next activity of their regularly scheduled day at the clinic.

8) Group Interaction Skills Activities: Individual and group instructions were interspersed throughout the intervention to teach children to: 1) follow individual instructions that were targeted towards them from an instructor, 2) follow group instructions together with other children, and 3) patiently wait when individual instructions were not targeted towards them. The individual instructions consisted of the robot instructor asking any single child within the group to complete a one-step instruction such as "clap your hands". The group instructions consisted of the robot requesting the entire group to complete an action such as "Everyone clap your hands". An intervention session consisted of each child in the group presented with three individual instructions and three group instructions. The total number of times a child had to patiently wait when other children were receiving instructions was dependent on the total number of children participating in the sessions and can be calculated by $3 \times (n-1)$ where n is the number of children participating in the session. Depending on the children's schedule, there were three to five students at each session.

C. Robot Control Interface

A customized Wizard-of-Oz graphical user interface (GUI), Figure 2, was used to operate the robot during intervention sessions. In the GUI, the researcher controlling the robot could navigate to a page specific to each activity or game by clicking on its tab at the top of the page. Each page

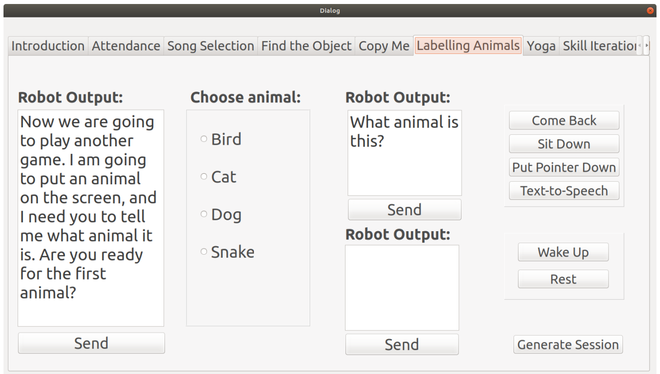

Fig. 2: Customized GUI.

has discrete robot behaviors related to each activity that the operator could use to conduct the activity with the robot. There were also buttons lining the bottom of the page that were available regardless of the selected tab to address any problem behavior that occurred during a session since problem behaviors can be common for individuals with ASD. These buttons included controls that allowed the robot to ask the participants to ”Sit down”, ”Come back to the circle”, and ”Put the pointer back on the ground”. The objects, animals, movements, and yoga poses of activities could be randomized by using the ”Generate Session” button.

IV. Human-Robot Interaction Study

A human-robot interaction study was conducted to compare children with ASD’s learning behaviors (i.e., engagement, communication, affect, and performance) during group-based interventions led by a human instructor and by a robot instructor as well as the effect of time on learning behaviors for each instructor type. Our hypotheses were:

- **H1** There will be no difference between children’s learning behaviors in the first and third human-led session.
- **H2** Children’s learning behaviors will be higher in the first robot-led session than the third robot-led session.
- **H3** Children’s learning behaviors will be higher in the first robot-led session than the first human-led session.
- **H4** There will be no difference between children’s learning behaviors in the third human-led session and the third robot-led session.
- **H5** There will be an interaction effect between session number and instructor type on children’s learning behavior.

In order to investigate these hypotheses, we conducted a 2x2 repeated measures experimental design where the two independent variables were type of instructor and time. Both independent variables were manipulated at two levels. Namely, the instructor variable could either be robot or human and the time variable could be the first intervention session or third intervention session. Overall, there were four conditions: human first session, human third session, robot first session, and robot third session. The learning behaviors we measured included children’s engagement, communication, affect, and performance in the different conditions. The rationale for a within-subject study is to reduce the variability between subjects because ASD is a spectrum with wide variation in the type and severity of symptoms experienced by individuals [15]. The procedures and protocols for this study were reviewed and approved by the Institutional Review Board at Oakland university (#IRB-FY2021-376).

A. Participants

Participants were recruited from a university-based ABA autism clinic. The criteria for inclusion in the study were children: 1) 3-8 years old, 2) diagnosed with ASD based on the Diagnostic and Statistical Manual of Mental Disorders (DSM-V) with severity’s ranging from one (i.e., requiring support) to three (i.e., requiring very substantial support) [16], and 3) already receiving ABA therapy prior to the start of this study. A total of nine children with ASD were recruited. Written informed consent was obtained from all participants’ parental guardians prior to the commencement of the study and participants’ parental guardians were informed that their children could withdraw from the study at any time.

B. Procedure

Each Circle Time session lasted about 15-20 minutes with 3-5 children present at each session. Children were assigned to Circle Time sessions based on the clinic’s pre-existing Circle Time schedule. Due to the differences in children’s schedules (i.e., number of therapy sessions per week) human and robot sessions were alternated based on the participant with the least number of robot or human sessions. In the present study, only the first three human and robot sessions for each participant was used for data analysis.

Human and robot sessions were made as similar as possible. Both utilized the same slideshow, which was controlled remotely by a researcher at the back of the room during both conditions. Human instructors followed a script to ensure their speech and actions were consistent with those of the robot. The only difference between the instructor conditions occurred during the Sing/Dance Along Activities, where the human instructor would move to stand and dance at the side of the screen so as to not obstruct the children’s view, while the robot was small enough that it would not obstruct their view of the screen and therefore remained stationary.

C. Data Collection

Both human-led and robot-led sessions were video recorded with three video cameras and coded to measure children’s learning behaviors including engagement, communication, affect, and performance.

1) Engagement: Engagement was measured by observing the child’s eye gaze and head orientation. There were three categories of engagement that were defined by where the child directed their gaze: 1) instructor/screen, 2) therapist/peers, or 3) off-target. The session was divided into ten-second intervals and each interval was observed to see where the child directed their gaze. If they directed it towards either of the first two categories for two consecutive seconds, the interval was classified as that category. An interval could be classified as engaged with both categories. The interval was categorized as off-target if the participant did not hold their eye gaze on either category for two consecutive seconds within the interval. The percentage of each engagement category during a single session was then calculated by dividing the number of intervals classified as that category by the total number of intervals in the session.

2) Communication: Communication was measured by observing the frequency of occurrences of any nonverbal or verbal communicative action and the target audience of that action. Instances of communication were categorized as either being directed at the instructor, therapists, or peers. Included in instances of communication were nonverbal or verbal responses to requests, physically indicating a choice such as pointing at a song during song selection, and any vocal utterances with words. Responses to one-step instructions, vocal utterances that were not words, and laughing were not counted as communication.

3) Affect: Affect was measured by observing when a participant displayed positive affect (laughing, smiling, making positive comments or gestures, and attempting to touch robot/instructor) or negative affect (crying, whining, frowning, whimpering, vocally protesting, attempting to leave, kicking/hitting themselves/others, throwing items) during a session. The session was divided into ten second intervals and coded for instances of positive and negative affect. Since both positive and negative affect could be displayed within a ten second interval, it was possible for an interval to be categorized as both negative and positive. An interval was classified as neutral if neither positive or negative affect was exhibited. The percentage of positive intervals in a session was subtracted by the percentage of negative intervals to obtain their overall affect during the session.

4) Performance: Performance was measured based on the child's correct response to individual or group instructions. Responding to a request directed to the individual child, responding to a request directed to the group, and refraining from responding when an instruction was directed towards another individual was considered a correct response. Correct responses could be prompted by a therapist as this is a common form of teaching in ABA therapy and prompted correct responses were differentiated from independent correct responses. Not responding to a request directed to the child or the group and responding to a request directed toward another individual was considered an incorrect response. Performance data was only recorded in regards to the one-step instructions directed to individuals and to the group that were delivered throughout and between the activities of the session. A performance score was calculated by assigning one point to correct responses and half a point to correct responses that were prompted:

$$P = \frac{r_c + (r_{cp} \times 0.5)}{n} \tag{1}$$

where P is the performance score, r_c is the number of independent correct responses, r_{cp} is the number of prompted correct responses, and n equals the number of questions during that session.

D. Interobserver Agreement

The performance, engagement, communication, and affect data were coded by two independent therapists. 30% of each metric was double-coded to ensure that interobserver agreement (IOA) was over 70%. IOA was scored using the interval-by-interval method for affect and engagement. This method measures the intervals agreed upon by the independent coders divided by the total number of intervals in a session. Performance and communication were scored using total count IOA. This method takes the smaller frequency of observations divided by the larger frequency of observations. IOA scores below 70% were corrected by the two behavior technicians coming together to re-code the session. The final IOA scores were 85.39% for affect, 85.09% for engagement, 82.11% for communication, and 95.49% for performance.

V. Data Analysis

We evaluated our hypotheses using a two-way repeated measures ANOVA (RM-ANOVA) to investigate any interaction between independent variables and paired samples t-tests to identify significant differences between the experimental conditions. The two independent variables for the two-way RM-ANOVA were instructor type and session number. The dependent variables used for our statistical analysis were children's engagement, communication, affect, and performance. Engagement was analyzed based on its three components which included engagement with instructor/screen, therapist/peers, and off-target. Prior to running the two-way RM-ANOVAs and paired sample t-tests for each dependent variable, we assessed the normality of the data using Shapiro-Wilk's test. In cases where the normality assumption was violated, arcsin, square root, and natural logarithm transformations were used for engagement, communication, and affect, respectively to normalize the data [17]. An $\alpha = 0.05$ was set for all tests.

VI. Results

The results of our study are summarized in Table I and found that H1, H3, H4, and H5 were partially supported and H2 was not supported.

From our study, children's average engagement with the instructor and screen in the first human-led session (μ = 0.757, σ = 0.087) was lower than the third human-led session (μ = 0.774, σ = 0.156) but this difference was not significant (t(8) = -0.684, p = 0.513). Children's average engagement with their peers or therapists in the first human-led session (μ = 0.146, σ = 0.084) was higher than the third human-led session (μ = 0.113, σ = 0.057) but this difference was not significant (t(8) = 0.651, p = 0.533). Children's average off-target engagement in the first human-led session (μ = 0.174, σ = 0.062) was higher than the third human-led session (μ = 0.162, σ = 0.134) but this difference was not significant (t(8) = 0.841, p = 0.425). Their average communication in the first human-led session (μ = 27.611, σ = 12.469) was higher than the third human-led session (μ = 18.648, σ = 7.632) and this difference was significant (t(8) = 2.564, p = 0.033). Children's average affect in the first human-led session (μ = 0.098, σ = 0.125) was lower than the third human-led session (μ = 0.111, σ = 0.055) but this difference was not significant (t(8) = -0.573, p = 0.582). On

Learning Behavior		H1 vs H3	R1 vs R3	H1 vs R1	H3 vs R3
Engagement	Instructor or Screen	$t(8)$ = -0.684, p = 0.513	$t(8)$ = -0.630, p = 0.547	$t(8)$ = 2.161, p = 0.063	$t(8)$ = 1.310, p = 0.227
	Peers or Therapist	$t(8)$ = 0.651, p = 0.533	$t(8)$ = 0.372, p = 0.720	$t(8)$ = -0.152, p = 0.882	$t(8)$ = 0.226, p = 0.827
	Off-target	$t(8)$ = 0.841, p = 0.425	$t(8)$ = 0.132, p = 0.898	**$t(8)$ = 3.249, p = 0.012**	$t(8)$ = 0.497, p = 0.633
Communication		**$t(8)$ = 2.564, p = 0.033**	$t(8)$ = -0.497, p = 0.632	$t(8)$ = 0.290, p = 0.779	**$t(8)$ = -2.705, p = 0.027**
Affect		$t(8)$ = -0.573, p = 0.582	$t(8)$ = 0.262, p = 0.800	$t(8)$ = -0.164, p = 0.874	$t(8)$ = 0.392, p = 0.705
Performance		$t(8)$ = -1.491, p = 0.174	$t(8)$ = 0.775, p = 0.460	$t(8)$ = 0.200, p = 0.846	$t(8)$ = 1.978, p = 0.083

TABLE I: Results of the paired sample t-tests comparing the first human-led (H1) vs the third human-led (H3), the first robot-led (R1) vs the third robot-led (R3), the first human-led (H1) vs the first robot-led (R1), and the third human-led (H3) vs the third robot-led (R3).

average, they demonstrated lower performance in the first human-led session (μ = 0.763, σ = 0.177) than the third human-led session (μ = 0.835, σ = 0.118) but this difference was not significant ($t(8)$ = -1.491, p = 0.174). These results partially support our first hypothesis because there were no significant differences between children's engagement, affect, and performance in the first and third human-led sessions. There was only a significant difference between the children's communication on the first and the third human-led sessions.

Children's average engagement with the instructor and screen in the first robot-led session (μ = 0.826, σ = 0.073) was lower than the third robot-led session (μ = 0.844, σ = 0.083) but this difference was not significant ($t(8)$ = -0.630, p = 0.547). Children's average engagement with their peers or therapists in the first robot-led session (μ = 0.138, σ = 0.073) was higher than the third robot-led session (μ = 0.123, σ = 0.061) but this difference was not significant ($t(8)$ = 0.372, p = 0.720). Children's average off-target engagement in the first robot-led session (μ = 0.116, σ = 0.034) was lower than the third robot-led session (μ = 0.119, σ = 0.077) but this difference was not significant ($t(8)$ = 0.132, p = 0.898). Their average communication in the first robot-led session (μ = 28.888, σ = 23.443) was higher than the third robot-led session (μ = 28.333, σ = 10.392) but this difference was not significant ($t(8)$ = -0.497, p = 0.632). Children's average affect in the first robot-led session (μ = 0.104, σ = 0.118) was lower than the third robot-led session (μ = 0.115, σ = 0.157) but this difference was not significant ($t(8)$ = 0.262, p = 0.8). On average, they demonstrated higher performance in the first robot-led session (μ = 0.752, σ = 0.261) than the third robot-led session (μ = 0.704, σ = 0.218) but this difference was not significant ($t(8)$ = 0.775, p = 0.460). These results do not support our second hypothesis that children's learning behaviors in the first robot-led session would be higher than the third robot-led session because there were no statistically significant differences between conditions for engagement, communication, affect, or performance.

Our results show that children's average engagement with the instructor or screen in the first robot-led session (μ = 0.826, σ = 0.074) was higher than the first human-led session (μ = 0.757, σ = 0.087) but this difference was not significant ($t(8)$ = 2.161, p = 0.063). Children's average engagement with their peers or therapists in the first robot-led session (μ = 0.138, σ = 0.073) was lower than the first human-led session (μ = 0.146, σ = 0.084) but this difference was not significant ($t(8)$ = -0.152, p = 0.882). Children's average off-target engagement in the first robot-led session (μ = 0.116, σ = 0.034) was lower than the first human-led session (μ = 0.174, σ = 0.062) and this difference was significant ($t(8)$ = 3.249, p = 0.012). Their average communication in the first robot-led session (μ = 28.888, σ = 23.443) was higher than the first human-led session (μ = 27.611, σ = 12.469) but this difference was not significant ($t(8)$ = 0.290, p = 0.779). Children's average affect in the first robot-led session (μ = 0.105, σ = 0.118) was higher than the first human-led session (μ = 0.098, σ = 0.125) but this difference was not significant ($t(8)$ = -0.164, p = 0.874). On average, they demonstrated lower performance in the first robot-led session (μ = 0.752, σ = 0.261) than the first human-led session (μ = 0.763, σ = 0.177) but this difference was not significant ($t(8)$ = 0.200, p = 0.846). These results partially support our third hypothesis that children's learning behaviors on the first robot-led session were higher than the first human-led session. Namely, children's off-target engagement in the first robot-led session was significantly lower than their off-target engagement in the first human-led session. However, there was no statistically significant differences in children's affect, communication, or performance

Children's average engagement with instructor or screen in the third robot-led session (μ = 0.844, σ = 0.083) was higher than the third human-led session (μ = 0.774, σ = 0.156) but this difference was not significant ($t(8)$ = 1.310, p = 0.227). Children's average engagement with their peers or therapists in the third human-led session (μ = 0.113, σ = 0.057) was lower than the third robot-led session (μ = 0.123, σ = 0.061) but this difference was not significant ($t(8)$ = 0.226, p = 0.827). Children's average off-target engagement in the third human-led session (μ = 0.162, σ = 0.134) was higher than the third robot-led session (μ = 0.119, σ = 0.077) but this difference was not significant ($t(8)$ = 0.497, p = 0.633). Their average communication in the third robot-led session (μ = 28.333, σ = 10.392) was higher than the third human-led session (μ = 18.648, σ = 7.632) and this difference was significant ($t(8)$ = -2.705, p = 0.027). Children's average affect in the third robot-led session (μ = 0.115, σ = 0.157) was higher than the third human-led session (μ = 0.111, σ = 0.055) but this difference was not significant ($t(8)$ = 0.392, p = 0.705). On average, they demonstrated lower performance in the third robot-led session (μ = 0.704, σ = 0.218) than the third human-led session (μ = 0.835, σ = 0.118) but this difference was not significant ($t(8)$ = 1.978, p = 0.083). These results partially support our forth hypothesis that there will be no difference between children's learning behavior on the third robot-led and the third human-led session because there were no statistically significant differences between conditions on children's engagement, affect, or performance. There was only a significant difference in communication between the third human-led and robot-led sessions.

The results of the two-way repeated measures ANOVA showed that there was only significant interaction between instructor type and session number on children's performance (F(1,8) = 7.430, p = 0.026, η_p^2 = 0.482). However, the main effects for instructor type (F(1,8) = 1.614, p = 0.240, η_p^2 = 0.168) and session number (F(1,8) = 0.058, p = 0.815, η_p^2 = 0.007) were not significant. This result also partially supports our hypothesis as interaction effects were not observed between instructor type and session number for engagement, communication, or affect.

VII. Discussion

In this study, we investigated the differences in learning behaviors of children with ASD when participating in a robot-led and human-led group intervention session. The primary learning behaviors we focused on included children's engagement, communication, affect, and performance during the sessions. Our primary hypotheses were that there would only be differences in children's learning behaviors between the robot's first and third session as well as the robot's first session and the human's first session. Furthermore, we expected there to be an interaction effect on children's learning behavior between the type of instructor and number of sessions. The rationale for these hypotheses were that a large body of HRI research suggests that there is a novelty effect when individuals interact with social robots and current research has conflicting evidence on the novelty effect of social robots on children's learning behaviors [8], [9], [18]. The human-led sessions served as a baseline of comparison as these interactions were not novel to the children.

Overall, our results suggested that **H1** was partially supported because there were no differences in children's learning behaviors in the first and the third human-led sessions except for children's communication. Namely, children demonstrated a higher communication level in the first human session than the third session. This variation in communication may have occurred due to the inherent variability in the sessions caused by differing children groups, activities, and human instructors. The children were randomly assigned classmates in each session, which may have led to children being randomly assigned a friend. These children may have felt more comfortable in a particular session, ultimately leading to more communicative events. In addition, certain children may prefer some of the activities more than others. For instance, many of the communicative events occurred during song selection and the children often requested that the child selecting the song to choose the song they liked on the screen. Lastly, there was variability of human instructors because there were different therapists leading circle time sessions according to the clinic's schedule and it is hypothesized that the variability in their expressivity alters children's reaction to activities.

The results from the analysis rejected **H2** because the children had similar learning behaviors in the first and third robot-led sessions. These preliminary results suggest that we did not observe novelty effect in the robot-led sessions. This may be due to the consistency in the robot's behaviors across sessions and it has been shown that children with ASD enjoy routines and are more comfortable with mechanical stimuli which may have led to similar learning behaviors across sessions [11], [19]. It should also be noted that these are preliminary results from 3 sessions and we intend to investigate whether the novelty effect may arise after our planned 10 sessions.

The results from the analysis partially support **H3** because there was significant difference in off-target engagement between the first human-led and robot-led sessions. The overall off-target engagement in the human-led sessions were higher than the off-target engagement in the robot-led session, which suggests that the children viewed the robot as more appealing. A theme of children with ASD is that they lack motivation for certain social reinforcers. This can be explained by the Social Motivation Theory of autism that states that children with ASD are less intrinsically inclined to engage in meaningful social relationships [19]. In addition, this theory argues that children would prefer nonhuman stimuli. This could be an explanation for difference in certain learning behaviors for children across metrics. Furthermore, the children may have initially viewed the robot as a toy and children with ASD have been shown to exhibit increased engagement and interest in toys as well as mechanical stimuli [19], [20]. Finally, the robot was often more animated than the human. For example, in the "Copy Me" activity. Pepper was programmed to include noises (e.g., gorilla roaring) and music (e.g., sound of guitar) with each imitation request while the human performances had minimal to no vocal noises (i.e., animal or musical sounds). This discrepancy is expected because human instructors are not always as animated.

The results from the analysis partially supported **H4** as there were no differences in learning behaviors between the third robot-led and human-led sessions except for communication. Namely, communication was higher in the robot-led session than human-led sessions. This could be due to the increase in comfort children experienced from repeated exposure to the robot. Since the children may have felt more comfortable in the presence of the robot, they may have expressed their thoughts and feelings more often. The increase in communication of one child may have also caused a chain reaction and improved the atmosphere for other children to begin communicating. This may also cause the children to interact with each other and improve joint attention which is a specific goal of ABA therapy [21]. Such observations of spontaneous communication with robots suggest that robots could be useful in group ABA therapy for children with ASD because one of the aims of ABA therapy is to increase unprompted independent communication [22]. Independent communication often requires many hours of therapist prompted instructions to increase communication skills and prompts are difficult to fade out or in some cases cannot be completely removed. Hence, such unprompted communication during robot-mediated group interactions could improve the outcomes of ABA therapy.

Finally, the analysis partially supported **H5**. As reported in the results, the independent variables instructor type and session number interacted to influence children's performance only. This may be explained by the child's increased exposure to the instructor over time. The more sessions that the child is exposed to the instructor, the more comfortable the child may feel in a session. This comfort and familiarity may influence the child's willingness to respond to instructor questions [23]. This can possibly lead to higher or lower performance. The child may feel more comfortable answering individual and group instructions, which may increase performance. In contrast, an increase in response to peer instruction would decrease performance. The performance scores are unlikely to be negatively affected by question complexity because the instructions are simple behaviors (e.g., touching nose). Non-significant interaction between the independent variables was seen for the other metrics. Interaction may be seen in future studies when more sessions are completed.

The results of this study suggests that children with ASD learn behaviors similarly with human and robot instructors in a group classroom setting. This is a positive result because it demonstrates the plausibility of deploying robots to real world classroom settings. School is a formative period in a child's life and it is important that children experience an inclusive environment allowing them to reach their full potential [24]. By having robot instructors in schools, children with deficits in responding to human instruction may have a viable alternative. This can apply to not only children with ASD but also to neurotypical children that enjoy the consistent and repetitive but fascinating nature of robots. This study serves as a springboard for research into robot-led group instruction beyond the classroom setting such as robot-led group play between children with ASD. Most studies in literature are limited to one-on-one interaction between a robot and a child [25] or a few children and a robot [26], [27]. Lastly, we aim to extend this study to more than ten sessions to further investigate the long-term effects of robot-led group interventions on learning behaviors. This is necessary because it is common for children with ASD to participate in classroom settings that span many months. This long-term analysis will hopefully elucidate any differences between human-led and robot-led group instruction during an extended time frame.

REFERENCES

[1] M. Cooney and W. Leister, "Using the engagement profile to design an engaging robotic teaching assistant for students," *Robotics*, vol. 8, no. 1, p. 21, 2019.

[2] E. Vrochidou, *et al.*, "Social robot nao as a self-regulating didactic mediator: A case study of teaching/learning numeracy," in *2018 26th international conference on software, telecommunications and computer networks (SoftCOM)*, 2018, pp. 1–5.

[3] H. Chen, H. W. Park, and C. Breazeal, "Teaching and learning with children: Impact of reciprocal peer learning with a social robot on children's learning and emotive engagement," *Computers & Education*, vol. 150, p. 103836, 2020.

[4] D. Sampsel, G. Bharwani, D. Mehling, and S. Smith, "Robots as faculty: student and faculty perceptions," *Clinical Simulation in Nursing*, vol. 7, no. 6, pp. e209–e218, 2011.

[5] A. Edwards, C. Edwards, P. R. Spence, C. Harris, and A. Gambino, "Robots in the classroom: Differences in students' perceptions of credibility and learning between "teacher as robot" and "robot as teacher"," *Computers in Human Behavior*, vol. 65, pp. 627–634, 2016.

[6] N. Shin and S. Kim, "Learning about, from, and with robots: Students' perspectives," in *RO-MAN 2007-The 16th IEEE International Symposium on Robot and Human Interactive Communication*, 2007, pp. 1040–1045.

[7] S. Srinivasan and A. Bhat, "The effect of robot-child interactions on social attention and verbalization patterns of typically developing children and children with autism between 4 and 8 years," *Open Access Author Fund Awardees' Articles. 18*, 2013.

[8] N. Rakhymbayeva, A. Amirova, and A. Sandygulova, "A long-term engagement with a social robot for autism therapy," *Frontiers in Robotics and AI*, vol. 8, p. 180, 2021.

[9] M. T. van Otterdijk, *et al.*, "The effects of long-term child–robot interaction on the attention and the engagement of children with autism," *Robotics*, vol. 9, no. 4, p. 79, 2020.

[10] M. I. Ahmad, O. Mubin, S. Shahid, and J. Orlando, "Robot's adaptive emotional feedback sustains children's social engagement and promotes their vocabulary learning: a long-term child–robot interaction study," *Adaptive Behavior*, vol. 27, no. 4, pp. 243–266, 2019.

[11] J. Korneder, *et al.*, "Robot-mediated interventions for teaching children with asd: a new intraverbal skill," *Assistive Technology*, pp. 1–10, 2021.

[12] J. E. Favell, J. E. Favell, and J. F. McGimsey, "Relative effectiveness and efficiency of group vs. individual training of severely retarded persons." *American Journal of Mental Deficiency*, 1978.

[13] C. S. Martin, "Exploring the impact of the design of the physical classroom environment on young children with autism spectrum disorder (asd)," *Journal of Research in Special Educational Needs*, vol. 16, no. 4, pp. 280–298, 2016.

[14] R. L. Koegel and A. Rincover, "Treatment of psychotic children in a classroom environment: I. learning in a large group 1," *Journal of Applied Behavior Analysis*, vol. 7, no. 1, pp. 45–59, 1974.

[15] C. Lord, M. Elsabbagh, G. Baird, and J. Veenstra-Vanderweele, "Autism spectrum disorder," *The lancet*, vol. 392, no. 10146, pp. 508–520, 2018.

[16] American Psychiatric Association, *Diagnostic and statistical manual of mental disorders : DSM-5*, 5th ed. American Psychiatric Association Arlington, VA, 2013.

[17] J. P. Stevens, *Applied multivariate statistics for the social sciences*. Routledge, 2012.

[18] I. Leite, C. Martinho, and A. Paiva, "Social robots for long-term interaction: a survey," *International Journal of Social Robotics*, vol. 5, no. 2, pp. 291–308, 2013.

[19] C. Chevallier, G. Kohls, V. Troiani, E. S. Brodkin, and R. T. Schultz, "The social motivation theory of autism," *Trends in cognitive sciences*, vol. 16, no. 4, pp. 231–239, 2012.

[20] A. S. Mcduffie, R. G. Lieberman, and P. J. Yoder, "Object interest in autism spectrum disorder: A treatment comparison," *Autism*, vol. 16, no. 4, pp. 398–405, 2012.

[21] J. Kourassanis-Velasquez and E. A. Jones, "Increasing joint attention in children with autism and their peers," *Behavior analysis in practice*, vol. 12, no. 1, pp. 78–94, 2019.

[22] M. H. Charlop, L. Schreibman, and M. G. Thibodeau, "Increasing spontaneous verbal responding in autustic children using a time delay procedure," *Journal of applied behavior analysis*, vol. 18, no. 2, pp. 155–166, 1985.

[23] S. Lindsay, M. Proulx, H. Scott, and N. Thomson, "Exploring teachers' strategies for including children with autism spectrum disorder in mainstream classrooms," *International Journal of Inclusive Education*, vol. 18, no. 2, pp. 101–122, 2014.

[24] R. F. Catalano, K. P. Haggerty, S. Oesterle, C. B. Fleming, and J. D. Hawkins, "The importance of bonding to school for healthy development: findings from the social development research group." *The Journal of school health*, vol. 74 7, pp. 252–61, 2004.

[25] J. Wainer, K. Dautenhahn, B. Robins, and F. Amirabdollahian, "A pilot study with a novel setup for collaborative play of the humanoid robot kaspar with children with autism," *International Journal of Social Robotics*, vol. 6, no. 1, pp. 45–65, Jan 2014.

[26] J. Albo-Canals, *et al.*, "A pilot study of the kibo robot in children with severe asd," *International Journal of Social Robotics*, vol. 10, no. 3, pp. 371–383, Jun 2018.

[27] B. Huskens, A. Palmen, M. Van der Werff, T. Lourens, and E. Barakova, "Improving collaborative play between children with autism spectrum disorders and their siblings: The effectiveness of a robot-mediated intervention based on lego® therapy," *Journal of Autism and Developmental Disorders*, vol. 45, no. 11, pp. 3746–3755, Nov 2015.