Int J of Soc Robotics (2018) 10:93–113
https://doi.org/10.1007/s12369-017-0433-8

Human–Robot Interaction in Autism Treatment: A Case Study on Three Pairs of Autistic Children as Twins, Siblings, and Classmates

Alireza Taheri[1,4] · Ali Meghdari[1] · Minoo Alemi[1,2] · Hamidreza Pouretemad[3,4]

Accepted: 3 October 2017 / Published online: 17 October 2017

Abstract In this paper, three pairs of children with autism include a pair of twins, two siblings, and two classmates were enrolled in a 12-session robot-assisted group-games program. As many environmental factors were for the most part the same for the siblings as well as genetic factors for the twins, we were able to observe/compare the effect of the designed games on the participants individually and in paired-groups. The results indicated that all participants' autism severity decreased after the course of the program. Improvement in social skills, social participation/avoidance, and detrimental social behaviors were also observed in the participants with high-functioning autism with close to being large Cohen's d effect sizes. Moreover, based on the video coders' observations the joint attention, gaze scores toward the robot, and verbal communications of the paired-groups increased significantly over the treatment time ($p < 0.05$). However, in general, the designed program effect on the subjects' behavior seems to be different for participants from different points on the autism spectrum; and even the high-functioning subjects showed different potential behavioral progress.

Keywords Autism spectrum disorders · Human–Robot interaction · Joint attention · Imitation · Humanoid robot · Social skills

✉ Ali Meghdari
meghdari@sharif.edu

Alireza Taheri
taheri@mech.sharif.edu

Minoo Alemi
alemi@sharif.edu

Hamidreza Pouretemad
h-pouretemad@sbu.ac.ir

1 Social & Cognitive Robotics Laboratory, Center of Excellence in Design, Robotics and Automation (CEDRA), Sharif University of Technology, Tehran, Iran

2 Faculty of Humanities, Islamic Azad University, West Tehran Branch, Tehran, Iran

3 Institute for Cognitive and Brain Sciences (ICBS), Shahid Beheshti University, Tehran, Iran

4 Center for the Treatment of Autistic Disorders (CTAD), Tehran, Iran

1 Introduction

Children with autism have impaired reactions to real world events and often avoid social interactions and communications [1]. Since 2000, it has been widely shown that using social robots in autism treatment can significantly increase the impact of intervention sessions for children with ASD[1] [2–19]. In this type of research, the impact of the robot-assisted games/scenarios on children's performance as well as their different behavioral skills are investigated over time using different qualitative/quantitative assessment tools (such as video coding of the sessions, questionnaires, etc.) [3–9]. Improvement in imitation [10,11], joint attention [2–4,11–14], communication and social interaction skills [2,3,15–17] of participants during (even a short number of) robotic interventions are the most frequent findings of previous studies. Nevertheless, roboticists have mostly studied children with high-functioning ASD [2,12,18] and simultaneous investigations of participants with high- and/versus low-functioning autism are less frequent in this field of research [11,19].

In particular, investigating siblings and twins with autism has always been an interesting topic for psychologists to dis-

1 Autism Spectrum Disorders.

cover more information about autism neurodevelopmental disorder [20–30]. In [21], genetic and environmental bases of nonverbal communication and social interaction have been studied in more than 1000 child/adolescent twins and siblings with pervasive developmental disorders. By investigating 664 infants, Ozonoff et al. [22] reported that 18.7% of younger siblings of individuals with autism spectrum disorders developed autism, which is a rate significantly higher than the general population. Huskens et al. [23] have studied the effectiveness of Lego-based interventions on a broad range of collaborative behaviors in three children with ASD and their siblings in a multiple-baseline design research. Although they did not observe significant changes in the collaborative behaviors of their participants with autism, they reported a potential for robot-assisted interventions to increase the responses and interaction initiations of ASDs. Alternatively, previous research conducted without robots on twins with autism has focused mainly on the relative effects of genetics and environment [24,25], autistic traits in twins [26,27], and comparing monozygotic (MZ) and dizygotic (DZ) twins with autism [28,29]. Contrary to the extensive basic studies on twins with ASD, there has been less research on the behavioral clinical treatment of autistic twins [20,30]. In 1975, the development of social skills of a pair of twin brothers with ASD has been investigated by Kean [20]. Hilton and Seal [30], studied the communication/behavioral performance of a pair of twins with ASD during sixteen sessions. After their trial interventions in DIR[2] and ABA,[3] and comparing the questionnaire results with their clinical observations, they observed a slight gain in CSBS[4] questionnaire results of the ABA participant. While there is very limited data for robot-assisted behavioral interventions on twins with autism in the literature so far [8], a study based on tablet apps has been done on a pair of 4-year-old twins with autism during 11 weeks in 2015 [31]; and the authors concluded that iPad/tablet devices can be effective in improving turn-taking behaviors of ASDs.

In this research, we have designed a set of robot-assisted therapeutic games and run them on 3 pairs of Iranian male individuals with autism during 12 intervention paired-group sessions as a case study. What makes this study different is that 2 out of the 3 pairs of participants are special cases which are rarely investigated in robotic-based studied in the literature; one pair is a set of twins with autism (includes one high- and one low-functioning 7-year-old subject); and the second pair are two (15- and 10-year-old) siblings with high-functioning autism. The strength of investigating twins and siblings are the control advantage provided by these special cases, including the same parents, food, clothing and education, all which are difficult subjects to otherwise control. The main goal of this research was to investigate how the effect of the designed robot-assisted protocol could differ for our different participants individually and in paired-groups. To this end, we have simultaneously investigated each subject's behavioral data inside and outside the study by performing: (1) content analysis of the video records of the interventions, (2) blind assessments of the participants' behaviors before and after the study by a clinical child psychologist, and (3) a collection of parent observations of their children's behavioral changes during the course of the program. Moreover, we will investigate changes in three subscales of the participants' social skills. Through this case study, we would like to find preliminary exploratory potentials of having a robot in paired-group autism classrooms.

2 Research Methodology

Study approvals Ethical approval for the protocol of this study was provided by the Iran University of Medical Sciences (#IR.IUMS.REC.1395.95301469), and the certification for Applied Behavioral Analysis (ABA) and Robot-Assisted treatment with children with autism was received from the Center for the Treatment of Autistic Disorders (CTAD), Iran. The participants' parents were neither paid nor had to pay for the sessions and were acting as volunteers. In addition to uphold moral obligations, both the parents and the researchers signed pledge/consent forms before the interventions began.

2.1 Participants with Autism

The six male participants with autism consisted of three pairs; (1) a pair of a 7-year old fraternal twins, one of whom with high-functioning and the other one with low-functioning autism, (2) two siblings with high-functioning ASD, one 15 years old and the other 10 years old, and (3) two high-functioning classmates ages 6 and 7 years old. The participants are referred to as "Twins: P1-A, P2-I", "Siblings: P3-K, P4-M", and "Classmates: P5-T, P6-H", respectively, throughout the manuscript.

Twins Participant P1-A (with high-functioning autism) has suffered from hyperactivity (ADHD) and eye-contact-avoidance from an early age, and has mild verbal skills; since he started speaking 1 year later than typically developing children. The other twin, P2-I, has a more severe case of autism. He has poor verbal skills, with a Mean Length of Utterance (MLU) < 3, and often engages in repetitive fingers fluttering, non-purposeful/stereotyped behaviors.

Siblings P3-K, the 15-year-old participant, started talking for the first time when he was 5. He usually avoids social interactions and also has a lack of eye contact. P4-M is more

[2] Developmental, Individual-Difference, Relationship-Based Model.

[3] Applied Behavioral Analysis.

[4] Communication and Symbolic Behavioral Scales.

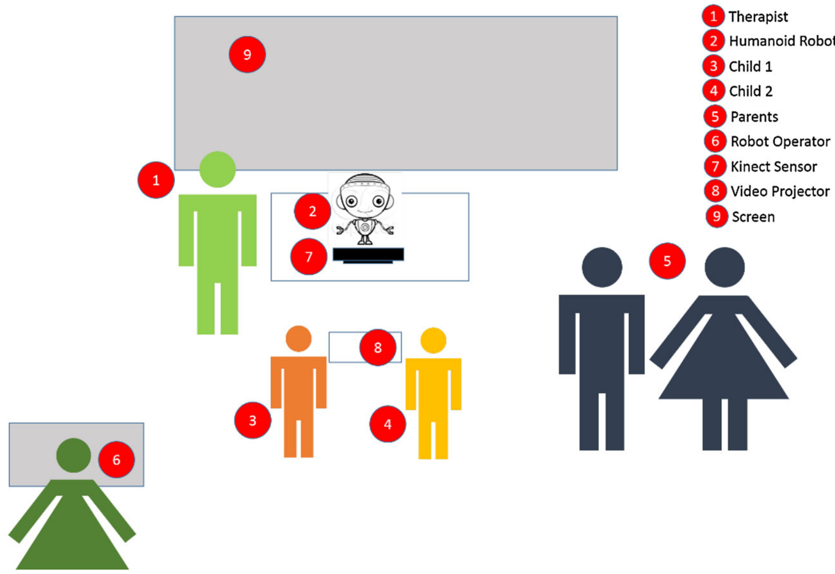


Fig. 1 Schematic of the experimental setup

sociable than his brother and usually join the group activities of his peers; however, he shows aggressive behaviors in school and also has some repetitive/stereotyped behaviors. Both of the siblings do not have a tendency to initiate a talk; and only communicate verbally when they have to answer the other's questions. P3-K and P4-M are students in regular schools; however, they do not perform well educationally and are the weakest students in their classes.

Classmates 7-year-old P5-T has the least deficit in social skills among our participants. His main problem is in attention span and making unclear sounds which negatively affects his learning process. P6-H did not have adequate communication with his parents during childhood; and he had a deficit in verbal development. His lack of eye contact is quite obvious. In this study, P5-T and P6-H were selected as peers/classmates to take part in the robot-assisted interventions.

2.2 Setup of the Study

The designed intervention scenarios include the humanoid robot(s), each pair of participants, their parents, therapist, and a robot operator, in order to engage the participants in different individual/group imitation and joint attention tasks. During this study, each pair attended 12 thirty-minute sessions during $\sim$ 3 months at the Social & Cognitive Robotics laboratory of Sharif University of Technology.

The sessions were held in a $5 \times 5 \times 3\,\text{m}^3$ room. Depending on the designed scenario, one or occasionally two humanoid robots were involved in the experimental setup. Two laptops, two cameras (for filming sessions), Microsoft Kinect Sensor, a video-projector, as well as a whiteboard and chairs were also included in the setup. Without having a control group, this case study is based on a Wizard of Oz style robot control and all of the instructions were given by the robot and/or the therapist. The schematic setup of the study is shown in Fig. 1.

2.3 Humanoid Social Robots

One of the humanoid robots used in our educational-therapeutic program was a NAO-H21 manufactured by the Aldebaran Company with 21 degrees-of-freedom (DOFs) [32]. It was renamed "Nima", a Persian boy's name, for use in the Iranian context. The other Humanoid Robot was an Alice-R50 created by the Robokind Company which has 32 degrees-of-freedom [33]. We renamed it "Mina", a Persian girl's name. These two robots have been used in different autism studies worldwide [3,7,12,18]. Although the majority of our scenarios were based on the Nima robot, we included the Mina robot because of its ability to show different facial expressions (having 8 DOFs in the face) as well as allowing us to determine the effect of robot form on the participant's results (Figs. 2, 3).

2.4 Designed Games and the Therapeutic Protocol

Imitation and motor skills [2,3], initiating interactions/communications [1,2,6], and initiating/responding to joint attention behaviors [1,4,13] are often impaired in children with autism. We have developed a variety of therapeutic games to investigate the potential benefits and different effects of the robotic-interventions on our subjects with high- and low-functioning autism based on their impairments. The games were designed based on Applied Behavioral Analysis (ABA) [30]. These scenarios, whose potential effectiveness was confirmed by clinical child psychologists, concentrate on affecting ASD's different cognitive skills, imitation, joint attention, social skills, eye-contact, and turn-taking. The participants took part in some of these games at each

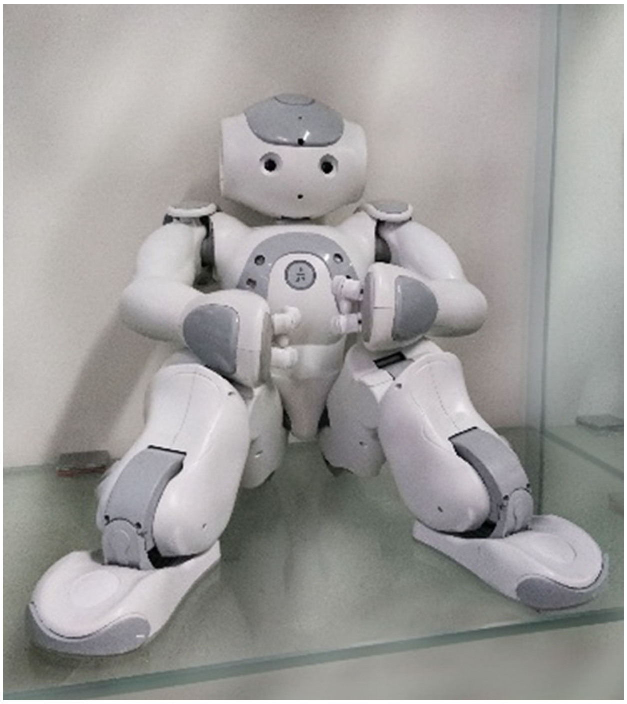


Fig. 2 The NAO (Nima) robot

Fig. 3 The Alice (Mina) robot

session in different modes: Robot–Child or Robot–Child–Peer/Parent/Therapist interactions. Robotic scenarios were programmed in C#, JAVA, NAO platform Choregraphe (by Aldebaran Co.) [32], and Workshop software (by Robokind Co.) [33].

The game scenarios are presented in Table 1 and the structure of the interventions is shown in Fig. 4. The games which concentrate on "imitation" and "joint attention" skills are referred to as type-I and type-JA, respectively. In games I/1 and I/3, the robot is more likely to be a follower, companion, or reinforcement tool; however, in the other games, the robot's engagement is much more which performs as a teacher. During the games, the desired actuator commands were sent to the robot in two ways: manually by the robot operator or automatically through the Kinect sensor or haptic Phantom-Omni encoders. It should be noted that due to ethical aspects, a debrief about the manual control of the robots was given to the parents after the interventions. Whether the child completed a task correctly or not, the robot provided verbal reinforcement encouraging the participants or gave them a big round of applause for their efforts.

2.5 Assessment Tools

To reach the goals of this study, four types of measuring instruments including Video Coding of Intervention Sessions, Questionnaires, Blind Human Assessments, and Interviews with the Parents were used at different times.

2.5.1 Gilliam Autism Rating Scale (GARS) Questionnaire

Developed by Gilliam in the 1990's, this is one of the most widely used scales for autism diagnosis and assessment [34]. The GARS scale is broken down into four addition subscales: Stereotyped Behaviors, Communication, Social Interactions, and Developmental Disturbances. It has been established that the subscale' Cronbach's alpha coefficients are 0.90 for stereotyped behaviors, 0.89 for communication, 0.93 for social interactions, 0.88 for developmental disturbances, and 0.96 for autism typology [35]. After studying the GARS in Iran on 100 children, the Cronbach's alpha for the four subsections and overall test were 0.74, 0.92, 0.73, 0.80, and 0.89, respectively [36].

2.5.2 Autism Social Skills Profile (ASSP) Questionnaire

The Autism Social Skills Profile (ASSP) was developed by Scott Bellini [37]; it is a comprehensive tool to measure social behaviors of children with autism. The items of this questionnaire are designed in such a way that it can be easily used in either single or group educational programs with social aims. ASSP can also be used to measure the level of improvement after the treatment and the items represent a wide range of social behaviors in children from 6 to 17 years old with autism. This questionnaire consists of 45 items. Higher scores indicate more positive social behaviors. ASSP has been translated to Farsi, referred to as NAMA, and validated in Iran with a Cronbach's alpha larger than 0.7, which is quite acceptable. This questionnaire can be filled in within 15–20 min by the children's parents, teacher, or any other adult familiar with the subjects' social behaviors. The items of this questionnaire are on a 4-point Likert scale (never, sometimes, often, and always) and responses are graded from 1 to 4.

Table 1 List of the designed games

Game type/#	Description of the game	Main purposes of the game
I /1	Real-time imitation by the robot in upper body movements of the child (in Robot–Child mode)	Robot as a companion/reinforcement tool: draw the attention of the child to the robot and therapist Reciprocal imitation
I /2	Teaching imitation/motor skills by the robot to the children through individual/paired-group exercise and dances (in Robot–Child and Robot–Child–Peer/Parent modes)	(Gross) imitation Dyadic/Triadic interactions
I /3	Tele-operating the humanoid robots' arms and head using a 6-DOFs haptic Phantom-Omni robot as a remote controller (in Robot–Child and Robot–Child–Peer/Parent modes)	Empowering the children/therapist to move the robots' joints arbitrary Dyadic/triadic interactions Turn-taking games Imitation
I /4	Playing a real xylophone (in Robot–Child mode)	(Fine) imitation Joint attention Turn-taking Eye-hand coordination
JA /1	Pointing to far/near points and showing the cards/objects by the robot/child (in Robot–Child mode)	Joint attention Turn-taking Gaze-shifting
JA /2	Kinect-based recognition game and Classification of animals, fruits, places, and objects by pointing to different baskets on the screen (in Robot–Child and Robot–Child–Parent modes)	Classification Joint attention and pointing Gaze-shifting
JA /3	Playing a developed Kinect-based virtual xylophone on the screen (in Child–Parent/Therapist modes)	Joint attention Imitation Gaze-shifting Visual pursuit

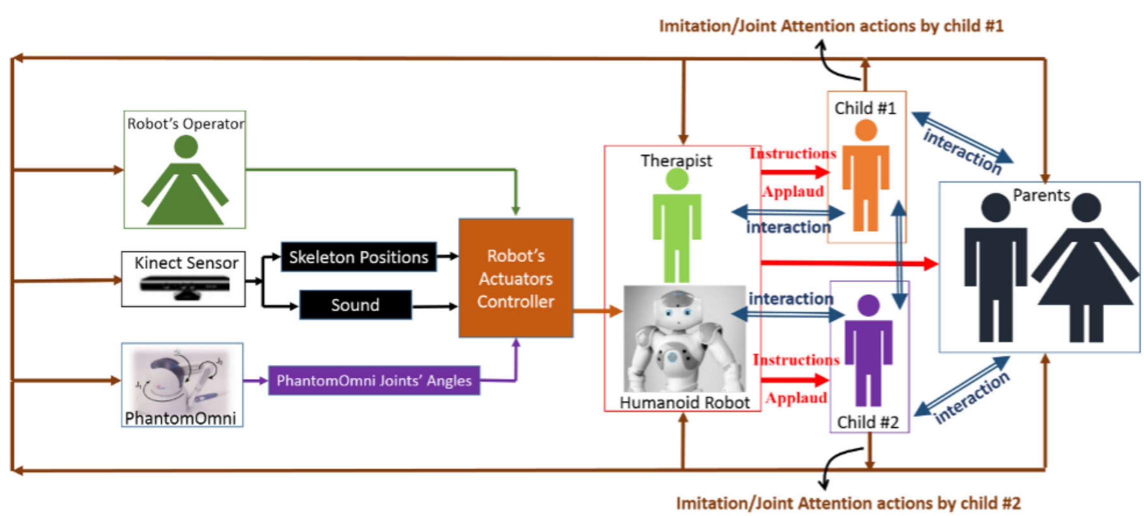


Fig. 4 The modular structure of the interventions for the paired-groups

There are three subscales of social skills including: Social Reciprocity, Social Participation/Avoidance, and Detrimental Social Behaviors.

Due to the small number of children (and paired-group members) in the current research, we need to be careful in running scientific statistical analysis significant tests on the questionnaires' data. Notwithstanding, we calculated the Cohen's *d* effect size [38] among the Pre- and Post-Test results of the ASSP questionnaire, for two situations: (a) considering all six participants as a group, and (b) excluding the low-functioning subject and only considering the five high-functioning subjects. We considered the Cohen's *d* effect size, which is independent of the sample size, as an estimation measure of the treatment effect. This has been used not for making strong claims and generalizing the observed results, but for finding the potential of the robot-assisted interventions' effect on ASDs in the sessions (Table 4).

2.5.3 Quantitative Content Analysis of Intervention Video Records

Quantitative content analysis is a powerful tool to analyze written texts, videos or other media. This method is systematic, flexible, and replicable and can be used by specialists in behavioral assessments [39]. During this research, the behavior of the participants during the intervention sessions had been recorded for further analyses. Although the content analysis of videos was time-consuming and costly, it gave us valuable behavioral patterns of the participants/paired-groups. Two psychologists separately observed the videos and rated the behavior of each participant for all the intervention sessions (with Pearson correlation $r = 0.719$ between the evaluators); then the mean of their scores were taken as each of the task score of the subjects. To this end, some items, among the different social and cognitive impairments of ASDs, were extracted based on three valid questionnaires: GARS [34], Autism Treatment Evaluation Checklist [40], and ASSP [37]. These items consist of: (1) Gaze-Shifting, (2) Joint Attention, (3) Imitation, (4) Maladaptive Behaviors, and (5) Verbal Communications.

Each participant received a score of +1 for every correct task. These tasks could be correctly done by following the games' instructions or by the child's creativity. The time interval of the tasks was not a criterion for the evaluations of our video coders. The evaluators did not have information about the research questions of the study; and they tried their best to reduce the amount of human errors in the quantitative content analysis scoring.

2.5.4 Human Assessment of Behaviors

A clinical child psychologist outside the study assessed the subjects' autistic behaviors both 1 week prior and 1 week after the program (as Pre- and Post-Tests) in order to observe the effect of the robot interventions on the participants.

2.5.5 Interview with the Subjects' Parents

Due to existence of opportunities for the children to show novel social behaviors [2] outside the study environment, we asked the parents who spent extensive time with their children to inform us of any behavioral changes they may have observed. In this regard, the parents were interviewed two times includes: immediately at the end of the program and a few months after the interventions to obtain information on (1) the immediate effects of using robots on their children, and (2) long term/durability effect of the experiment.

3 Results and Discussions

The participants started with an introduction session with the humanoid robots showing their capabilities to each paired-group and their families. This session began with the robots greeting the children and continued with cheering them up by dancing, singing songs, calling their names, shaking hands, etc. At the end of the first session, a short description of the lesson plans for the following intervention sessions was given. During the sessions, each participant was involved in at least one of the imitation/joint attention games' type/modes. Table 2 presents the schedule of our intervention sessions for the three pairs. In order to observe the effect of the designed games' package on the participants with autism, the order of the played games were shuffled for the three paired-groups; We have run both type-I and type-JA games in each four-session's section; while the number of total games for different groups were kept the same as much as possible. Table 2 shows that the participants were engaged in individual/group modes of the therapeutic games. It should be noted that in Table 2, the games were sorted according to their types/#; and chronological orders of the played games for the children are not mentioned here (Fig. 5).

The four following measurement instruments were used to observe the effects of the interventions: (1) Quantitative content analysis of the videos, (2) GARS and ASSP questionnaires, (3) Human assessments, and (4) Interview with the parents. In order to identify possible changes on the participants inside and outside of the study, these different assessment tools were used simultaneously.

3.1 Content Analysis of Video Records

Based on the movies of the intervention sessions, the video coders rated different behaviors of each subject separately in each game/session. Using Anderson–Darling test [41], we observed that the behavioral scores for each group

Table 2 Intervention sessions schedule including games type/#, the playmate in the games, and the total game numbers during the treatment time for each paired-groups

Treatment Time	Twins (P1-A, P2-I)				Siblings (P3-K, P4-M)				Classmates (P5-T, P6-H)			
	Games Played by each child	Playmates			Games Played by each child	Playmates			Games Played by each child	Playmates		
		Robot	Parent / Therapist	Peer		Robot	Parent / Therapist	Peer		Robot	Parent / Therapist	Peer
Sessions #1-4	I/1				I/1				I/1			
	I/2				I/2				I/2			
	I/3				JA/2				JA/2			
	JA/1											
	JA/2											
	JA/3											
	# of Games per Participant			7	# of Games per Participant			6	# of Games per Participant			6
Sessions #5-8	I/2				I/2				I/2			
	I/3				I/3				I/3			
	I/4				I/4				I/4			
	JA/1				JA/1				JA/1			
	# of Games per Participant			8	# of Games per Participant			7	# of Games per Participant			7
Sessions #9-12	I/2				I/2				I/2			
	JA/2				I/4				I/4			
					JA/1				JA/1			
					JA/2				JA/2			
					JA/3				JA/3			
	# of Games per Participant			6	# of Games per Participant			8	# of Games per Participant			8

had approximately normal distribution. In order to compare/examine the effect of two categorical independent variables: “participants” and “treatment time” (and their probable interaction) on different children’s “behavioral scores”, two-way ANOVA analysis tests were applied in two separate conditions using Minitab Software [41].

For the factor “participants”, two conditions were taken into accounts; (1) considering each of the 6 participants individually (i.e. having six levels: P1 to P6), and (2) considering each playmate pair as a separate group; therefore, we will have three levels: “Twins”, “Siblings”, and “Classmates”. It should be noted that due to the nonhomogeneous performance of the participants (especially the twin brothers), considering them as paired groups is not necessarily meaningful and the prerequisite of having similar performance on the desired behavioral scores is required. In addition, in both conditions, there were only three levels for the factor “treatment time”: “Sessions #1–4”, “Sessions #5–8”, and “Sessions #9–12”. In other words, the effectiveness of this program for the subjects is studied during three time intervals as the representation of their performance in each month of the robot-assisted interventions. For both of the mentioned conditions, the significant p values of two-way ANOVA tests according to each factor, as well as their interaction, are presented in Table 3 which will be discussed separately in the next subsections. Fortunately, the two-way ANOVA test is quite robust to violations of normality and approximate normally distributed data [41].

3.1.1 Gaze-Shifting

The gaze-shifting score is the total number of the participant’s gazes toward the robot and/or the other individuals (i.e. the therapist, parents, or his playmate) in the room during the intervention sessions. The purposes of studying this Visual Attention-based parameter are to find out (a) whether there is correlation between the rate of the total gaze-shifts of different participants/paired-groups during the treatment time, and (b) the gaze-shifting scores’ trend toward the robot (i.e. the item with the most important role in our therapeutic sessions). In order for the data to be comparable, the scores of each subjects were normalized by dividing the number of his total gaze-shifts to the play time; therefore, the unit of this parameter is “number per minute”. Unfortunately, our video coders could not score the time duration of eye contact/gaze fixation of the subjects with the robots and other individuals

Fig. 5 Snapshots of the intervention sessions; **a** the twins in the Robot–Child–Peer mode of game I/2, **b** P1-A in the Robot–Child mode of game JA/2, **c** P4-M responding to the pointing of the robot in game JA/1, **d** P3-K tele-operating the Nima robot via the haptic robot (game I/3) and his brother, P4-M imitating the robot's movements (game I/2) simultaneously, **e** P5-T imitating the robot in Robot–Child–Parent mode, **f** P6-H playing the real xylophone (game I/4), **g** P1-A tele-operating the robots via the haptic robot (game JA/2), and **h** P6-H playing the virtual and real xylophones (games I/4 and JA/3) in Robot–Child–Parent mode

in the class, manually. They also did not score the number of children's glances toward the objects in the class.

The interaction plot for total gaze-shifting scores in Condition 1 is presented in Fig. 6. According to Table 3, there was a statistically significant difference between the participants' total gaze scores; however, no significant monotonic trend was observed during the treatment time. Figure 6 also shows increase in the scores of the twins, i.e. P1-A and P2-I; while the total gaze-shifting scores of the other 4 participants were almost constant over time. To study whether the increase in total gaze-shifting of the twins during the sessions is significant, an ANOVA test was applied on simple linear regression models of their scores separately (during the whole sessions). The associated p values for the overall

Table 3 Two-way ANOVA; the table of significant p values of the independent variables "participants/paired-groups" and "treatment time", and their interaction for different subjects' behaviors using Minitab Software

Item	Dependent variable	Condition 1: individual participants P1 to P6			Condition 2: participants in three paired groups: "Twins", "Siblings", and "Classmates"		
		Participants	Treatment time	Interaction of the factors	Paired-groups	Treatment time	Interaction of the factors
Gaze-shifting	Total gaze score	**0.000**	0.079	0.078	0.622	0.110	**0.021**
	Gaze toward the robot	0.201	**0.010**	0.146	0.327	**0.004**	**0.006**
	Gaze percentage toward the robot	**0.001**	**0.038**	0.319	**0.017**	**0.024**	0.146
	Gaze toward the humans	**0.000**	0.401	0.112	0.475	0.578	0.106
Joint-attention (JA)	Pointing score	0.310	**0.003**	0.391	0.083	**0.001**	0.301
	Responding to JA score	0.932	**0.008**	0.831	0.703	**0.005**	0.338
	Total score of JA	0.449	**0.000**	0.264	0.115	**0.000**	0.281
Imitation	Imitation success rate	0.483	0.910	0.932	0.288	0.878	0.756
Behavioral problem	Maladaptive behaviors score	**0.000**	0.698	0.425	0.471	0.676	0.213
Verbal communication (VC)	Total VC score	**0.000**	**0.038**	0.200	**0.002**	**0.046**	0.169
	Amount of VC with robot	**0.042**	**0.017**	0.052	0.051	**0.021**	0.113
	VC percentage with robot	**0.023**	**0.030**	**0.028**	0.647	0.250	0.736

Numbers that represent the 95% confidence intervals (i.e. statistical significance at the $p < 0.05$ level) are in bold in the table

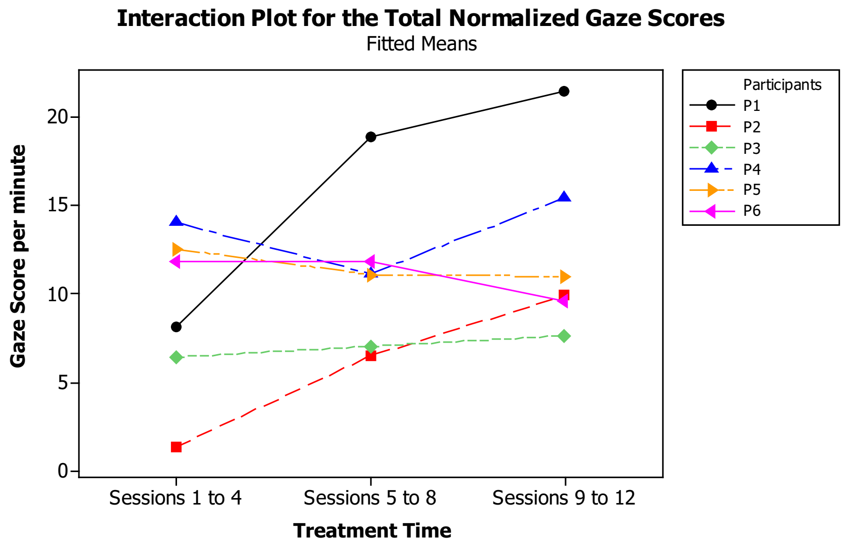
F-test of the twins' models are 0.02 and 0.02, respectively; therefore, we can be confident that some linear relationship exists between the gaze scores of P1-A and P2-I and the total session numbers. Considering that the treatment time has no significant effect on the gaze scores of either the participants or paired-groups, the observed improved performance for the twins could also be due to the order of the games played for them.

Figure 7a, b show the interaction plot for the normalized number and the portion of gaze-shifts toward the robot(s) in Conditions 1 and 2, respectively. Table 3 interestingly reveals that the normalized gaze numbers toward the robot(s) are significantly different among the participants/paired-groups (in both Conditions 1 and 2). On the other hand, both factors "participants/paired groups" and "treatment time" had significant effect on the gaze portion while there was not a significant interaction between the effects of factors 1 and 2. Based on the results, we can be confident that the participants'/paired-groups' visual attention toward the robot(s) improved overall during the course of the program. It can also be hypothesized that during the sessions, the friendship between the robots and the subjects was getting stronger; therefore higher attention rates toward the robot were expected over time. Improvement in gaze scores of the subjects toward the robot means that the robot was socially acceptable for our participants and could facilitate their social interactions; which is in line with the findings of [6,42]. Moreover, this observation could be the foundation for increase in verbal/non-verbal communication of the participants with the robot (Sect. 3.1.5) which is also reported in [42] while they used their simple robot, Keepon. According to Fig. 7a, the most improvement for the number of gazes toward the robot is observed for P1-A and P2-I (i.e. the twins). Also, Fig. 7b indicates that the performance of the siblings' pair in gaze portion toward the robot(s) is $\sim$ 50% (steady) during the program.

3.1.2 Joint Attention (JA)

In this study, the video coders scored the following behaviors as the Joint Attention (JA) scores of the participants: (a) following the pointing/gaze-direction of the robot or other individuals in the room correctly, or gaze-shifting between a person and other situation/object; which is known as the Responding to Joint Attention (RJA) score, and (b) showing an object, or pointing to far/near points was called the Pointing score. The total JA score is the sum of RJA and Pointing activities that have been observed in the interventions' movies. The scores have been normalized/reported by dividing the number of correct tasks to the game time (i.e. number per minute). In this research, joint attention behaviors occurred more often in the Kinect-based recognition game, pointing game, and playing the virtual/real xylophone game.

Fig. 6 Interaction plot for the total normalized gaze-shifting scores considering the effect of treatment time and the participants individually (Condition 1)



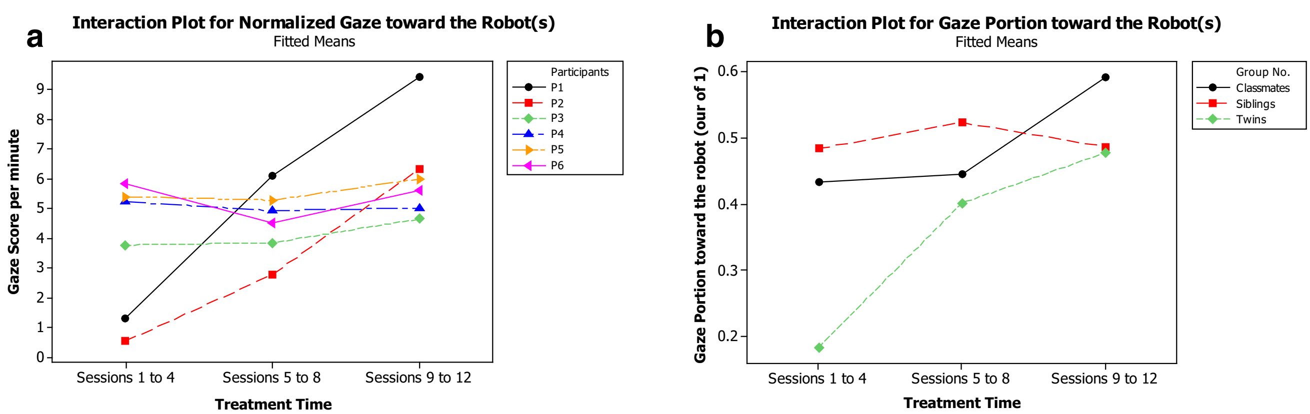


Fig. 7 Interaction plot for the **a** normalized gaze-shifting scores toward the robot in Condition 1, and **b** portion of gaze toward the robot (out of 1) in Condition 2; considering the effect of treatment time and the participants/paired groups

To clarify the definitions/differences of Sects. 3.1.1 and 3.1.2, Gaze-shifting in itself is only a form of attention shifting in most cases. It only becomes a form of joint attention when it is synchronized between interacting partners.

The interaction plots for RJA and JA scores in Conditions 2 and 1 are presented in Fig. 8a, b, respectively. According to Table 3, although there is no significant increasing trend in the JA scores versus treatment time, both the individual participants and paired-groups had significant effects on the joint attention and its subscales. We can observe that the Classmates paired-group performed better in the case of RJA. This is a sign that the participants of this study had nonhomogeneous levels of joint attention skill. It should be noted that the lower scores of the JA in Sessions #5–8 in comparison to the other two session-parts are mostly because of the unbalanced JA situations provided for the participants and does not necessarily mean a decrease in the JA skills of the children. However, the observed increasing trends for some of the subjects such as P1-A, P2-I and P6-H seem to be meaningful and are in agreement with the human assessments' observations on JA behaviors (which is presented in Sect. 3.3). The improvement in joint attention skills of ASDs has been confirmed in many robot-assisted studies [4,13,43]. In [4], the authors observed that their robot provided a mediator for joint attention skills of their subjects with autism as well more the possible contribution of this technology in autism research.

3.1.3 Imitation

In the imitation games, the humanoid robot performed a movement and asked the participant/paired-group to do the same action. The Imitation tasks of the robot by the subjects included a wide range of easy and difficult actions such as simple gross movement of arms and head, daily exercise, one-leg balance, and hitting the real/virtual xylophone's bars. During the exercise/dance games, we faced the ceiling effect in "gross" imitation skills of P1-I, P3-K, P4-M, and P5-T.

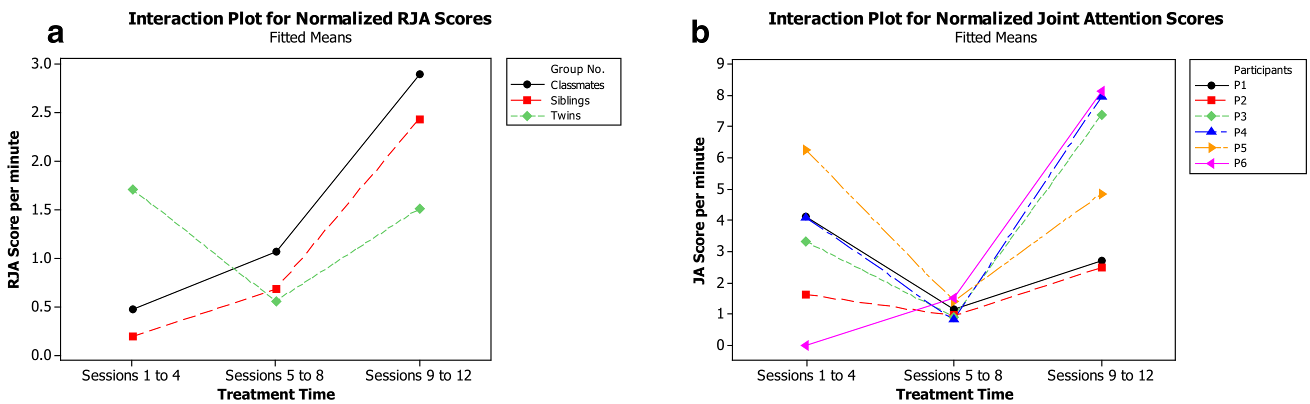


Fig. 8 Interaction plot for the **a** normalized RJA scores in Condition 2, and **b** joint attention scores in Condition 1; considering the effect of treatment time and the participants/paired groups

Imitation of the robot's actions in playing a real xylophone was not difficult for the participants; however, their incorrect performance were in rhythm perception/imitation when the robot hit three or more bars with different time intervals. Moreover, the assignment of playing the virtual xylophone game was new and appealing for the participants with autism. To hit the same virtual xylophone bars as the therapist/robot, they had to do a complex cognitive imitation/joint attention task; i.e. looking at the screen and remembering the bar color hit by the therapist/robot and then trying to play the same note by virtually touching the right bar.

Based on clinical psychologists' advice, we considered this fact that the "learning effect" most likely occurs for the participants after being involved in a game; and repeating that game with the same difficulty level may seriously affect their scores/performance. Therefore, to prevent this issue, our strategy was to conduct hierarchical and easy-to-difficult level games as much as possible over time.

Figure 9 shows the interaction plot of the percentage success rate of the participants during the imitation games. The success rate in each session was calculated by dividing the number of each participant's correct tasks by the number of his whole tasks in that session. According to Table 3, Fig. 9 does not seem to show an improvement trend in the imitation performance of the whole participants. It should be noted that the performance of the children could also be due to the change in session's games/instructions level and their moods (i.e. the performance of P6-H). According to the psychologist's assessment (Sect. 3.3), the slight increasing trends for P4-M and P5-T's performance are due to their progress in fine imitation skills. Also, we have investigated the performance of the low-functioning subject during the whole sessions; P2-I did not do very well in the hard exercise in session ten. His low performance was because of his problem in one-leg balance imitation for 15 s, long sequence of the actions, as well as his low mood during that session. Disregarding P2-I's performance in session ten (as an outlier data), linear regression analysis was carried out using Minitab Software to see whether the improvement trend observed for his imitation versus intervention time was statistically significant. After applying the ANOVA statistical test, we indicated that the regression model was statistically significant, $F = 61.5$ ($p = 0.001 < 0.05$). We noticed that $R^2 = 93.9\%$ which means the session numbers explains the quite good variability of his imitation success rate. This observation could serve as a helpful comparison to the psychologist's pre- and post-assessments of P2-I's imitation skills.

Improving imitation/motor skills of children with autism through similar types of (gross/fine) imitation games including robotic-assisted studies [2,7,10,11] and research without robots [44–46] have been also mentioned which confirm the tentative finding for the low-functioning and two of the high-functioning subjects.

3.1.4 Maladaptive Behaviors

In this paper, the video coders quantified the maladaptive behaviors of the participants by counting the total number of activities include: lack of cooperation, aggression/jangling, intervening group games, stereotyped behavior, meaningless repetition of words and echo, ecstasy and inattention to the group, and engaging in solitary interests and hobbies. The results are presented in Fig. 10. Stereotyped behaviors could also happen due to the increase level of children's excitement, which would mean that the children enjoy the robot-assisted intervention. However the only situation that meaningless repetition of words and echo are positive is when the child does not have imitation/verbal skills at all. For the participants of this study, this situation does not exist and the meaningless repetition of tasks are considered as their maladaptive behaviors.

Table 3 shows that among the ANOVA analysis in the two conditions, the only significant factor on the maladaptive behaviors score is the participants. P1-A, P3-K, and (inter-

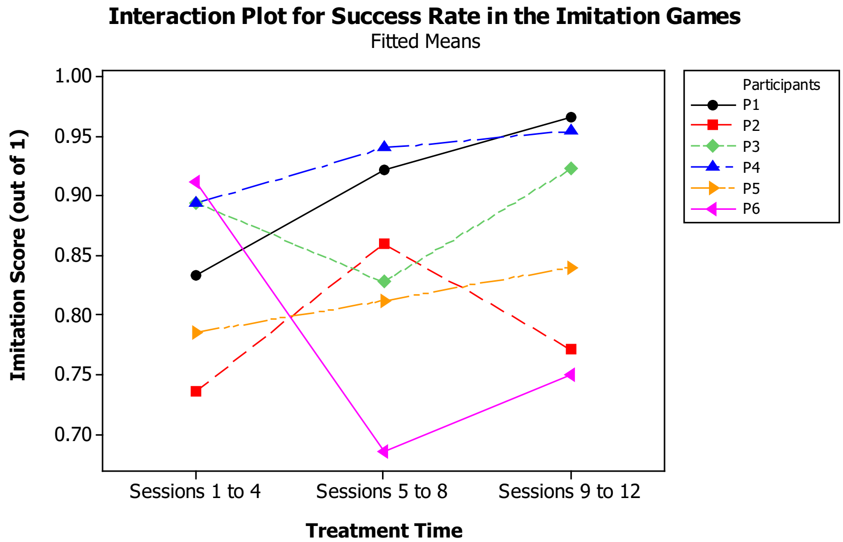


Fig. 9 Interaction plot for the percentage success rate in the imitation games considering the effect of treatment time and the participants individually (Condition 1)

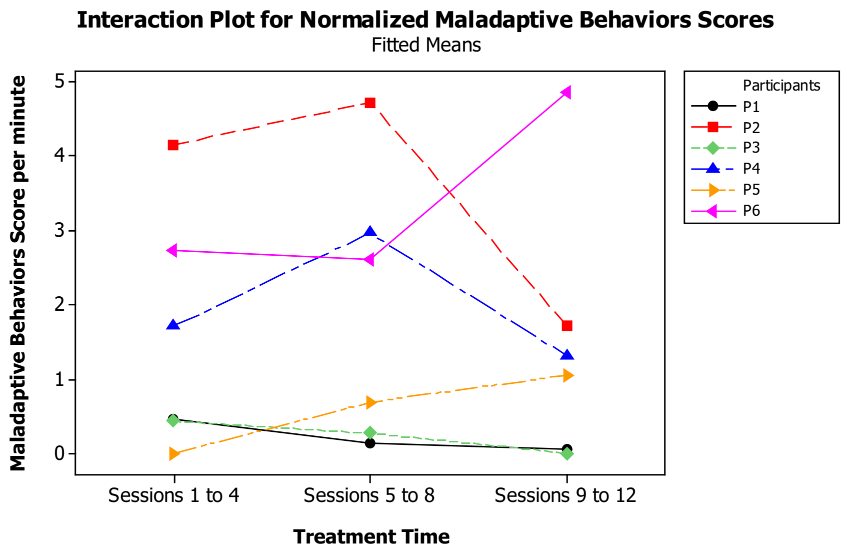


Fig. 10 Interaction plot for the normalized maladaptive behaviors scores considering the effect of treatment time and the participants individually (Condition 1)

estingly) P5-T showed almost no stereotyped behaviors and their other maladaptive behaviors during the sessions are negligible. P4-M had the almost fixed score of 2 stereotyped behaviors per minute in the sessions. Among the participants of this study, P6-H showed the most aggression maladaptive behaviors (such as attacking the robot/mother) in the sessions. The noticeable observations of this subsection happened for P5-T and P2-I. Unlike P5-T's behavioral problem in his real life, he was very motivated to behave maturely in the presence of Nima as he did not want to ruin friendship with the robot with inappropriate behavior and he showed negligible maladaptive behaviors. Moreover, although Fig. 10 does not show a monotonic trend in P2-I's maladaptive behaviors, his score in Sessions #9-12 was less than the first and the middle four intervention sessions. In Sects. 3.2.1 and 3.3, this observation is also investigated (and confirmed). We believe that the humanoid robot was able to attract the child's attention and enhance his motivation in participating in the games over time (in line with the findings of [42]). In other words, the robot's attendance and the vivacious educational-therapeutic games caused the subject with low-functioning autism to forget some of the internal problems and repetitive meaningless stereotyped behaviors. Although we did not design any games to control maladaptive behaviors directly, we observed a reduction of P2-I's echo and stereotyped behaviors throughout the robot-assisted program. Tapus et al. [47] also indicated the decrease in stereotyped behaviors of children with autism while interacting with the NAO humanoid robot.

3.1.5 Verbal Communications

One of the parameters investigated during the movies' content analysis was the number of Verbal Communication (VC) of the participants with the robot and the other individuals in the intervention room (Fig. 11a, b). This parameter is the

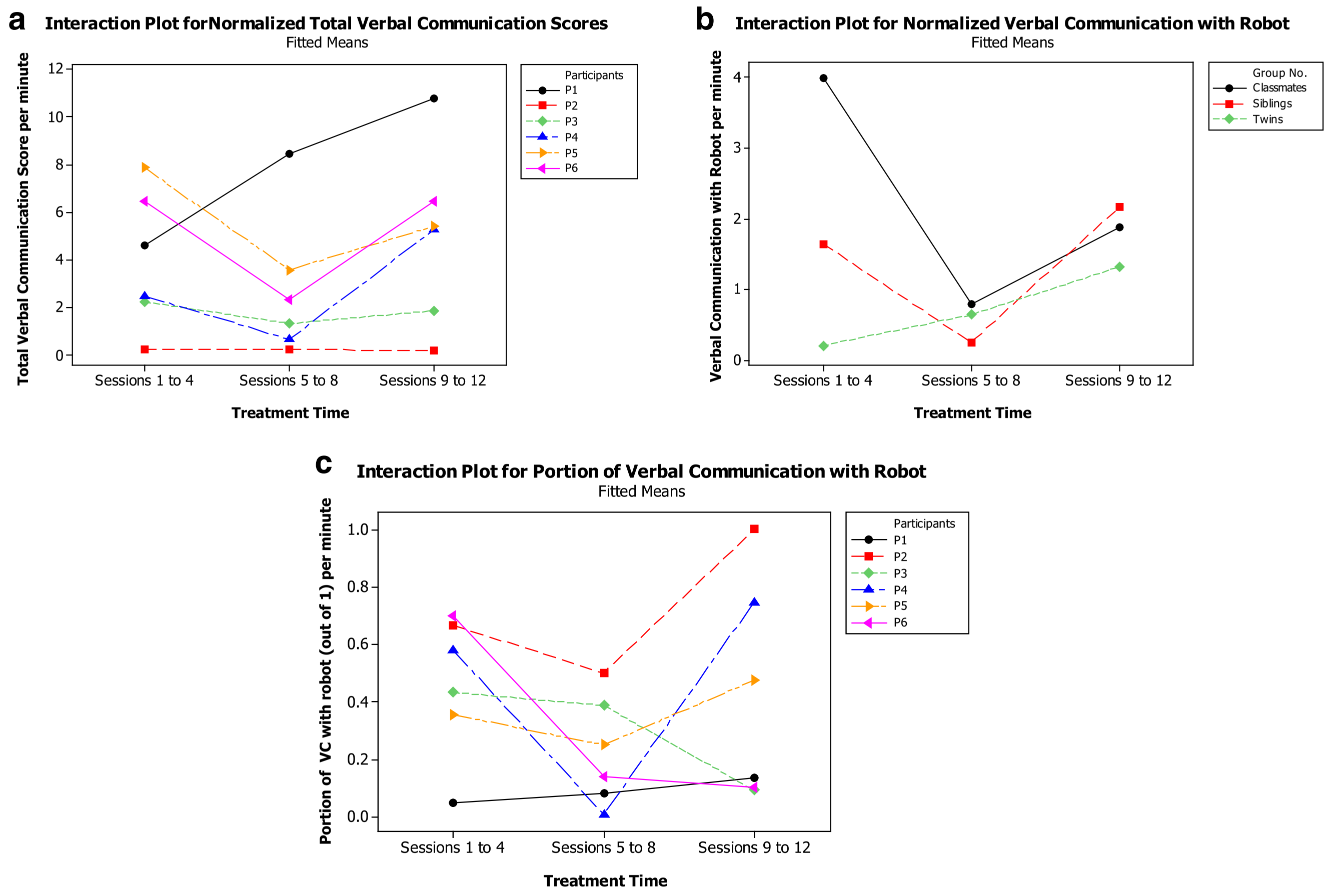


Fig. 11 Interaction plot for the **a** total verbal communication scores in Condition 1, **b** verbal communication with robot in Condition 2, and **c** portion of verbal communication with robot (out of 1) in Condition 1; considering the effect of treatment time and the participants/paired groups

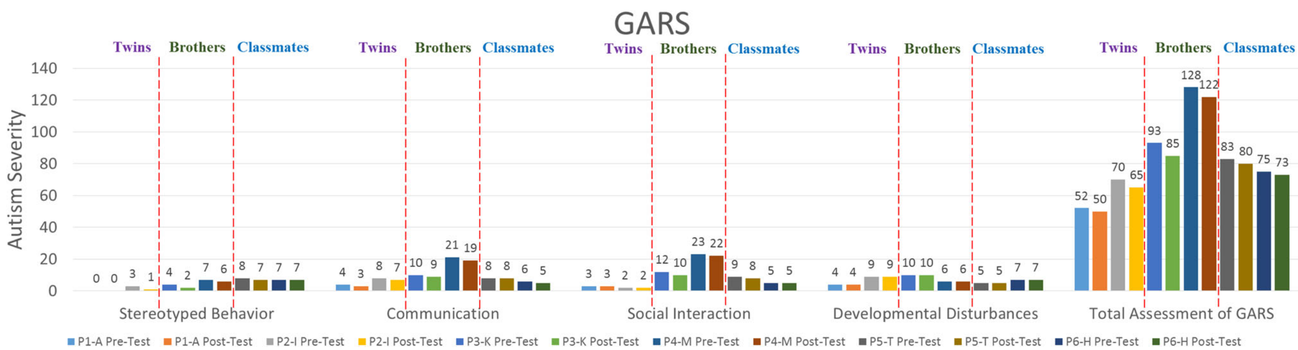


Fig. 12 GARS' subscales and overall scores (autism severity) of all six participants in the Pre-Test and Post-Test

number of situations of using a word, phrase, or a question to communicate with the robot, therapist, parents, or peer. Table 3 indicates that both of the factors "participants/paired-groups" and the "treatment time" have significant effects on the total VC scores. It can be concluded that the overall verbal communications of the participants/groups have been improved during the course of the program. Similar to the current research, improvement in communication/interaction skills of children with autism in robot-assisted interventions were confirmed in [2,6,15,16,42]. It seems that the middle-sessions' games provided less opportunities to show communication for the sibling and classmate pairs. As an example, we observed that the program's scenario motivated P4-M to have an overall increasing trend in verbal communication with the individuals in the class and as a result, the self-confidence of the child improved in comparison to the past. In particular, to investigate whether there was a significant relationship between P1-A's communication total scores and intervention time, linear regression analysis for the whole sessions was carried out using Minitab. Since the results of analysis of variance generated are $F = 5.64$, p value $= 0.045 < 0.05$, the linear model is statistically significant. Therefore, P1-A's verbal communication increased over the course of the intervention; and in the last four sessions he showed appropriate amounts of verbal communication with the therapist/parent, his twin brother, and the robot. In Sect. 3.3, the mentioned observation is also investigated (and confirmed).

Moreover, a detailed look into the scores of verbal communication with the robot(s) shows us that this parameter is significantly dependent to the paired-groups; and also has been increased (very close to being) significantly over time. Also, P1-A was the subject with the highest number of the verbal communications among the participants; however, less than 20% of his VCs were with the robot (Fig. 11c); therefore, considering the very low scores of P2-I in verbal communication (Fig. 11a), the overall performance of the twins-group with the robot was the least among the other two paired-groups due to the nonhomogeneous skills of the group members (Fig. 11b). Improvement in communication skills of ASDs with robots in such studies has been also reported in [6,42].

Interestingly, games I/3, J/3, I/2, J/2, and I/2 were selected as the best game by the twins, P3-K, P4-M, P5-T, and P6-H, respectively at the end of the program; which shows us different children's interest toward the designed games.

3.2 Questionnaires

3.2.1 GARS Questionnaire

The GARS questionnaire was filled in by the participants' mothers 1 week before and 1 week after the intervention sessions. The autism severity of the subjects are presented in Fig. 12. It should be mentioned that higher subscales/total scores indicate a higher autism severity.

A decreasing trend in the autism severity of the subjects from Pre- to Post-Test was reported for all six participants. The amount of decease in GARS scores however seemed to be different; and was much more for the Siblings group than the other two groups. According to Fig. 12, P3-K showed improvement in all three subscales of the GARS: stereotyped behaviors, communication, and social interaction during the course of the program.

An interesting point of the autism severity questionnaire, which may not have happened accidentally, is the observation of progress for all six participants in the subscale "communication" (outside of the study). This observation is in line with the video coders' reports of the subjects' communication. The robot-assisted intervention sessions seemed to have a promising effect on the subjects' socio-cognitive skills and as a result, their autism severity. It is a signal that the presence of the robots, as attractive communication tools, as well as being involved in group games could be possible ways to affect the communication skills of the children with autism even in their real life. The protocol of this study included many imitation and joint attention tasks; and it has been shown that improvement in imitation and joint attention

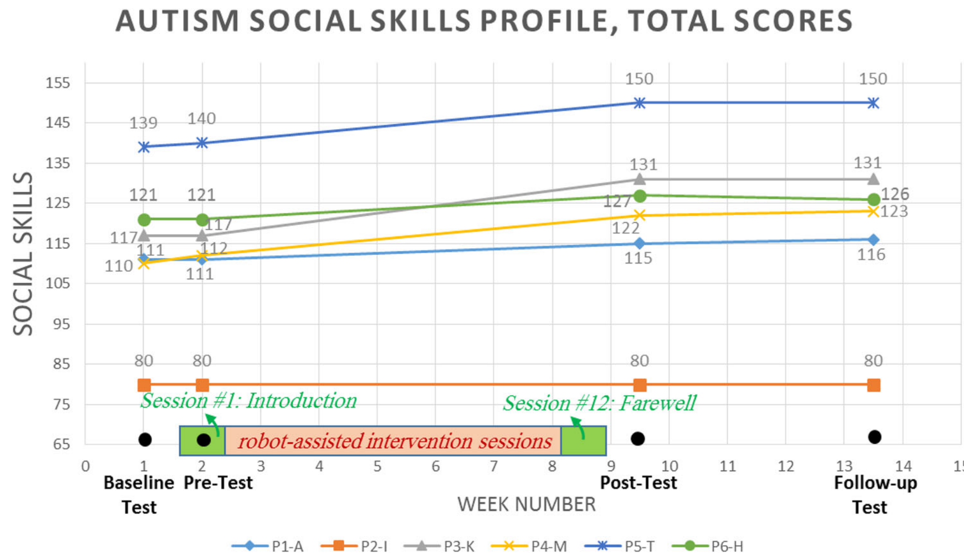


Fig. 13 Total ASSP scores of the participants in the Baseline-Test, Pre-Test, Post-Test, and Follow-up Test

skills can significantly affect social/communication skills of individuals with ASD [1–3,7,15–17].

According to Fig. 12, the stereotyped behaviors decreased for P2-I, P3-K, and P4-M from Pre-Test to Post-Test. Moreover, no stereotyped behaviors is reported for P1-A which is also in agreement with the observations of our video coders.

It should be noted that the questions supporting the Developmental Disturbances subscale in GARS are related to the first 36-month of the children's life; therefore, there is no change in the Pre-Test and Post-Test in that subscale. Among the participants, P3-K had the most problems in his life regarding the developmental disturbances subscale before the age of 3.

3.2.2 ASSP Questionnaire

The ASSP questionnaire are presented in Fig. 13. Based on the parents' reports, the social skills of all high-functioning participants improved somewhat during the robot-assisted sessions, while no improvement was reported for the low-functioning subject P2-I. A detailed overview of the ASSP gives us the following information: (a) "social reciprocity" did not change for either of the twins, but the other subjects showed improvement in it (with the overall changes' mean and standard deviation, 2.67 and 3.08, respectively) with small effect size for both conditions (Table 4); (b) "social participation/avoidance" increased for P1-A, P3-K, P4-M, and P6-H (total changes' mean: 1.83, changes' standard deviation: 3.06); and c) "detrimental social behaviors" increased for all high-functioning participants indicating that the nature of the designed protocol may have a positive effect on this subscale.

According to Table 4, we observed that the educational robot-assisted program improved social skills, social participation/avoidance, and detrimental social behaviors in our participants with high-functioning autism with close to being large effect sizes (i.e. Cohen's d $\sim$ 0.7) which is in line with the findings of [9,16,17,48,49].

A review on the principles and basic factors of comprehensive "reciprocal imitation" programs include: being face to face with the child, actions with toys, imitating body movements, postures, and vocalizations, as well as symbolic/imagination games can shed light on why children with autism showed improvements in social skills and its subscales [48]. It is observed that the designed games could still affect children's social interactions. To explain the findings of the current study, we should consider the first function of imitation, e.g. using body movements in order to develop social interaction and reciprocity with other people in society [50]. Another function of imitation is to make the child able to be aware of the interlocutor's activities and intentions which will lead to his/her social learning [51]. Imitation brings about increased amount of playing games, joint attention, and social accountability in children with autism [1]. Therefore, with such scenarios, we could positively affect the social interactions of the participants with high-functioning autism. The designed games were developed and conducted with the aim to increase their social interaction. Similarly, Feil-Seifer and Matarić, [6] successfully used the Bandit robot in robot-assisted interventions as a catalyst for social behavior and investigated the interactions of children with autism.

Moreover, comparison of the ASSP overall scores between the Post- and Follow-up Tests shows the retention and the stability of the robot-assisted program's impact on social skills of the participants.

3.3 Human Assessment

The participants were assessed by a clinical child psychologist 1 week before and 1 week after our robot-assisted

Table 4 Cohen's *d* effect size measurement between Pre-Test and Post-Test of the ASSP questionnaire' results

Questionnaire	Subscale	Cohen's d effect size between Pre- and Post-Test Scores	
		All 6 participants	5 high functioning children
ASSP	Overall scores	0.342	**0.704**
	Subscale 1: social reciprocity	0.230	0.389
	Subscale 2: social participation/avoidance	0.333	**0.694**
	Subscale 3: detrimental social behaviors	0.463	**0.698**

The numbers that represent the large or close to being large effect size (i.e. $\sim$ or > 0.8) are in bold in the table

program. The criteria for this assessment consisted of more than 25 items on self-help skills, social interaction, verbal communications, mathematical concept, motor skills, joint attention and some cognitive skills (based on ESCS[5] [1,52], regular imitation tests, ABA treatment, etc.).

P1-A Based on the psychologist reports, P1-A showed more admissible progress in verbal communications and joint attention skills than in other tested skills. However, he still has problems in verbal descriptions, storytelling, visual memory, and using pronouns in sentences making. His gross/fine imitation skills were totally acceptable in both Pre-Test and Post-Test.

P2-I Before starting our program, P2-I's obvious problems were in verbal and perceptuo-motor skills, he could not perceive most of the verbal items. He had also a deficit in 3-d spatial navigation as well as understanding simple concepts, practically. P2-I was languid and his muscles were loose He showed low instruction perception and considerable behavioral problems. Verbal communication is difficult for P2-I and he usually avoids speaking. According to the psychologist's report, P2-I made progress in instruction perceptions and cooperation, gross imitation and motor skills; however, he still hardly can play hopscotch. The future tasks for his motor skills could be massage of arms/feet and balancing exercise on the dominant foot with both closed/open eyes. He showed progress in joint attention skills and his stereotyped behaviors decreased in comparison to his past. Moreover, she reported that P2-I's other weaknesses are mental skills and verbal communications. P2-I's ability in using propositions depends on his 300-words' vocabulary. His family should concentrate on teaching semantic relations (i.e. subjects+verbs/objects/adjectives) to the low-functioning son. His deficit is in abstract concepts which need more complex processing.

P3-K The eye contact of P3-K in story-telling has been improved in comparison to the Pre-Test. His main deficit is still the exhibition of abnormal speech prosody. His progress in facial expressions and cognitive skills was fairly satisfactory. Regarding joint attention, the ceiling effect was observed for P3-K.

P4-M He has a problem in turn-taking group games even in the Post-Test. His verbal communication skills has been improved slightly since the Pre-Test; however, it is not still appropriate for his age. He spoke louder in the Post-Test, which could be the sign of his improvement in self-confidence.

P5-T As one of his behavioral problem, P5-T is still headstrong towards behavioral requests. Being in a new environment usually makes him anxious. He was occasionally screaming in the Pre-Test. Based on his medical report, P5-T did not have the tendency to start verbal communications with older people; however, in the Post-Test the psychologist reported some verbal initiations by the child. He successfully passed some of the fine imitation tasks (e.g. matching the ribbon/matches games) which he previously did not do (in the Pre-Test).

P6-H In the Pre-Test, the psychologist reported that P6-H's situation is not acceptable with regard to his age in alphabets/numbers/week days' perception, answering strangers, and finding similarities/differences in abstract concepts. Moreover, he had a deficit in eye contact (especially in descriptive item) and social communication. P6-H showed a problem in eye-hand coordination in fine motor skills as well as in the item of complex gross imitation. Having lack of cooperation, he was (intentionally) trying to waste the time of the assessment session in the Pre-Test. Putting his hands in his mouth, chewing his sleeves, and eliciting vocal echoes were other observations of the Pre-Test session. The overall overview was his behavioral problems exceeded his cognitive deficits. The most valuable observation for P6-H in the Post-Test was that he showed better social participation and elicited semi-matured behaviors/sentences while he was engaging in an effective communication. He performed appropriately in the task of following the objects in the book in the Post-Test; while he had a lack of cooperation in the same item in the Pre-Test. He still has a serious problem in counting. His eye contact and attention span got slightly better but is still very lower than usual. Regarding verbal

[5] Early Social Communication Scales; a comprehensive clinical measure of joint attention behaviors, behavioral requests, and social interaction behaviors of children.

communication skills, his current deficit is his inability in keeping the story topic while he is telling a story and trying to make the sentences as simple as he can. The instruction perception of P6-H is improved. The psychologist described some potential tasks to the family in order to control the behavioral problems of the child. He has still problems in auditory memory and reasoning.

The parts of the reports including the children's progress in joint attention and communication skills as well as the low-functioning subject's improvement in stereotyped behaviors are pretty in line with the observed data trend from the video coding and the GARS questionnaire presented in Sects. 3.1 and 3.2.1. We can also add that involving the paired-groups with autism in such programs may cause them to get out of their solitary inner world a bit; so they could experience some helpful social and communication situations (which has also been confirmed in [1,2,7,42]). Fortunately, no retrogression has been reported by the clinical child psychologist for any of the participants.

3.4 Interview with the Parents

We had an interview with the parents after our last clinical session. To find out more reliable evidence for this study, a delayed follow-up interview was performed with each parent in order to investigate the long-term effects of the robot-assisted program on the participants. The most interesting parts of the interview are as follows:

Twins "In contrast to their ABA classes, our kids showed inexplicable interest in taking part in imitation and turn-taking games. P1-A often danced like Nima at home and sang the robot's song. For the first time since their birth, we have seen the twins playing a meaningful turn-taking game together with the table-soccer at home. They never understood that robots' actions occurred because of commands sent by an operator to the robots". The mother stated, "We believed that robotic clinical intervention would have a positive effect on our children's social interaction and their communication toward each other during these 2 months; however, we did not expect a miracle in their progress! Bringing my children to this different intervention program, I think I am doing my mother's duties better than the past". In the delayed interview, the twins' mother said: "P1-A has put the pictures of the twins and robots in the shelf. He has a good memory of the robot-assisted interventions. Right now, he assumes that the robots were as some of his previous toys that he does not work with them anymore. P2-I does not talk about the robots; he is currently going to a regular school with a shadow caregiver and his behaviors are influenced by school's trainings. Recently, he has a new stereotyped behavior and move his hands as he is writing something".

Siblings "I believe that the self-confidence and motivation of P4-M has been improved somewhat; unlike his school assignments, he feels that he can do the robot's tasks correctly. During this program, we saw that P3-K wrote down in his notebook that he will have a class with Nima tomorrow! Both of my sons know that the robot has been controlled by the operator. They both usually do not start a verbal communication with their classmates in the school; therefore, they haven't talked about the robot and this program anywhere until I or their father ask them about the sessions". In the delayed interview, the mother stated that "Both of the siblings are trying more serious in their school assignments. P4-M sometimes shows headstrong behaviors; however we believe that's because he is approaching his age of puberty. Now, one of their dreams is to design such a robot in the future".

Classmates P5-T's mother said: "P5-T is not resistance to the education programs at all; however, I felt that he is more eager to take part in this class. He believed that he has something that the others do not have access to (i.e. the cute robot); and he sometimes showed off for that at school. Being deeply involved in the sessions and the games, I felt that my son feels more secured and closer to me". P6-H's mother stated: "My son is more cooperate in his ABA classes and this program in doing the items than at home; by threatening P6-H not to bring him to the robot's class, we sometimes tried to improve his obedience at home. He has a serious behavioral problem and I am really concern about my son's ability to learn". In the delayed interview, one of the mothers stated: "P5-T sometimes asked to join him in dancing and doing exercise. He thanked me for bringing him to the robots' classes". It should be noted that we could not find the mother of P6-H for the delayed interview.

One interesting point of this study observed for the twins was that the children selected the tele-operation game (I/3) as their preferred experiment. On the hand, they did not perceive the robots as being operated by a human during the rest of the sessions. We believe that this is due to ASDs' lack of generalization skills [1] and importantly, their impairment in theory of mind [53] which means looking at something from someone else's perspective, considering his/her feelings and thoughts.

The twins' mother's statement of "P1-A danced like Nima at home" needs to be clarified: it is common that the parents seeing their child with autism imitates his/her teacher's behaviors at home; if the child could perform the similar tasks (or the generalized version of them) in other appropriate situations, it means that the learning process has occurred for him/her; on the other hand, there is a possibility that some children with autism repeat those tasks over and over as a new stereotyped behaviors and use them irrelevantly to the situations which is not a positive event indeed. The twins' mother emphasized that she has not observed the taught skills in robotic classes as stereotyped behaviors for her children. Actually, similar to Sect. 3.3, no behavioral retrogression has been reported by the mothers for the subjects and the mother confirmed their previous statements/claims. Most of them thanked us again for making their children happy during the 3 months of the program.

3.5 Summary of the Results Section, Limitations, and Future Works

What stands out in the results obtained from the quantitative content analysis of the video records is *paired-groups' probable improvement in joint attention and verbal communication/interactions, and P2-I's decrease of autistic detrimental behaviors*. The participants were happy/eager to attend the robotic classes and some of them considered the robots as their close friends; which means that some kinds of relationship with the social robots have been started developing [54]. The ethical aspects of relations in human–robot interaction have been investigated in [54] and the authors mentioned that human–robot interactions are constructed based on the human–human interaction rules. Moreover, P6-H's lack of cooperation and maladaptive/aggressive behaviors affected his individual/paired-group performance of doing the tasks especially in Sessions #5–8. Among the participants, P2-I and P6-H's had more difficulties in the hard-level imitation games. In this study, the twins-group showed the most progress among the paired-groups. In addition, we also faced the ceiling effect for the siblings-group in comparison to the other two groups in the designed game.

The overall findings of this study show that using robots in treatment of ASDs could be quite effective. However, the effects seem to be different for children from different points on the autism spectrum. The overall observations of this study is similar to the findings of [23] which no unique significant change were observed for all of their participants with autism; while a potential for their robotic program to increase the responses/interaction initiations of children with autism has been reported. This research was a pilot case study to obtain proof for the concept of a short-term and compact interventions, and similar to the other case-study research [2,55], generalizing the findings would require further research with more participants/paired-groups. It should be considered that the progress in cognitive and social skills of the participants/paired-groups is based on the existence of two simultaneous factors: (a) the robots' role as a co-therapist in the intervention sessions, as well as (b) the nature of the designed therapeutic games. At the moment, we cannot separate the effects of each factor on the subjects' improvements individually. In order to have a comprehensive basic analysis of whether the games or the robots are more effective, future research could be done by replacing the robots with cartoon characters or humans during the same intervention scenarios. Quantifying the behavioral analysis of the participants was one of the noteworthy aspects of this study.

The focus of this study was on the effects of the overall designed scenarios (including all the games) on the participants. In future works, the effects of the robot's involvement level in game scenarios can be studied. Also, for investigating the joint attention and/or imitation skills of ASDs in detail, it is recommended to reduce the number/diversity of the games for future studies. Moreover, we are still unable to answer the following question of whether conducting the same scenarios with a human-therapist (without robots) would be more/less effective than the robot-assisted interventions; and to this end, multiple-baseline single-subject-design studies should be done in the future.

Although the most important limitations of our study were (a) the small number of the participants, (b) participants' maturation, (c) potential effects of the other classes outside our sessions, (d) unpredicted behaviors of the subjects (especially P2-I and P6-H) during the sessions, (e) engineering technical issues, and (f) the small number of the intervention sessions and limited number of quantitative data as the inputs for our ANOVA tests, the positive results signal potentials discussed in this preliminary exploratory study are promising and could shed light for continuing autism treatment as well as other applications of social robots in Iran [56,57].

4 Conclusion

Various assessment tools of this study made it possible for us to investigate the participants' behavioral changes inside and outside of the current program. As an overall finding of this study, based on the GARS questionnaire, the autism severity decreased for all subjects of this study. Moreover, the ASSP questionnaire's results showed that the total scores of social skill of the 5 high-functioning participants as well as its subscales, i.e. social reciprocity, social participation/avoidance, and detrimental social behaviors, improved during the almost 3 months of robot-assisted interventions. Although no change was reported for the social skills of the low-functioning twin by the mother (based on the ASSP), his stereotyped behaviors decreased during the intervention time. Due to the design of this research, the assessments' focus was comparing each child's/group's observed behaviors to his/its previous performance, but there were many limitations in quantifying the behavioral data as well as analyzing it statistically. Nevertheless, applying two-way ANOVA tests on the content analysis of the video records in Condition 2 indicated that significant increasing trends exist for the total gaze score toward the robot, joint attention, and verbal communication of the three paired groups. According to the analysis of variance in Condition 1, even the high-functioning participants of this research showed different potential significant behavioral progress. As one of the main limitations of this study, having homogenous participants on the autism spectrum is very difficult (and somewhat impossible); therefore, it was predictable that the intervention sessions affected different participant' behavioral change differently. The assessments of the children by the clinical child psychologist showed us some improvement for the subjects that were, interestingly, in line with the video coding results and the mothers' claims. It should be noted that because of the small number of the subjects in case studies, there are no strong claims on generalizing the findings to other children with autism.

Acknowledgements Our profound gratitude goes to the "Center for the Treatment of Autistic Disorders (CTAD)" and its psychologists for their contributions to the clinical trials with the children with autism. This research was funded by the "Cognitive Sciences and Technology Council" (CSTC) of Iran (http://www.cogc.ir/) (Grant No. 103). We also appreciate the Iranian National Science Foundation (INSF) for their complementary support of the Social & Cognitive Robotics Laboratory (http://en.insf.org/).

Funding This study was funded by the "Cognitive Sciences and Technology Council" (CSTC) of Iran.

Compliance with ethical standards

Conflict of interest Author Ali Meghdari has received research grants from the "Cognitive Sciences and Technology Council" (CSTC) of Iran. The authors Alireza Taheri, Minoo Alemi, and Hamidreza Pouretemad declare that they have no conflict of interest.

Ethical approval Ethical approval for the protocol of this study was provided by Iran University of Medical Sciences (#IR.IUMS.REC.1395. 95301469), and the certification for ABA and robot-assisted treatment with children with autism was received from the Center for the Treatment of Autistic Disorders (CTAD), Iran.

References

1. Pouretemad H (2011) Diagnosis and treatment of joint attention in autistic children. Arjmand Book, Tehran (in Persian)
2. Scassellati B, Admoni H, Mataric M (2012) Robots for use in autism research. Ann Rev Biomed Eng 14:275–294
3. Taheri AR, Alemi M, Meghdari A, PourEtemad HR, Basiri NM (2014) Social robots as assistants for autism therapy in Iran: research in progress. In: 2014 second RSI/ISM international conference robotics and mechatronics (ICRoM). IEEE, pp 760–766
4. Robins B, Dickerson P, Stribling P, Dautenhahn K (2004) Robot-mediated joint attention in children with autism: a case study in robot-human interaction. Interact Stud 5(2):161–198
5. Kozima H, Nakagawa C, Yasuda Y (2005) Interactive robots for communication-care: a case-study in autism therapy. IEEE international workshop on robot and human interactive communication, ROMAN 2005:341–346
6. Feil-Seifer D, Matarić MJ (2009) Toward socially assistive robotics for augmenting interventions for children with autism spectrum disorders. In: Experimental robotics: The eleventh international symposium. Springer, Berlin. doi:10.1007/978-3-642-00196-3_24
7. Taheri A, Meghdari A, Alemi M, Pouretemad H, Poorgoldooz P, Roohbakhsh M (2016) Social robots and teaching music to autistic children: myth or reality?. In: International conference on social robotics. Springer, pp 541–550
8. Taheri A, Alemi M, Meghdari A, Pouretemad H, Basiri NM, Poorgoldooz P (2015) Impact of humanoid social robots on treatment of a pair of Iranian autistic twins. In: International conference on social robotics. Springer, pp 623–632
9. Khosla R, Nguyen K, Chu MT (2015) Socially assistive robot enabled home-based care for supporting people with autism. In: PACIS, p 12
10. Liu C, Conn K, Sarkar N, Stone W (2008) Online affect detection and robot behavior adaptation for intervention of children with autism. IEEE Trans Robot 24(4):883–896
11. Duquette A, Michaud F, Mercier H (2008) Exploring the use of a mobile robot as an imitation agent with children with low-functioning autism. Auton Robots 24(2):147–157
12. Mavadati SM, Feng H, Gutierrez A, Mahoor MH (2014) Comparing the gaze responses of children with autism and typically developed individuals in human–robot interaction. In: 2014 14th IEEE-RAS international conference on humanoid robots (humanoids), pp 1128–1133
13. Kajopoulos J, Wong AHY, Yuen AWC, Dung TA, Kee TY, Wykowska A (2015) Robot-assisted training of joint attention skills in children diagnosed with autism. International conference on social robotics. France. Springer, Paris, pp 296–305
14. Boccanfuso L, Barney E, Foster C, Ahn YA, Chawarska K, Scassellati B, Shic F (2016) Emotional robot to examine different play patterns and affective responses of children with and without ASD. In: 2016 11th ACM/IEEE international conference on human–robot interaction (HRI). IEEE, pp 19–26
15. Pioggia G, Ferro M, Sica ML, Dalle Mura G, Casalini S, De Rossi D, Muratori F (2006) Imitation and learning of the emotional behaviour: towards an android-based treatment for people with autism. In: Proceedings of sixth international conference on epigenetic robotics: modeling cognitive development in robotic systems
16. Robins B, Dautenhahn K, Te Boekhorst R, Billard A (2005) Robotic assistants in therapy and education of children with autism: can a small humanoid robot help encourage social interaction skills? Univers Access Inf Soc 4(2):105–120
17. Werry I, Dautenhahn K, Ogden B, Harwin W (2001) Can social interaction skills be taught by a social agent? The role of a robotic mediator in autism therapy. In: Proceedings of the 4th international conference on cognitive technology: instruments of mind, August 06–09, 2001, pp 57–74. doi:10.1007/3-540-44617-6_6
18. Salvador M, Silver S, Mahoor M (2015) An emotion recognition comparative study of autistic and typically developing children using the Zeno robot. In: IEEE on international conference on robotics and automation, Seattle, USA
19. Burack JA, Volkmar FR (1992) Development of low-and high-functioning autistic children. J Child Psychol Psychiatry 33(3):607–616
20. Kean JM (1975) The development of social skills in autistic twins. N Z Med J 81(534):204–207
21. Mazefsky CA, Goin-Kochel RP et al (2008) Genetic and environmental influences on symptom domains in twins and siblings with autism. Res Autism Spectr Disord 2(2):320–331
22. Ozonoff S, Young GS, Carter A, Messinger D, Yirmiya N et al (2011) Recurrence risk for autism spectrum disorders: a Baby Siblings Research Consortium Study. Pediatrics 128(3):e488–e495
23. Huskens B, Palmen A, Van der Werff M, Lourens T, Barakova E (2015) Improving collaborative play between children with autism spectrum disorders and their siblings: the effectiveness of a robot-mediated intervention based on Lego therapy. J Autism Dev Disord 45(11):3746
24. Wong CCY, Meaburn EL, Ronald A, Price TS, Jeffries AR et al (2014) Methylomic analysis of monozygotic twins discordant for autism spectrum disorder and related behavioral traits. Mol Psychiatry 19(4):495–503

25. Mitchell SR, Reiss AL, Tatusko DH, Ikuta I, Kazmerski DB, Botti JAC, Kates WR (2009) Neuroanatomic alterations and social and communication deficits in monozygotic twins discordant for autism disorder. Am J Psychiatry 166(8):917–925
26. Ho A, Todd RD, Constantino JN (2005) Brief report: autistic traits in twins vs. non-twins–a preliminary study. J Autism Dev Disord 35(1):129–133
27. Constantino JN, Todd RD (2003) Autistic traits in the general population: a twin study. Arch Gen Psychiatry 60(5):524–530
28. Folstein S, Rutter M (1977) Infantile autism: a genetic study of 21 twin pairs. J Child Psychol Psychiatry 18(4):297–321
29. Steffenburg S, Gillberg C, Hellgren L, Andersson L, Gillberg IC, Jakobsson G, Bohman M (1989) A twin study of autism in Denmark, Finland, Iceland, Norway and Sweden. J Child Psychol Psychiatry 30(3):405–416
30. Hilton JC, Seal BC (2007) Brief report: comparative ABA and DIR trials in twin brothers with autism. J Autism Dev Disord 37(6):1197–1201
31. Kim S, Clarke E (2015) Case study: an iPad-based intervention on turn-taking behaviors in preschoolers with autism. Behav Dev Bull 20(2):253
32. https://www.aldebaran.com/en (March 2016)
33. http://www.hansonrobotics.com/ (March 2016)
34. Gilliam JE (1995) Gilliam autism rating scale. ProEd, Austin, TX
35. Ashburner J, Ziviani J, Rodger S (2008) Sensory processing and classroom emotional, behavioral, and educational outcomes in children with autism spectrum disorder. Am J Occup Therapy 62(5):564–573
36. Ahmadi SJ, Safari T, Hemmatian M, Khalili Z (2012) Exploring the criterion of diagnosing autism (GARS). J Res Cognit Behav Sci 1(1):87–104 (in Persian)
37. Bellini S, Peters JK (2008) Social skills training for youth with autism spectrum disorders. Child Adolesc Psychiatr Clin N Am 17(4):857–873
38. Cohen J (1988) Statistical power analysis for the behavioral sciences, 2nd edn. Lawrence Earlbaum Associates, Hillsdale, NJ
39. Riff D, Lacy S, Fico F (2014) Analyzing media messages: using quantitative content analysis in research. Routledge, Abingdon
40. Rimland B, Edelson S (2000) Autism treatment evaluation checklist: statistical analyses. Autism Research Institute, California
41. Minitab 17 Statistical Software [Computer software] (2010) Minitab, Inc., State College, PA. www.minitab.com
42. Kozima H, Nakagawa C, Yasuda Y (2007) Children-robot interaction: a pilot study in autism therapy. Prog Brain Res 164:385–400
43. Warren ZE, Zheng Z, Swanson AR, Bekele E, Zhang L, Crittendon JA, Sarkar N (2015) Can robotic interaction improve joint attention skills? J Autism Dev Disord 45(11):3726–3734
44. Whalen C, Schreibman L (2003) Joint attention training for children with autism using behavior modification procedures. J Child Psychol Psychiatry 44(3):456–468
45. Ingersoll B, Gergans S (2007) The effect of a parent-implemented imitation intervention on spontaneous imitation skills in young children with autism. Res Dev Disabil 28(2):163–175
46. Ingersoll B, Schreibman L (2006) Teaching reciprocal imitation skills to young children with autism using a naturalistic behavioral approach: effects on language, pretend play, and joint attention. J Autism Dev Disord 36(4):487–505
47. Tapus A, Peca A, Aly A, Pop C, Jisa L, Pintea S, David DO (2012) Children with autism social engagement in interaction with Nao, an imitative robot–a series of single case experiments. Interact Stud 13(3):315–347
48. Ingersoll B (2008) The social role of imitation in autism: implications for the treatment of imitation deficits. Infants Young Child 21(2):107–119
49. Ingersoll B, Walton K, Carlsen D, Hamlin T (2013) Social intervention for adolescents with autism and significant intellectual disability: initial efficacy of reciprocal imitation training. Am J Intellect Dev Disabil 118(4):247–261
50. Carpenter M, Pennington BF, Rogers SJ (2002) Interrelations among social-cognitive skills in young children with autism. J Autism Dev Disord 32(2):91–106
51. Uzgiris IC (1999) Imitation as activity: its developmental aspects. In: Nadel J, Butterworth G (eds) Imitation in infancy. Cambridge University Press, Cambridge, pp 186–206
52. Mundy P, Delgado C, Block J, Venezia M, Hogan A, Seibert J (2003) Early social communication scales (ESCS). University of Miami, Coral Gables, FL
53. Kazdin AE (2011) Single-case research designs: methods for clinical and applied settings. Oxford University Press, Oxford
54. Baron-Cohen S (2001) Theory of mind in normal development and autism. Prisme 34(1):74–183
55. Graaf M (2016) An ethical evaluation of human-robot relationships. Int J Soc Robot 8(4):589–598
56. Alemi M, Ghanbarzadeh A, Meghdari A, Moghadam LJ (2016) Clinical application of a humanoid robot in pediatric cancer interventions. Int J Soc Robot 8(5):743–759
57. Alemi M, Meghdari A, Ghazisaedy M (2015) The impact of social robotics on L2 learners' anxiety and attitude in English vocabulary acquisition. Int J Soc Robot 7(04):523–535

Alireza Taheri is a Ph.D. candidate in Mechanical Engineering with emphasis on Social Robotics at Sharif University of Technology, Tehran, Iran. During 2015–2017, he spent a total of 1-year sabbatical as a research scholar at Yale University and University of Denver, USA. Email: taheri@mech.sharif.edu.

Ali Meghdari is a Professor of Mechanical Engineering at Sharif University of Technology (SUT) in Tehran. Professor Meghdari has performed extensive research in various areas of robotics; social and cognitive robotics, mechatronics, bio-robotics, and modelling of biomechanical systems. He has been the recipient of various scholarships and awards, the latest being: the 2012 Allameh Tabatabaei distinguished professorship award by the National Elites Foundation of Iran (BMN), the 2001 Mechanical Engineering Distinguished Professorship Award from the Ministry of Science, Research & Technology (MSRT) in Iran, and the 1997 ISESCO Award in Technology from Morocco. He is currently the Director of the Centre of Excellence in Design, Robotics and Automation (CEDRA), an affiliate member of the Iranian Academy of Sciences (IAS), and a *Fellow* of the American Society of Mechanical Engineers (ASME). He is also the Editor of Transactions on Socio-Cognitive Engineering, *Scientia-Iranica International Journal*. Email: meghdari@sharif.edu. Home page: http://meghdari.sharif.edu.

Minoo Alemi received her Ph.D. in Applied Linguistics (TEFL) from Allameh Tabataba'i University in 2011. She is an Assistant Professor and Division Head of Applied Linguistics in Islamic Azad University, West Tehran Branch. She is the founder of Robot-Assisted Language Learning (RALL), and the co-founder of Social Robotics in Iran which she achieved as a Post-Doctoral research associate at the Social & Cognitive Robotics Laboratory of Sharif University of Technology. Her areas of interest include discourse analysis, interlanguage pragmatics, social robotics, and RALL. Dr. Alemi has been the recipient of various teaching and research awards from Sharif University of Technology, Allameh Tabataba'i University, Islamic Azad University, and International Conference on Social Robotics (ICSR2014), Sydney-Australia. She has published over 75 papers and books in refereed national and international conferences, and journals. She is also on the editorial boards of the British Journal of Educational Technology (BJET), and the *Scientia-Iranica*: Transactions on Socio-Cognitive Engineering. Email: alemi@sharif.edu.

Hamidreza Pouretemad is a Professor of Clinical Neuropsychology and founder and Head of the Institute for Cognitive and Brain Sciences, Shahid Beheshti University, Tehran, Iran. He is also the founder of Iranian Society for Cognitive Science and Technology, and Director of the Center for the Treatment of Autistic Disorders (CTAD) in Tehran, Iran. He has done extensive research on Children with Autism Spectrum Disorder. He has published over 150 papers in refereed national and international conferences, and journals. He is also on the editorial board of the *Scientia-Iranica*: Transactions on Socio-Cognitive Engineering. Email: h-pouretemad@sbu.ac.ir.