

**Please cite the Published Version**

Ali, Sara, Mehmood, Faisal, Dancey, Darren, Ayaz, Yasar, Khan, Muhammad Jawad, Naseer, Noman, Amadeu, Rita De Cassia, Sadia, Haleema and Nawaz, Raheel  (2019) An Adaptive Multi-Robot Therapy for Improving Joint Attention and Imitation of ASD Children. IEEE Access, 7. pp. 81808-81825. ISSN 2169-3536

**DOI:** <https://doi.org/10.1109/access.2019.2923678>

**Publisher:** Institute of Electrical and Electronics Engineers (IEEE)

**Version:** Published Version

**Downloaded from:** <https://e-space.mmu.ac.uk/623505/>

**Usage rights:**  [Creative Commons: Attribution 3.0](https://creativecommons.org/licenses/by/3.0/)

**Additional Information:** This is an Open Access article published in IEEE Access. (c) 2019 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other users, including reprinting/ republishing this material for advertising or promotional purposes, creating new collective works for resale or redistribution to servers or lists, or reuse of any copyrighted components of this work in other works."

**Enquiries:**

If you have questions about this document, contact [openresearch@mmu.ac.uk](mailto:openresearch@mmu.ac.uk). Please include the URL of the record in e-space. If you believe that your, or a third party's rights have been compromised through this document please see our Take Down policy (available from <https://www.mmu.ac.uk/library/using-the-library/policies-and-guidelines>)

Received May 18, 2019, accepted May 31, 2019, date of publication June 19, 2019, date of current version July 8, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2923678

# An Adaptive Multi-Robot Therapy for Improving Joint Attention and Imitation of ASD Children

SARA ALI<sup>1</sup>, FAISAL MEHMOOD<sup>1</sup>, DARREN DANCEY<sup>2</sup>, YASAR AYAZ<sup>1,3</sup>,  
MUHAMMAD JAWAD KHAN<sup>1</sup>, NOMAN NASEER<sup>4</sup>,  
RITA DE CASSIA AMADEU<sup>1</sup>, HALEEMA SADIA<sup>5</sup>,  
AND RAHEEL NAWAZ<sup>1,2</sup>

<sup>1</sup>School of Mechanical and Manufacturing Engineering (SMME), National University of Sciences and Technology, Islamabad, Pakistan

<sup>2</sup>School of Computing, Mathematics and Digital Technology, Manchester Metropolitan University, Manchester, U.K.

<sup>3</sup>National Center of Artificial Intelligence (NCAI), Pakistan

<sup>4</sup>Department of Mechatronics, Air University, Islamabad, Pakistan

<sup>5</sup>Autism Resource Centre, Shifa International Hospitals Ltd., Islamabad, Pakistan

Corresponding author: Raheel Nawaz (R.Nawaz@mmu.ac.uk)

**ABSTRACT** Robot-mediated therapies for autism spectrum disorder (ASD) have shown promising results in the past. We have proposed a novel mathematical model based on an adaptive multi-robot therapy of ASD children focusing on two main impairments in autism: 1) joint attention and 2) imitation. Joint attention intervention is based on three different least-to-most (LTM) cues, whereas the adaptive imitation module uses joint attention for activation of the robot. The proposed model uses a multi-robot system as a therapist without any external stimuli (from the environment) to improve the skills of the ASD child. Another novel aspect of this paper is the deployment of a multi-robot system for introducing the ASD child to the concept of multi-person communication. This is particularly useful as, unlike humans, robots can be more consistent and relatively immune to fatigue. Two different therapies of human–robot interaction (i.e., with and without inter-robot communication) have been conducted. The model has been tested on 12 ASD children, eight sessions for each intervention over a period of six months. The effectiveness of the model is validated by analyzing the cognitive state of the brain before and after the intervention with electroencephalogram (EEG) neuroheadsets. Moreover, results obtained using the childhood autism rating scale (CARS) to measure the effectiveness of therapy also support the conclusions firmly. The statistical results with the  $p$ -value =  $3.79E-07 < 0.05$  and the  $F$  value =  $23.93 > 3.28$  show reliability and significance of the data. The results strongly indicate significant improvements in both modules, along with a notable improvement in multi-communication skills of the participating children.

**INDEX TERMS** Autism spectrum disorder (ASD), imitation, joint attention, multi-robot.

## I. INTRODUCTION

Autism spectrum disorder (ASD) is a chronological condition that causes an impairment in social interaction, developmental language and communication skills in children. According to a survey, ASD is considered as a neurodevelopmental disorder that affects 1 out of 68 neonates [1]. In daily life, the autistic children may not be able to interact socially or express themselves to others that may result in frustrated behavior [2], [3]. One among many issues in autistic children is lack of focus or joint attention (JA) towards verbal or non-

verbal communication stimuli given by others [4]. A child with ASD lacks a tendency to establish visual coordination with another person [5]. Research has shown that there are two main divisions in joint attention: (i) the response to joint attention (RJA), and (ii) initiation of joint attention (IJA). RJA is shifting of the visual attention following the cues like pointing and gaze whereas IJA refers to seeking others attention using one's own gestures and gaze [6]. Considering the same narrative, several researches have been carried out over past two decades in early identification and rehabilitation of ASD patients. It was preempted that the most common approach for treatment of ASD children is through psychiatric therapy. A psychologist examines the actions of the child

to identify the level of autism. Based on the spectrum level of the child, different cognitive therapies are implemented to improve the condition of the child [7]. Recently, robots are being involved in these cognitive behavioral therapies to enhance the focus and interest of the autistic child.

The research conducted on remedial measures for ASD using human-robot interaction (HRI) has shown that majority of the individuals suffering from autism are more inclined towards robots rather than human therapists [8]. In case of robotics therapy, joint attention [9], ability to imitate [10], verbal communication [11] and social activities [12] are mostly targeted to measure the improvement in ASD. Humanoid robots are gaining more attention for autism therapy as they are controllable, accurate, low cost and adaptive to environment [13], [14]. The current research trends are inclined towards development of novel robot-based therapies due to the inquisition of ASD children in robot [10], [12]. The child's engagement is a key prerequisite to improve the adaptive ability of robots in intervention [13]. Most of the research done in robot-based therapeutic measures focuses on the physical features of the robot [15], control architecture [15], different evaluation criteria [16], and several HRI based algorithms [17]. A child's behavior may vary according to the size, shape and looks of the robots as when interacting with robots, its appearance matters to the subject [18]. Research has proved that clinical use of robots is helpful towards eliciting positive social behavior of an ASD child [19].

There are two types of robots that are used in ASD intervention: anthropomorphic [11], [20] and zoomorphic [21], [22]. Zoomorphic robots are animal like robots used for studying the behavior of an ASD child. One such example of a zoomorphic robot is "Keepon" robot that is famous for positive social interaction with ASD children because of its cute and simple appearance [23].

Robots that look like human are called as anthropomorphic robots. These robots are recently being used for research involving social interaction and facilitating collaborative play. Examples of these are "KASPAR" [10] and NAO [24].

Anthropomorphic robots are specially preferred for developing social skills in autistic children [25]. Despite having functional limitations, robots having physical appearance close to a humans can play a vital role in significantly improving the child's behavior [26], [15]. The interventions using these robots have proven to be more successful if they address the core deficits of ASD rather than choosing free play as mode of interaction [24]. The distinguished features of the robots are their high repeatability and willingness to interact without any complaints and fatigue [27].

Perception in autism is also discussed in latest Japanese research using robots. It is based on theory of mind and [28] is discussed via using different animate and inanimate entities. Non superiority visual processing of autistic children over typically developed children has been reported in [29] cognitive wise. In perception, humanness nature of the robot is also important which significantly affects [30].

Features which are good to increase the child-robot interaction are presented in [31]. The engagement of autistic children has been measured during the occipital therapy and the relationship between task driven valance and arousal conditions has also been studied [32].

An autism diagnostic protocol based on Autism Diagnostic Observation Schedule (ADOS) using a NAO humanoid robot has been presented in [34], [35]. However our proposed model represents an adaptive therapy using multi-robot system for improvement in joint attention and imitation of an ASD child. Table 1 shows the differences between our current research and previously work done in [33], [34].

Moreover another concern in application of such systems is the end users' preference for evidence based practice (EBP) which is not generally catered for in robotic therapy solutions [20].

### A. CONTRIBUTION

Three important contributions of this research are: 1) Design and development of a single mathematical model for adaptive multi-robot based therapy of ASD children for both LTM-based joint attention as well as imitation. 2) Validation and effectiveness of MRIS system based on user study using CARS scale. 3) Notable improvement in multi interaction of an ASD child.

The multi-robot based adaptive model presented in this article satisfies the concern of EBP called *MRIS (Multi-robot-mediated Intervention System)*. These multi-robots act as non-human partners in order to improve the social communication skills between multiple persons at the same time. Moreover the robots themselves act as therapist as well as the stimulator for an intervention without the use of any body worn sensor during intervention. Based on the results of intervention for the improvement of joint attention and imitation, it has proven to be the robot-mediated interventions (RMIs) as an evidence-based practice (EBP) in autism. This is achieved using a variety of sensors. First, instead of recording data manually, sensors have been integrated into the system so as to ensure the correctness of results and avoidance of human error. Moreover the results from EEG headset before and after intervention also verify the success. The intervention was planned in a way that all participants took part in the therapies. Furthermore it was ensured that sensors should not touch the body during intervention as it may make the ASD child feel uncomfortable [11]. The ultimate aim of this research is to find the parameters in collaboration with clinical experts that can improve the multi communication skills of an ASD child using adaptive robotic interventions.

### II. MRIS ARCHITECTURE

Our MRIS architecture as shown in Fig. 1 is based on the model proposed by Zheng *et al.* [35]. In the previous research [35], the model is adaptive for JA module only. Moreover the previous model does not focus on interaction of an ASD child with multiple agents simultaneously. Whereas, our proposed model introduces

TABLE 1. Comparison between existing and proposed models.

Existing research models in [35] and [36]	Proposed model
Presented robot assisted autism therapy for diagnosis of autism using joint attention and imitation tasks	Presenting robots assisted therapy for improvement of joint attention and imitation
Non-adaptive	Completely adaptive
Diagnosing autism	Improving social impairments of Autism
Uses Autism Diagnostic Observation Schedule (ADOS)	Uses Childhood Autism Rating Score (CARS)
ALFaceDetection: Uses face detection for eye contact of child- a major drawback	Used ALGazeAnalysis: Proper eye contact detection along with face detection. Moreover KINECT was integrated for head / face tracking ensuring direction of the face
In case of no response from ASD child, robot simply increases sentences to get attention	In our proposed therapy is based on increasing the effectiveness of stimuli i.e. LTM based approach
Two robots are used one active and one passive, Order of stimuli: Waving, flashing leds, and sounds making (neither LTM nor MTL)	Both are active robots and continuously noticing the responses. Order of stimuli: Rasta cues, speaking and waving (LTM model)
Offline tracking of object to estimate the moves of the ASD child	Online / real time tracking of joints of ASD child's skeleton to estimate the accuracy of the action in imitation.
4 ASD children, number of experiments not specified	12 ASD children. Multiple sessions done for each experiment
Vocal recognition- Reduced performance	Vocalization not used because of such issues.
Used vocalization utterances for classification based on voice	Used joint attention to measure the response to vocalization (speech task of robot)
Dealing in believe space (3D, no ASD, HF ASD, LF ASD)	Dealing with minimal and mild cases of autism

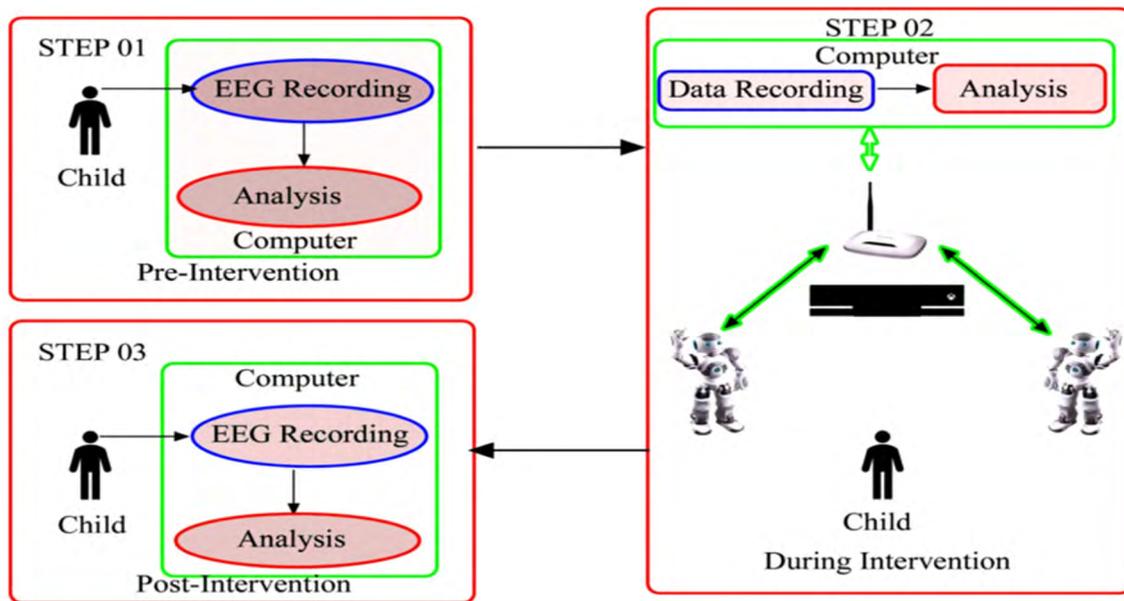
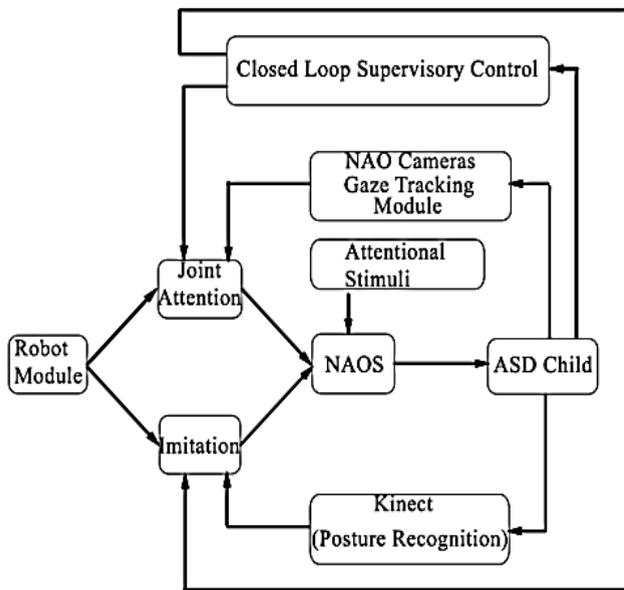


FIGURE 1. The 3-step system architecture explaining the robots and child arrangement in intervention area. The dotted line shows the inter-robot communication done in second phase of experimentation.

two adaptive modules i.e. for joint attention as well as imitation to improve the multicomunication of an ASD child. These two modules are discussed in detail in following sections.

**A. LTM-BASED JOINT ATTENTION PROTOCOL**

The interaction protocol for joint attention of MRIS uses least-to-most (LTM) cues as shown in Fig. 2. LTM has been used extensively as a tool for screening and diagnostics



**FIGURE 2.** Design Protocol of joint attention and imitation module. The joint attention module is based on LTM protocol. Both of these modules use closed loop supervisory control.

of ASD [36]. The child is introduced to the least intruding stimulus. If required the child is moved to the next level that is the more prominent intruding stimulus than the previous one [37], [38]. Robot mediated interventions have been using this protocol for teaching imitation skills to ASD children [39]. LTM protocol assists only when required.

Our designed protocol is based on three steps: 1) Visual Cues: In current joint attention, the first protocol is of visual cues. Two types of visual cues are developed: 1) “Rasta” (changing eye color of robot in cyclic manner) and “Blinking”, considered as least intruding stimulus. 2) Speech cues with these visual cues are added in the second protocol. Speech cues added are “Hi” and “Hello”.

These speech cues are more prominent hint for the ASD child compared to visual cues [38]. 3) Motion cues: Level three comprising of visual, speech and motion cues all combined together. The motion cues added are “Move forward”, “Move backward”, “Stand-up”, and “Sit-down”. As it can be seen the cues are ordered as per LTM approach [35].

While previous researches have only used a single robot with non-adaptive model for improvement of either joint attention or imitation in ASD child. We have introduced the first mathematical model based on multi-robots for improvement in both joint attention as well as imitation based on joint attention. This model can be used to improve joint attention and imitation skills in ASD children with the help of prompt only when required.

#### 1) NETWORKING PROTOCOL FOR JOINT ATTENTION

Networking protocol for LTM-based joint attention is shown in Fig. 3. Two transmission control protocol (TCP) servers are implemented in the control computer. The control computer

#### Algorithm 1 MRIS-LTM Joint Attention Model

Robot\_Action\_List = {“Visual”, “Visual + Speech”, “Visual + Speech + Motion”}

**Initialization:**  $n = 1$ , Count = 0, Index = 1;

**Load PQ** (form Subject.xlsx) **OR**  $PQ = \{\}$  (First Time)  
[Value, Index] = **Max** (PQ)

**Step 1:**  $RA(n) = \text{Robot\_Action}(\text{Robot\_Action\_List}(\text{Index}))$

$RB(n) = \text{Robot\_Behavior}(RA(n))$

$RESP(n) = \text{Participant\_Joint\_Attention}()$ ;

**IF**( $RESP(n) == \text{ExpResp}$ )

Reward

**INSERT** (PQ,  $RESP(n)$ , Index)

**GO TO STEP03**

**Step 2:**

Index = **Next\_Robot\_Action** ( $RESP(n)$ , PQ)

$n++$ ;

**Go to step 01**

**Step 03:**

**Write** (**SORT**(PQ), Subject.xlsx) & Terminates

communicates with NAO robots for activation of stimuli and feedback of data through the TCP servers during the experiment. The control modules are 1) Eye contact module which records the eye contact duration of the child and 2) Reinforcement stimuli module that gives cues for the joint attention module. The two cues given in this are rasta and blinking, as already discussed.

Both modules run in parallel. In Fig. 3 the modules are represented by numbers. Reinforcement stimuli modules are represented by  $C_{11}$  and  $C_{12}$  whereas eye contact modules are represented by  $C_{21}$  and  $C_{22}$ . Server sends commands to both clients i.e. both NAO robots at the same time via router and receives feedback during the experiment as shown in Fig. 3. This holds true for both modules. Moreover file writing is done in two separate files.

#### 2) MATHEMATICAL MODEL FOR JOINT ATTENTION

Various prompts are used in the intervention modeling for therapies. The prompts used are usually verbal and motion cues from the robot along with some environmental factors. The environmental factors are usually the medium towards which the robot points e.g. LCD screen etc. [35]. In this research no external factor has been introduced. The prompt cues were given by the robot in LTM order i.e. visual, speech and motion cues. These cues differ in level of complexity for obtaining the child’s response. This model starts with the least prominent cue for measuring the joint attention of the child. {V}, {S} and {M} are the libraries used for representing different reinforcement stimuli where {V} represents visual, {S} is for speech and {M} represents motion. By combining these libraries we get different stimuli ranging from least to most in its effectiveness e.g. using visual stimulus stand alone has least effect as compared to using it in combination with

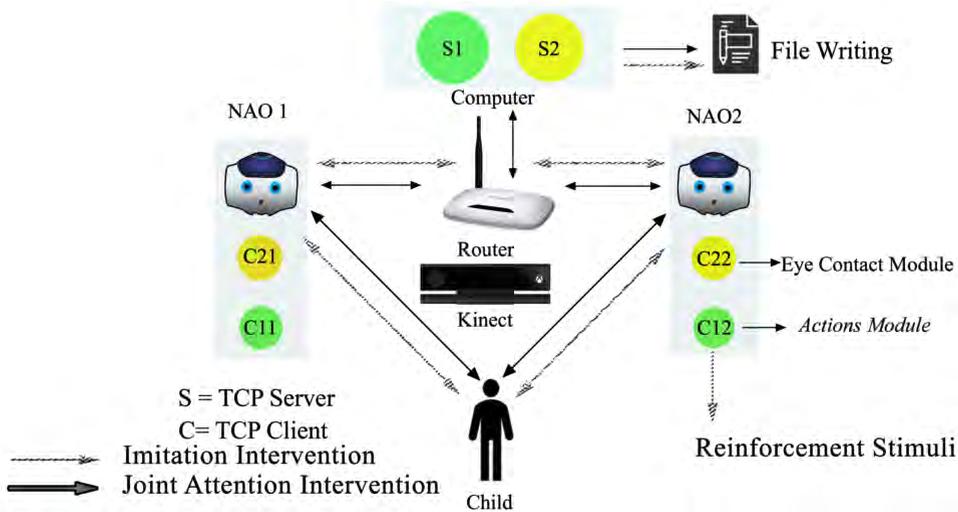


FIGURE 3. Networking protocol of MRIS imitation attention module.

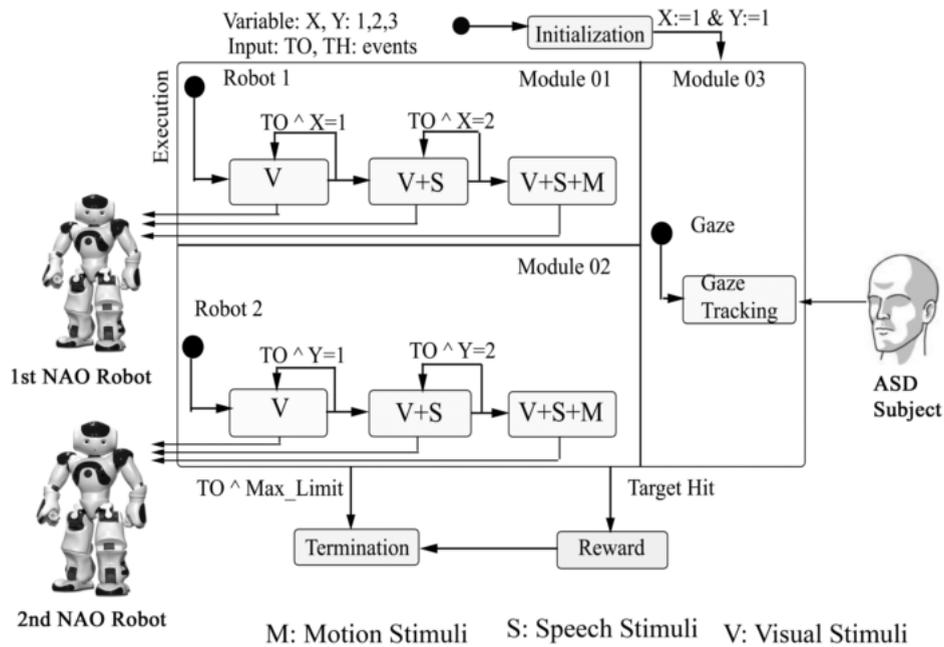


FIGURE 4. Harel statechart model of MRIS joint attention module.

speech and motion stimuli. MRIS-LTM model follows the following steps:

Step 1: Index is passed to Robot\_Action\_List which gives a robot action to be performed and hence Robot\_Behavior defines a behavior of the robot. Then Participant\_Joint\_Attention () function starts which records the joint attention and gives an associated current response. If current response matches with expected response, then reward is given i.e. when  $RESP(n) = ExpResp ()$  and control is transferred to step 3. If current response is not the expected response, then step 2 activates and step 1 is repeated till Max\_Limit is

reached i.e. when  $RESP(n)! = ExpResp ()$ . Data is recorded, and code terminates.

Step 2: Max\_Limit is checked first, if it has been reached the code terminates otherwise current response is given to Next\_Robot\_Action () function and step 1 is repeated till PQ is filled or any condition is met in step 1.

Step 3: In step 3, code terminates after saving the PQ list in sorted manner in an excel file. Max\_Limit represents the number of attempts without setting the expected response. The Harel state chart of joint attention model is presented in Fig. 4.

MRIS-LTM mathematical model for joint attention is shown in (1), (2), (3) and (4).

$$S_1 = XOR \{Initialization, Execution, Termination, Reward\} \quad (1)$$

$$S_2 = AND \{Robot1, Robot2, GazeModule\} \quad (2)$$

$$S_3 = OR\{V, V + S, V + S + M\} \quad (3)$$

Here “ $S_i$ ” denotes the output from different hierarchy levels. “ $i$ ” denotes the hierarchy number. All possible operands combination refers to a state from state machine diagram as shown in Fig. 4. All the states are mentioned in hierarchical level i.e.  $S_1$  is the top level state or parental state.  $S_2$  is intermediate state and  $S_3$  is the leaf node state. Equation (1), (2) and (3) indicate the control operator during the experiment. The depiction of where these control operators are applied is given in Fig. 4 showing which state will be active. For example in case of  $S_1$ , XOR gate represents that only for single high input the output of parental state will be high. Possible combinations for this module are  $\{1,0,0,0\}$ ,  $\{0,1,0,0\}$ ,  $\{0,0,1,0\}$  and  $\{0,0,0,1\}$  as only one input can be high at a time. States  $S_2$  and  $S_3$  are applicable only in case when  $S_1 = \{0,1,0,0\}$  i.e. in execution state only. Similarly for  $S_2$  an AND gate is implemented referring that all the three inputs should be active in order to actuate the intermediate state. Similarly in (3),  $S_3$  is represented by an OR gate such that if any stimulus in input is high the leaf node state  $S_3$  will execute. The stimulus will be executed in LTM order. However in suggested model only one state can work at a time. It can be seen that  $S_1$ ,  $S_2$ ,  $S_3$  are only used to trigger the respective level in the module once conditions are met. However the outputs themselves are analog and no data is discarded. Therefore the results in Table 3 and 4 in the Experimental design section are the results of therapies being performed within the state itself and are therefore not Boolean numbers.

$S_2$  is running in parallel under execution stage. Two signals i.e. timeout (TO) and target hit (TH) along with threshold values determine whether the system needs to shift to the next cue or not. This model not only places a check on time in which gaze of the child should be directed towards the robot but also the time duration for which it should establish the eye contact in order for it to be claimed as a target hit. For activation of joint attention module, the minimum time for eye contact should be at least 5 secs. TO triggers if no action is done by ASD child in that particular time of module activations.

To represent module 1 and 2 in execution stage the first step of MRIS LTM protocol i.e. the least prompt cue level is when each robot starts with visual cues i.e. rasta and blinking to measure child’s joint attention. This is represented by  $\{V\}$ . If the threshold value for joint attention is not achieved by the child, it moves to the next level represented as  $\{V+S\}$ . If the child does not meet the threshold value of this level too, the therapy is moved to third stage i.e. the highest level  $\{S+V+M\}$ . In execution stage, all the modules i.e. robot 1, robot 2 and gaze are working in parallel. Depending on

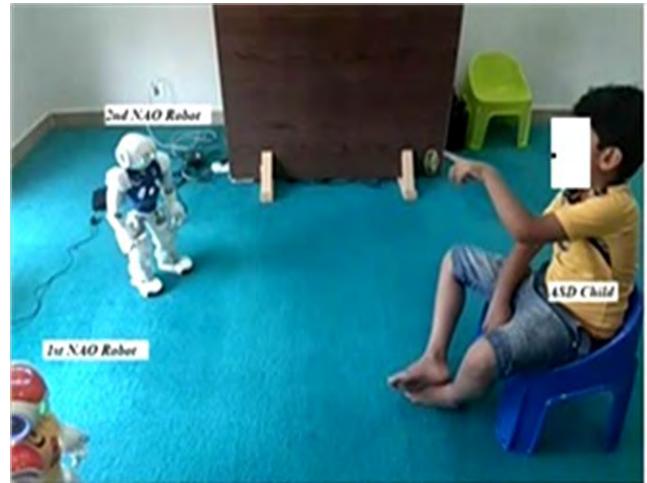


FIGURE 5. ASD child interacting with NAO robots during joint attention module.

the child’s performance at any stage he/she is rewarded at the same stage after the completion of therapy. This model also records the particular stage the child has to start when introduced to the therapy next time. The threshold value is the hyper parameter of this model. For this research the threshold value is 50%. Fig. 5 shows the interaction of child with the robots for the joint attention module.

The joint attention module is further explained using mathematical equations represented in (4) and (5). The joint attention module is linked with the reinforcement stimulus. A reinforcement stimulus is given by the robots to measure the joint attention of each subject. In order to execute this task two modules are running in parallel under  $O_{JA}$ , a module to measure joint attention and stimulus module. These two operands for an “AND” operation represent parallel execution for true state as shown in (4). First operand deals with joint attention recording for both robots while the second operand deals with reinforcement stimulus being given by the robots. This is represented in (4). Mathematical model of (4) is further explained in (5) where “ $i$ ” denotes the robot number, “ $k$ ” denotes the number of eye contacts and “ $j$ ” denotes the type of reinforcement stimulus. “ $n$ ” and “ $m$ ” belong to real numbers. In our case we have presented three stimuli denoted by  $ST_j$  as shown in (6) where  $R_i$  represents the robots presented in (7).

$$O_{JA} = AND \{Jointattention, Reinforcement Stimulus\} \quad (4)$$

$$O_{JA} = AND \left\{ \sum_{k=1}^n \left( \sum_{i=1}^2 \left( R_i, \int_{t=1}^m dt \right) \right), \sum_{j=1}^3 \left( \sum_{i=1}^2 (R_i, ST_j) \right) \right\}; \quad (5)$$

where  $n, m \in \mathbb{R}$

$$ST_j = \begin{cases} V & ; \text{if } j = 1 \\ V + S & ; \text{if } j = 2 \\ V + S + M & ; \text{if } j = 3 \end{cases} \quad (6)$$

$$R_i = \begin{cases} Robot \ one & ; \text{if } i = 1 \\ Robot \ two & ; \text{if } i = 2 \end{cases} \quad (7)$$

Equation (5) can be further explained by illustrating it in iterative manner:

1<sup>st</sup> Iteration:

$$j = 1 \text{ and } i = 1, 2 \left\{ (R_1, x_1) + (R_2, x_1) \right\}$$

$$k = 1 \text{ and } i = 1, 2 \left\{ (R_1, ST_1) \text{ and } (R_2, ST_1) \right\}$$

where  $x_1$  is the duration of eye contact noted by the robot and  $ST_1$  denotes the visual stimuli given by the robot.

2<sup>nd</sup> Iteration:

$$j = 2 \text{ and } i = 1, 2 \left\{ (R_1, x_2) + (R_2, x_2) \right\}$$

$$k = 2 \text{ and } i = 1, 2 \left\{ (R_1, ST_2) \text{ and } (R_2, ST_2) \right\}$$

where  $x_2$  is the duration of eye contact noted by the robot and  $ST_2$  denotes visual + speech stimuli launching on robot one and robot two. Similarly for third iteration  $ST_3$  will denote visual + speech + motion stimuli launching on robot one and robot two. This process continues till completion of the therapy.

### B. IMITATION PROTOCOL

The interaction protocol of MRIS imitation module uses the child’s joint attention to activate the robot. This is done by allotting a certain time limit (5s) for which the child should focus towards the robot in order to activate it. The threshold time not only ensures activation of only one robot at a time but also makes the module adaptive. After eye contact is established with a particular robot, the robot starts its imitation tasks i.e. Move Forward, Move Backward, Raise Hands, Hands Down. These motion gestures are imitated by the child and are measured using Kinect to calculate the success rate.

Introducing the second robot in the experiment helps impart multi-agent communication skills along with improvement in imitation. The functionalities of imitation can also be used with only one robot in the experiment however having an additional robot helps impart communication skills in multi-agent scenario.

#### 1) NETWORKING PROTOCOL FOR IMITATION

The networking protocol of MRIS imitation module is almost the same as joint attention module, shown in Fig. 3. The modules are represented by  $C_{ij}$ , where  $i$  is the server number and  $j$  is the client’s serial number. The action module is dependent on the eye contact module as discussed above and is represented by  $C_{11}$  and  $C_{12}$ .

#### 2) IMITATION-BASED MODEL

During initialization a priority list (PQ) is loaded. PQ list tells us about the last successive action of the robot which had imitated by the ASD child. In step 1 the joint attention of the ASD child is captured with gaze tracking module. If an eye contact is established for 5s, the imitation module starts. The selected time period of 5s is based on observations made during experimentation i.e. the time period should not be so short that it could start overlapping with the second robot’s activation and not so long that the child may lose focus before the response is conveyed.

TABLE 2. Functions in pseudo code.

Function	Description
Load PQ ()	load’s Priority Queue List
Max ()	returns the most effective stimulus existing in PQ
Robot_Action ()	Gives action details to robot to be performed
Robot_Behavior ()	Represents the current behavior of robot
Participant_Joint_Attention ()	Gaze tracking module evaluating joint attention of subject
Insert ()	Inserting current response and index to PQ list.
Filled	check whether PQ has been filled or not?
Next_Robot_Action	Defining the action of next robot to be handled
Sort	Sorting PQ list
Write	Writing sorted PQ list to excel file for future reference.

If the child’s response matches with the robot’s response, the robot gives reward by saying “Good Job” and performs the next action which should be imitated by child as per the protocol of the therapy. In discussion with the therapist, verbal response for encouragement was only added in case of correct imitation. In case of incorrect action performed by the child, discouraging response is not produced by the robot. The description of functions for both algorithms is presented in Table 2. The Harel state chart of imitation model is presented in Fig. 6. Fig. 7 shows the child engaged in imitation module with both robots.

MRIS-LTM mathematical model for imitation based on joint attention is shown in (8), (9), (10) and (11). Here “Si” denotes the output from different hierarchy levels. “i” denotes the hierarchy number. All the states are mentioned in hierarchical level i.e. S1 is the top level state or parental state. S2 is the intermediate state and S3 is the leaf node state as discussed for (1), (2) and (3). The only difference is in the leaf nodes of execution state i.e. S3 which reflects that only one robot will perform either of its two imitation task. This can be seen in Fig. 6 in execution stage. All possible combinations of operands refer to a state from state machine diagram as shown in Fig. 6. The depiction of the control operators are applied is given in Fig. 6 showing which state will be active.

$$S_1 = \text{XOR} \left\{ \begin{array}{l} \text{Initialization, Execution,} \\ \text{Termination, Reward} \end{array} \right\} \quad (8)$$

$$S_2 = \text{AND}\{\text{Robot1, Robot2, Gaze Module}\} \quad (9)$$

$$S_3 = \text{OR}\{\text{OR}\{\text{Forward, Backward}\}, \text{OR}\{\text{Raise hands, Hands down}\}\} \quad (10)$$

Imitation module is further explained using mathematical equations represented in (11) and (12). Equation (11) represents imitation module based on joint attention of the child that triggers this module on establishing eye contact.

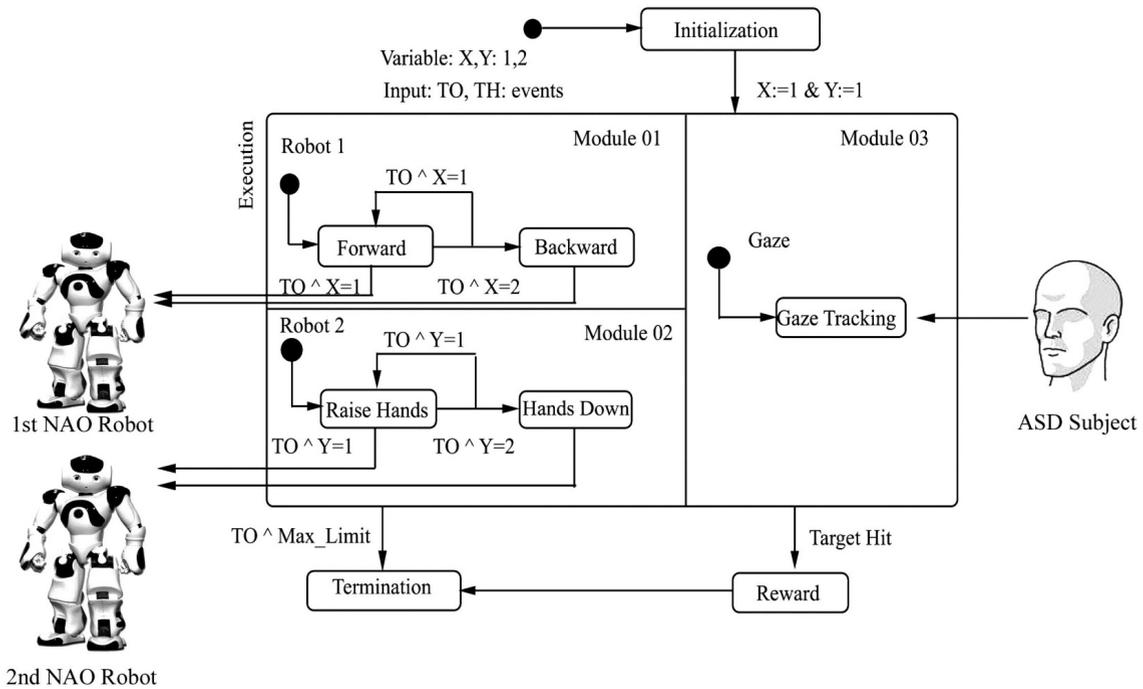


FIGURE 6. Harel statechart model of MRIS imitation module.

Therefore in order to execute this task, joint attention module along with imitation module is running in parallel. For this purpose an “AND” operation is considered best to represent (11). Mathematical model of (11) is further explained in (12).

$$O_{JA \rightarrow IM} = AND \{Joint\ attention, imitation\} \quad (11)$$

$$O_{JA \rightarrow IM} = AND \left\{ \sum_{k=1}^n \left( \sum_{i=1}^2 \left( R_i, \int_{t=1}^m dt \right) \right), \sum_{i=1}^2 \left( \sum_{j=1}^2 (R_i, IM_j) \right) \right\};$$

where  $n, m \in \mathbb{R}$  (12)

Here  $O_{JA \rightarrow IM}$  denotes the output from joint attention integrated with imitation module (state machine). “ $i$ ” denotes robot number, “ $k$ ” denotes the number of eye contacts and “ $j$ ” denotes the type of imitation. “ $n$ ” and “ $m$ ” belongs to real numbers. Where  $IM_j$  is the imitation sequence executed by robots:

$$IM_j = \begin{cases} R_1 \rightarrow \begin{cases} forward & \text{if } j = 1 \\ backward & \end{cases} \\ R_2 \rightarrow \begin{cases} raise\_hands & \text{if } j = 2 \\ hands\_down & \end{cases} \end{cases} \quad (13)$$

For further explanation of the mathematical model, (12) is illustrated below in an iterative manner:

1<sup>st</sup> Iteration:

$$\left. \begin{aligned} i &= 1 \text{ and } j = 1, 2 \left\{ (R_1, x_1) + (R_2, x_1) \right\} \\ k &= 1 \text{ and } i = 1, 2 \left\{ (R_1, IM_1) \text{ and } (R_1, IM_2) \right\} \end{aligned} \right\}$$

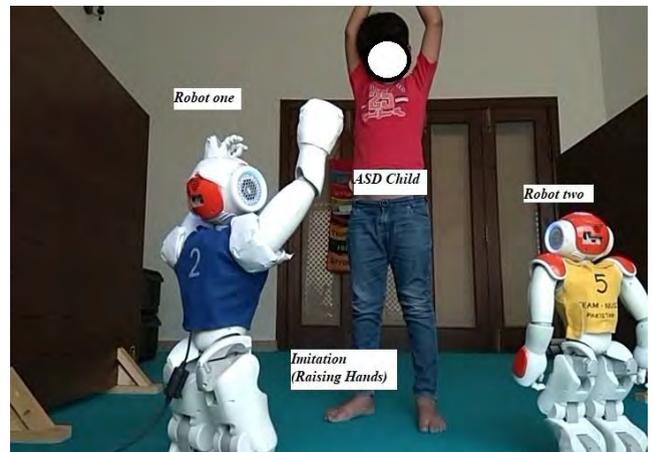


FIGURE 7. Engagement of an ASD child during imitation module with robots.

where  $x_i$  is the duration of eye contact noted by the robot and  $IM_j$  denotes the imitation tasks performed by robot 1.

2<sup>nd</sup> Iteration:

$$\left. \begin{aligned} i &= 2 \text{ and } j = 1, 2 \left\{ (R_1, x_2) + (R_2, x_2) \right\} \\ k &= 2 \text{ and } i = 1, 2 \left\{ (R_2, IM_1) \text{ and } (R_2, IM_2) \right\} \end{aligned} \right\}$$

where  $x_i$  is the duration of eye contact noted by the robot and  $IM_j$  denotes the imitation tasks performed by robot 2. The iterations continue till completion of therapy session.

### C. HARDWARE

MRIS uses two NAO humanoid robots for ASD therapy. Due to its anthropomorphic appearance and high

programmability, it is widely used for the purpose of therapy [40]. The robot controller uses the built in function along with LTM protocol for joint attention. Two kind of interventions are embedded in the controller itself i.e. joint attention intervention module and imitation module. First joint attention intervention is accomplished i.e. an LTM-based protocol therapy followed by imitation module using child's eye contact for activating the robot. Both modules are autonomous i.e. adaptive.

Gaze tracing and posture recognition modules are used for joint attention and imitation interventions respectively. Gaze tracking is done using NAO robots cameras. During the joint attention module two things are noted for gaze attention: 1) Delay in making eye contact with the robot. 2) Time duration for which eye contact is made. Imitation of the child is recorded using Kinect that measures the body posture of the child to match with the robot's posture during imitation therapy.

#### D. ADAPTIVE CLOSED LOOP SUPERVISORY CONTROL

This is the central module of MRIS as it controls the cues based on the child's response and then sends that command to the robot so that it can change the behavior accordingly. In case of adaptive module, the algorithm decides on its own if the child is following the command or not. If the imitation is performed correctly, the algorithm switches to the next level or command according to the protocol of the therapy. In joint attention module, the adaptive closed loop supervisory control gets the feedback from NAO robots camera whereas for imitation module, the correct posture recognition information is recorded by Kinect.

### III. MATERIALS AND METHODS

#### A. SUBJECTS

MRIS system was tested on 12 ASD children including 11 males and 1 female. They were recruited from Autism Resource Center (ARC). The study was approved by the autism specialist and director board of ARC. The recruited participants were already evaluated clinically based on Childhood Autism Rating Scale Schedule (CARS) criteria by the experts. The statistical characteristics of the participants are presented in Table 5 and Table 7. Parents of these children also signed consent form for the discussed therapy.

#### B. ENVIRONMENT SETUP

Fig. 5 and Fig. 7 show the environment setup for therapy. The two robots are placed in front of the child while the child sits on a comfortable plastic chair during the joint attention module so as to attain a height at which he/she can make an eye contact with robot. For imitation module the child stands in front of the robots. The robots are 1 m away from the child and from each other as well. They were placed in an arc like arrangement facing the child.

### C. EXPERIMENT DESIGN

MRIS follows the experiment architecture explained below for 1) human-robot interaction (without any inter robots' interaction) and 2) human-robot interaction along with the inter robot communication. These two strategies are presented as individual experiment architectures and the focus of this paper is not to inter-compare the two of them.

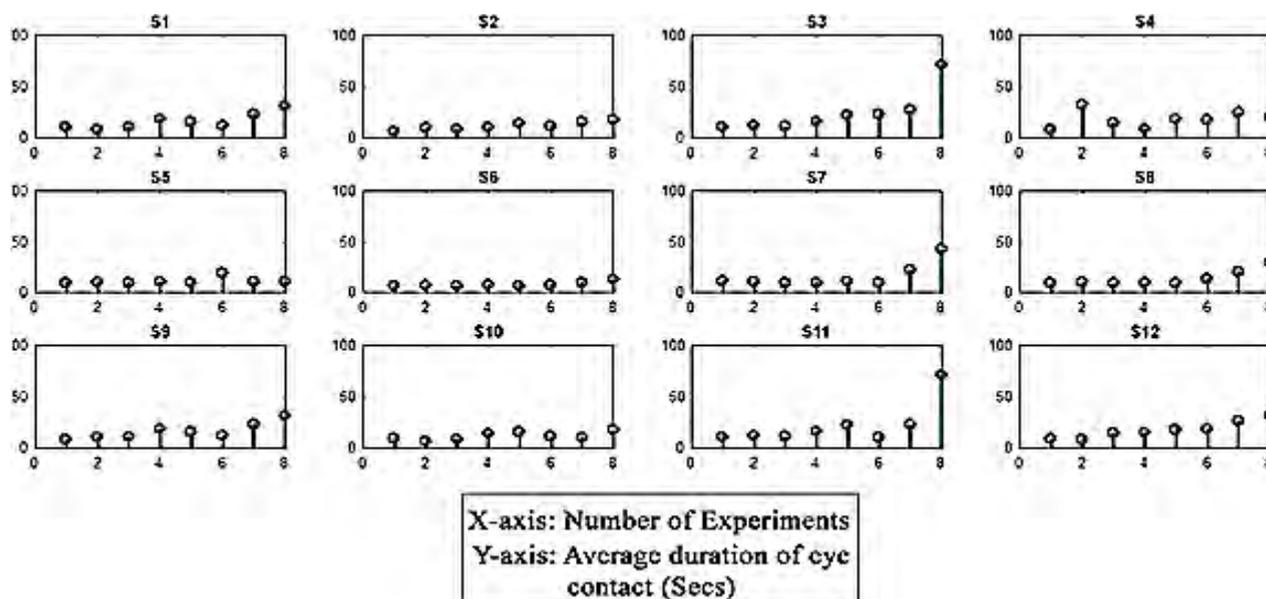
#### 1) HUMAN-ROBOT INTERACTION WITHOUT INTER ROBOT COMMUNICATION

The ASD child is taken to the EEG area before the start of intervention to measure the brain activity and attentiveness. For that he/she sits on a comfortable chair and counts from 1 to 10. After delay of 30s, he/she is asked to read the alphabets. After this reinforcement activity the child's brain activity is measured using EEG neuroheadset. Following that a child is seated on a chair facing both the robots in the intervention area. The robots start with their first intervention therapy i.e. LTM-based adaptive joint attention module. The child's response is noted by NAO robots' cameras. Subsequent to intervention completion, the child is again taken to the EEG room for measurement of brain state after therapy. Moving ahead, the child is then introduced to second intervention i.e. imitation. This intervention follows the same protocol for measuring brain activity before and after the therapy. In this therapy, as the child enters the intervention area, both the robots give stimuli by flashing their eyes (same color was used for both robots) after which the robots wait for the child to maintain eye contact with either robot for at least 5s. Once eye contact was established, the robot is activated for imitation activity. This intervention basically utilizes the joint attention module in a way that imitation of the robot is activated by the eye contact of the ASD child.

#### 2) HUMAN-ROBOT INTERACTION WITH INTER ROBOT COMMUNICATION

In this therapy we have introduced the concept of inter-robot communication along with human-robot interaction of ASD child. The main rationale for introducing inter-robot communication is that in a daily multi-communication setting, one may also need to listen to/watch others' communication. This inter-robot communication is carried out at the start of experiment. Initially both the robots are sitting and facing each other. One of them stands up and says "hello" along with waving action to the partner robot. The partner robot shows the response by standing up, saying "hello/hi" coupled with a waving action.

During this time period, the ASD child's response is recorded as a listening task. Following this, the robot turns towards the child and start communicating in a similar manner. The robot communicating with the child is randomly selected. In return the response of child is noted in terms of joint attention (child's eye contact), imitation (waving of hand) and speech.



**FIGURE 8.** Improvement in eye contact of each subject over the experiments. X-axis represents the number of experiment and Y-axis represents the time for eye contact in seconds for each subject.

It was also observed during experimentation that children were paying attention to the robots and also shifted gaze properly between robots during their inter-communication. The whole arrangement discussed is shown in Fig.1. The dotted line shows inter robot interaction done for these experiments.

Eight sessions for each intervention were conducted over a period of two months for each type of experiment. The entire experimentation was carried out for a period of four months. Weekly progress of the child was recorded for both interventions. Each session involved various trials, of which average was taken to calculate the overall success. The therapy was scheduled in a way that all participants were involved in both interventions.

However, for EEG recording before and after interventions, some participants felt uncomfortable in wearing EEG headset. Their data has not been included in the research. To record the initial engagement of the child with robots as a baseline measure, the participants' attention towards the robot was recorded for several trials until a stable baseline measure was achieved. Therefore, the first experimental value for each intervention was taken as the baseline parameter to measure the success over the intervention. Moreover the results for both types of experimentation i.e. joint attention and imitation module were compared using CARS score before and after the intervention as shown in Table 5. Table 5 explicitly shows the improvement in joint attention and imitation skills of each child before and after the experimentation along with the overall CARS improvement.

### 3) EEG SIGNAL PROCESSING

EEG data was acquired at a sampling rate of 128 Hz. The four EEG bands were measured i.e. alpha, beta and theta by applying the band pass filter on the data. The power

in each band was taken and difference was estimated to measure attentive and non-attentive state of the child. In this way power of each band was used to estimate the cognitive brain state of child [41]. The power of alpha, beta and theta bands state of child [41]. The power of alpha, beta and theta bands was measured during the time interval for which the person interacted with the child. We have used a video based recording and stopwatch to measure the time interval at which the stimuli was given. Average SNR calculated before was 1.06 and after was  $-11.28$ . Second order filter was used for which frequency before filtering  $F1 = 8\text{Hz}$  and after filtering was  $F2 = 12\text{Hz}$ .

### 4) DATA PROCESSING USING NAO

For measuring the eye contact, we used upper camera of the NAO robot. The color space was BGR and frame rate was 15 f/s. NAO's API, "ALGaze Analysis" was used for two events that were associated with closing and opening of eyes.

### 5) DATA PROCESSING USING KINECT

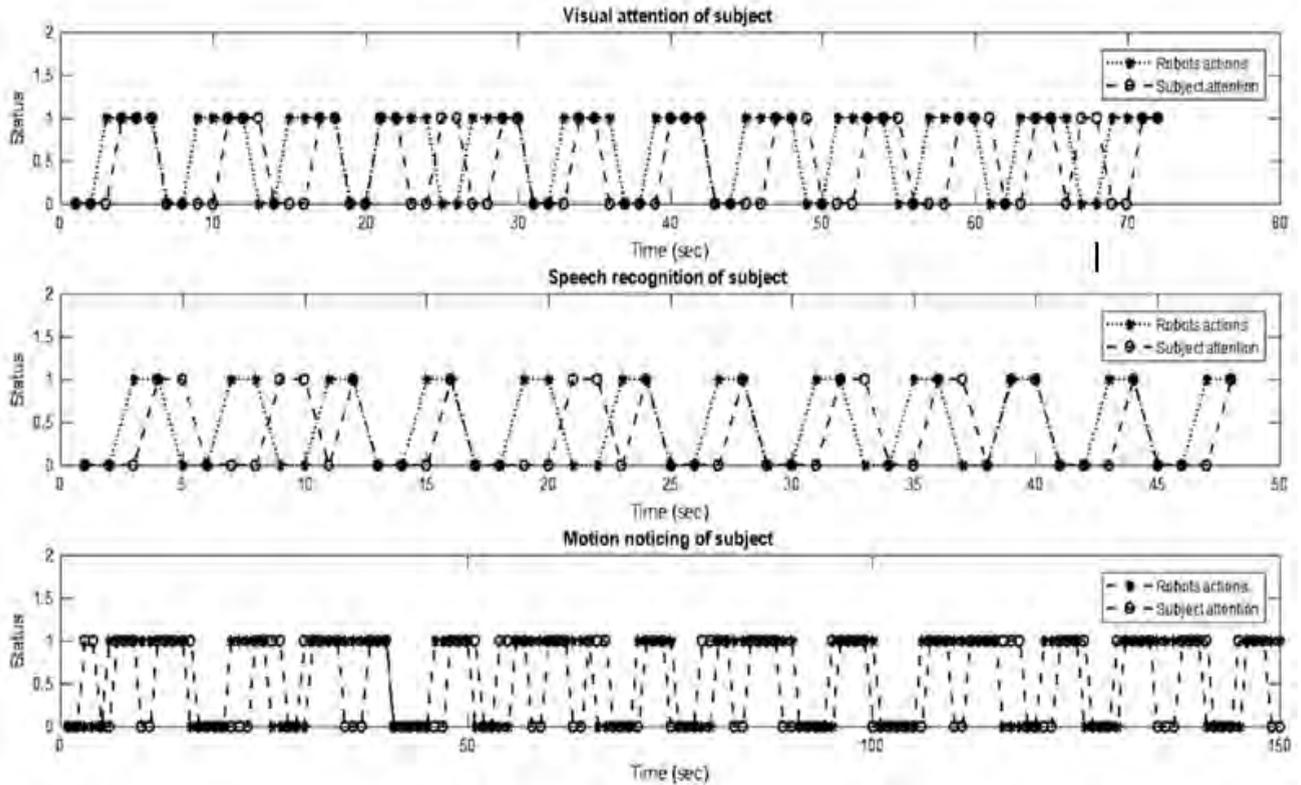
It was used to track the skeleton of the subject. Its frame rate varies depending upon the speed of processing device (laptop). In Kinect it ranges from 15-30 fps.

## IV. RESULT

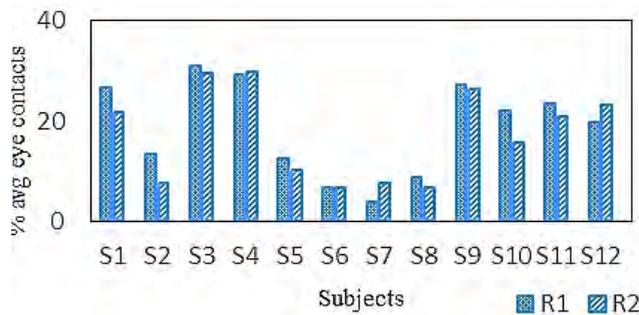
### A. RESULTS OF HUMAN-ROBOT INTERACTION WITHOUT INTER-ROBOT COMMUNICATION

#### 1) RESULTS OF JOINT ATTENTION MODULE

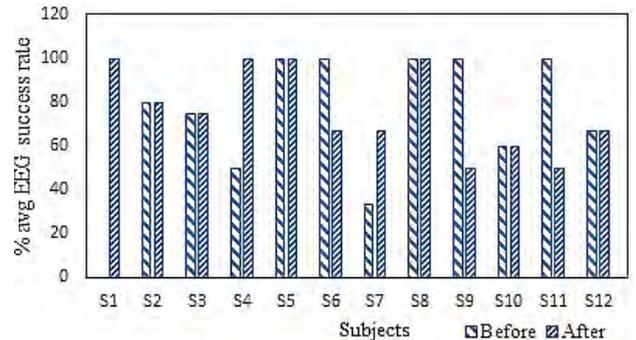
The results of joint attention module were recorded as: 1) eye contact of the ASD child to measure the attentiveness of the child towards a stimulus given by a robot as shown in Fig 8, 2) Delay in shifting the gaze of the child from one robot to



**FIGURE 9.** Pulse plot showing delay in gaze shifting for joint attention module. First plot is dealing with the visual attention of subject (Total number of cues: 12. 6 Blinks and 6 Rasta cues. Duration of each cue: 4 seconds). Second plot is dealing with the speech recognition of subject (Total number of cues: 12. 6 Hello and 6 Hi cues. Duration of each cue: 2 seconds) and third plot deals with the motion-based noting of an ASD subject (Total number of cues: 12. 6 standing & waving and 6 sitting cues. Duration of each cue: standing & waving= 10 seconds; sitting= 5).



**FIGURE 10.** Average number of eye contacts of each subject with robot 1 (R1) and robot 2(R2) over the experiments.



**FIGURE 11.** Average EEG success rate of each individual before and after joint attention module.

the other based on a given stimulus measures improvement in social and cognitive developments. Fig. 9 shows the gaze shifting behavior of S1 for session 2, 3) the biasness of the child towards any of the robots to see the improvement in multi communication as shown in Fig.10 and 4) Interest level of the participants in joint attention module was measured before and after the intervention using EEG as shown in Fig. 11. Based on above parameters the improvement in behavior of the child is shown in Table 3. The table shows the improvement in joint attention of each child compared to the joint attention measured during first week.

## 2) RESULTS OF IMITATION MODULE

The results of this module were assessed after the measurement of interest level of the child using EEG in which a robot was triggered when the child established eye contact: 1) the imitation performed by the child when the robot gives a stimulus to measure the motor skills is shown in Fig. 12, 2) the social interaction based on the stimuli given by the two robots and measuring biasness based on actuation of robots is shown in Fig. 13. 3) The results in Fig. 14 correspond to the joint attention along with the imitation module. The overall

TABLE 3. Joint attention table.

Subjects	Type of eye contact	Joint attention accuracy of week one	EEG accuracy week one	Combined accuracy of week one (EEG + JA)	Joint attention accuracy of all weeks	EEG accuracy of all weeks	Combined accuracy of all weeks (EEG + JA)
S1	Delayed	47.5	0	23.8	63.6	50	56.8
S2	Delayed	53.5	0	26.7	68.8	80	74.4
S3	Delayed	23.1	0	11.6	59.6	75	67.3
S4	Delayed	42.18	0	21.1	52.5	75	63.8
S5	Delayed	73.1	0	36.5	67.9	100	83.9
S6	Delayed	32.6	0	16.3	66.1	83.3	74.7
S7	Immediate	78.8	0	39.4	63.1	50	56.5
S8	Immediate	93.4	100	96.7	85.5	100	92.7
S9	Delayed	28.9	50	39.5	43.4	75	59.2
S10	Delayed	66.7	0	33.3	72.2	60	66.1
S11	Delayed	52.4	0	26.2	76.1	75	75.6
S12	Delayed	51.5	0	25.8	67.6	66.7	67.1

TABLE 4. Imitation table.

Subjects	Type of eye contact	Imitation accuracy of week one	EEG accuracy week one	Combined accuracy of week one (Imitation+ JA)	Imitation accuracy of all weeks	EEG accuracy of all weeks	Combined accuracy of all weeks (Imitation + JA)
S1	Immediate	83.3	0	41.7	74.5	50	62.2
S2	Immediate	75	0	37.5	73.3	75	74.2
S3	Immediate	75	0	37.5	58.9	83.3	71.1
S4	Delayed	70.8	0	35.4	64.2	75	69.6
S5	Immediate	100	0	50	90.6	87.5	89.1
S6	Immediate	87.5	0	43.7	81.3	100	90.6
S7	Immediate	75	0	37.5	67.2	83.3	75.3
S8	Immediate	100	100	100	75	87.5	81.3
S9	Immediate	83.3	0	41.7	91.1	75	83.1
S10	Immediate	100	0	50	82.3	75	78.6
S11	Immediate	100	50	75	92.2	83.3	87.8
S12	Immediate	100	0	50	89.1	83.3	86.2

TABLE 5. Cars table for human-robot interaction without inter robot communication.

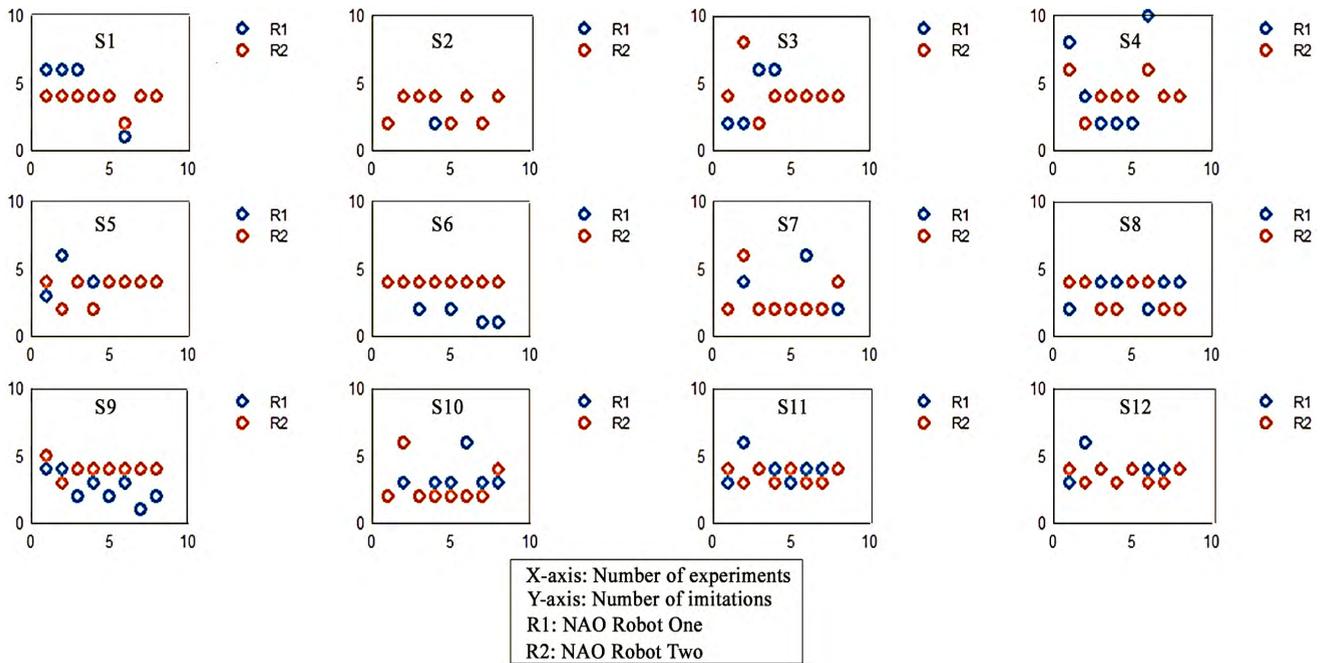
Subjects	Age (YEARS)	avg_imi Before	avg.ja Before	CARS Before	avg_imi After	avg_ja After	CARS After
S1	7.5	2	2.3	30.5	1.3	1.5	28
S2	8	2.3	1.8	30.5	2	1.8	27.5
S3	9	2.3	2.5	35	2.5	2	33
S4	10	2	2.8	40	2	2.8	37
S5	5	2.5	2.5	33.5	1.3	1.8	28
S6	8.5	2.3	2.3	33.5	1	1	33
S7	4.3	1	1.8	22.5	1.3	1	19.5
S8	3.7	1.5	1.3	20.5	1.3	1.3	19
S9	9.9	1.5	1.5	22.5	1.3	1.3	21
S10	9.8	1.3	1.5	27.5	1.5	1.3	26
S11	10.4	2.3	2.3	36	1.3	1.3	32.5
S12	9.4	1.8	2.3	34.5	1.5	2.3	31

improvement in imitative behavior of the child from week one is shown in Table 4.

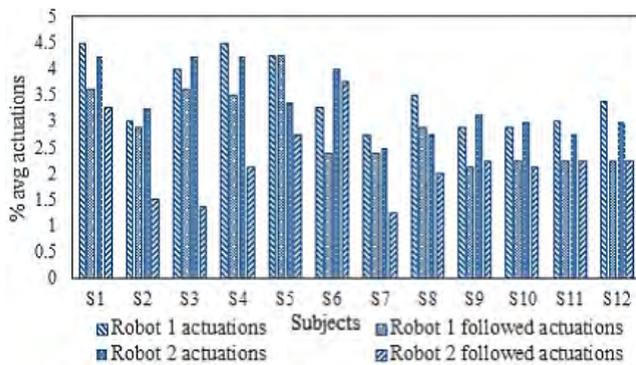
Results for both types of experimentation i.e. joint attention and imitation modules along with joint attention were verified using CARS score before and after the intervention as shown in Table 5 and Table 8.

**B. RESULTS OF HUMAN-ROBOT INTERACTION WITH INTER-ROBOT COMMUNICATION**

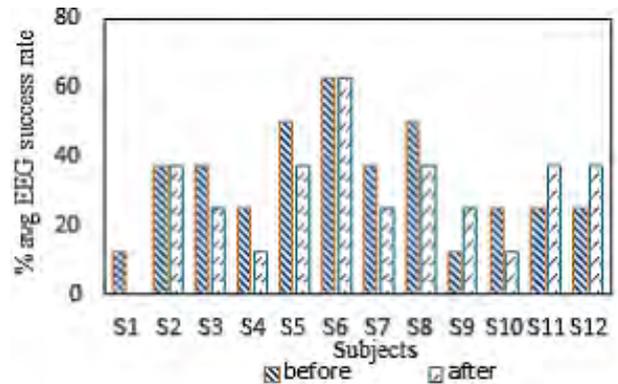
Four different parameters were evaluated for all the sessions conducted over a period of 8 weeks. The average values of parameters were plotted for each subject as shown in Fig. 15. Fig. 16 shows the detailed distribution of eye contact of an



**FIGURE 12.** Improvement in imitation of each subject over the experiments. X-axis represents the number of experiment and Y-axis represents the number of imitation.



**FIGURE 13.** Average number of followed actuations by each ASD child and actuations given by both robots in all experiments.



**FIGURE 14.** Average EEG success rate of each individual before and after joint attention along with the imitation module.

ASD child with both robots along with an average eye contact time and number of eye contacts maintained.

Table 7 shows the results of different parameters which are considered in updated therapy. In this therapy we are considering waving and speech response of an ASD child towards both robots along with attention paid by the ASD child towards intercommunication of robots. Last three columns represent the percentage of success of different subjects. Further definitions of different acronyms used in Table 7 have been given in Table 6.

The impact of this intervention can be seen by pre and post intervention CARS score represented in Table 8.

### V. STATISTICAL ANALYSIS

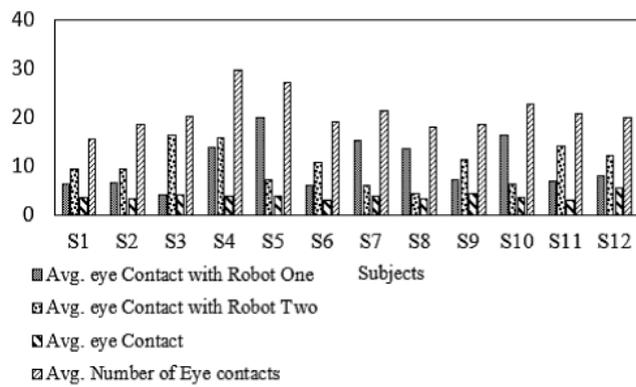
Statistical analysis of this research is done for each intervention module. We have used ANOVA (single factor) for

this purpose. The result for joint attention and EEG module was  $F$  value = 20.36,  $p$ -value = 1.74E-06 and  $F$  critical value = 3.28.

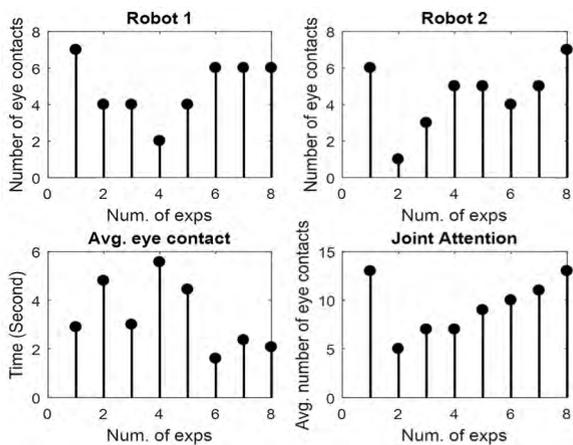
Fig. 17 shows the graph of ANOVA for joint attention and EEG modules without inter-robot communication. Results for joint attention and imitation were  $F$  value = 23.93,  $p$ -value = 3.79E-07 and  $F$  critical value = 3.28. Fig. 18 shows the graph of ANOVA for a particular module without inter-robot communication. The results for last intervention module i.e. measuring the joint attention and imitation skills of an ASD child with inter-robot communication in the intervention was  $F$  value = 4.52,  $p$ -value = 0.0185 and  $F$  critical value = 3.28. Fig. 19 shows the graph of ANOVA for this particular module.

**TABLE 6.** Parameters description evaluated during inter-robot communication.

Parameter	Description
WR	Waving response of ASD child towards any robot
$R_1$ -ASD	Robot one is interacting with ASD child
$R_2$ -ASD	Robot two is interacting with ASD child
SR	Speech Response of ASD child towards a robot
COMM	Communication / interaction between the robots presented to ASD child
$R_1 - R_2$	Robot one is interacting with robot two
$R_2 - R_1$	Robot two is interacting with robot one
WR %	Accumulative waving response %age of ASD child for robot one and two
SR %	Accumulative speech response %age of ASD child for robot one and two
COMM %	Attention paid towards the robot-robot communication

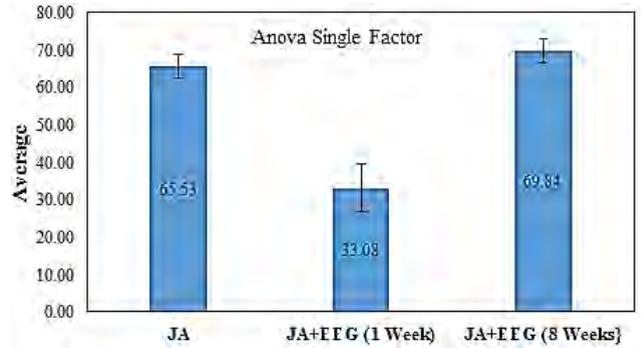


**FIGURE 15.** Overall results for 12 ASD children for joint attention, imitation and response to inter-robot communication.

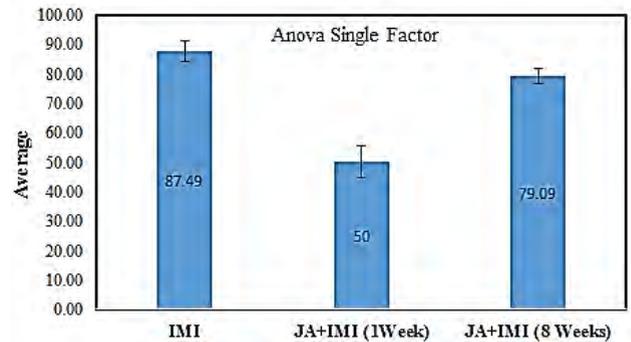


**FIGURE 16.** Details of joint attention of a subject over all the experiments.

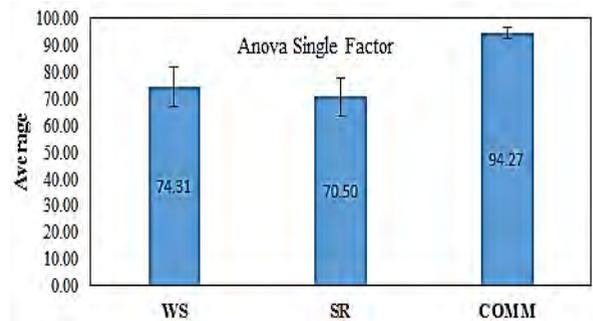
Since we have used single factor ANOVA for statistical analysis therefore we have p-value along with F statistic. In our case, we selected  $\alpha = 0.05$  as a threshold and we got p-value lower than  $\alpha$  i.e.  $p\text{-value} = 1.74E-06$ , showing that our data is reliable. Moreover in both interventions, our calculated F value is greater than critical F value, thus rejecting the null hypothesis.



**FIGURE 17.** Average values calculated using ANOVA single factor for EEG and joint attention of overall sessions.



**FIGURE 18.** Average values calculated using ANOVA single factor for joint attention and imitation of overall sessions.



**FIGURE 19.** Average values calculated using ANOVA single factor for inter-robot communication module of overall sessions.

## VI. DISCUSSION

Unlike previous research, our designed modules for joint attention as well as imitation are adaptive. Various studies have been carried out related to this LTM-based prompt method showing that it is not restricted to only imitation or joint attention but can be used generally for any robot mediated therapy. In a research presented in [42], the child was asked to imitate the robot’s gesture. If the child fails then the robot points in order to improve the gesture.

In another research, the robot therapy is based on asking open question initially, if the child is unable to answer it correctly then the robot adds a hint of correct answer in it [43]. Similarly ARIA system uses the LTM-based protocol

**TABLE 7.** Results of evaluated parameters during inter-robot communication.

Subj.	WR		SR		COMM		WR %	SR %	COMM %
	$R_1$ -ASD	$R_2$ -ASD	$R_1$ -ASD	$R_2$ -ASD	$R_1 - R_2$	$R_2 - R_1$			
S1	7	5	5	6	8	8	75	68.75	100
S2	8	8	8	8	8	8	100	100	100
S3	8	8	8	7	8	8	100	93.75	100
S4	7	8	5	4	8	8	93.75	56.25	100
S5	8	7	7	8	8	8	87.5	93.75	100
S6	6	6	4	3	7	8	80	46.66	93.75
S7	7	3	7	6	8	7	66.66	86.66	93.75
S8	3	0	3	1	8	7	20	26.66	93.75
S9	2	4	2	2	6	6	50	33.33	75
S10	8	8	7	8	8	8	100	93.75	100
S11	4	3	6	4	7	7	50	71.42	87.5
S12	6	5	7	5	6	8	68.75	75	87.5

**TABLE 8.** Cars table for human-robot interaction with inter robot communication.

SUBJECTS	AGE (YEARS)	CARS BEFORE	CARS AFTER
S1	7.5	28	25.5
S2	8	27.5	27
S3	9	33	32.5
S4	10	37	35
S5	5	28	24
S6	8.5	33	19.5
S7	4.3	19.5	17.5
S8	3.7	19	18
S9	9.9	21	19.5
S10	9.8	26	25.5
S11	10.4	32.5	26.5
S12	9.4	31	30

for the model based on joint attention improvement of an ASD child [42]. Only one research shows a single robot based adaptive model for improvement of joint attention only [40].

Moreover the stimuli generated by both sources are the same and hence no biasness is introduced for an ASD child unlike previous studies that have used screen or other sources as stimuli to measure joint attention of a child. Also there are no external environmental factors involved in our prompts as included in NORRIS [35].

The advantage of this model is that it does not require continuous engagement of the human therapist. It is difficult for any person to work for extended continuous hours as a therapist unlike robots. Moreover these robot based therapies can be conducted at home. Keeping in view the non-human involvement it has certain disadvantages particularly if the child gets frustrated, how to manage the situation.

In addition to that another significant factor is willingness of the child for EEG recordings. Therefore it would be better to use some other device instead of EEG as children are sometimes reluctant in wearing the device. Moreover the proposed research does not compare the two models i.e. human-robot

interaction without inter-robot communication and human-robot interaction with inter-robot communication. Therefore the focus of this study is not to show which therapy is better than another as the protocol of both the therapies are different and depends on the intervention to be conducted.

The proposed future work for this research is implementing the proposed model on a larger set of ASD children. Secondly this model can be extended for more than two robots. Moreover the effectiveness of therapy can be evaluated for human-human interaction as well.

**VII. CONCLUSION**

Based on the results, this research has three main contributions. 1) Design and development of a single mathematical model for adaptive multi-robot based therapy of ASD children for both LTM-based joint attention as well as imitation. 2) Validation and effectiveness of MRIS system based on user study using CARS scale as shown in Table 5 and 8. This gives an insight into how effective the designed therapy is. 3) Notable improvement in multi interaction of an ASD child.

In this article, we have proposed the first autonomous multi-robot based mediated therapy for joint attention and imitation called MRIS. Two humanoid robots (NAO) were used as interaction partners of an ASD child. For the first intervention, interaction of a child was recorded in two different modules i.e. joint attention and imitation module. In joint attention module the child’s gaze tracking was acquired using NAO camera to observe eye contact and delay in making contact after the stimulus is given. The implemented prompts of this module are based on LTM-RI hierarchy. In imitation module, the activation of module was dependent on eye contact of an ASD child, hence making the module itself adaptive. The child’s imitation was measured over a period of experiments to observe any improvement in the child’s behavior.

The second intervention involved inter-robot communication during which the child’s behavior was recorded when the robots were communicating with each other. This is a normal protocol in daily life communication when one may also need to watch or listen to others’ communication.

The improvement in multi-communication skills of the child with robots was recorded during intervention.

The child was introduced to eight sessions of each intervention. Each intervention was carried out for two months. The therapy was spread over a period of 6 months. All 12 subjects participated in each intervention. The participation of each individual was made sure by scheduling in such a way that each session for both interventions was carried out over a whole week. We had instances when sessions could not be conducted because of an unexpected reason or child's absence. Therefore the experiment was conducted on any other feasible day of the same week as per schedule. This is how all 12 subjects participated in all the sessions. Full participation was also ensured through meetings with the parents and therapist.

Results show that eye contact duration of each participant has improved over the experiments. Some degree of improvement was shown by every participant. Moreover the delay in making eye contact with the robot after the stimulus is given has been reduced. i.e., the subjects became more responsive to the stimuli. For imitation module it was observed that the participant actuated both robots almost equally in recurring experiments. Therefore the therapy proves to be successful for multi-interaction as shown in Fig. 15. However while testing the system and gathering data, it was noticed that the percentage of success varied from child to child as each individual was responsive towards different type of stimuli based on the level of autism they fall in.

The mathematical model for MRIS was validated by the cognitive brain state measured before and after the experiments using EEG headset (Fig 12 and Fig 14). Moreover the CARS score before and after the therapy shows a significant improvement in communication skills of an ASD child. The statistical analysis performed on the results also supports the conclusion firmly.

The advantage of this model is that it does not require any body worn sensors during intervention that can make the child uncomfortable. Additionally the improvement in child's behavior is recorded using sensor integration therefore reducing the chance of error and ensuring correctness of results.

## REFERENCES

- [1] D. L. Christensen, D. A. Bilder, W. Zahorodny, S. Pettygrove, M. S. Durkin, R. T. Fitzgerald, C. Rice, M. Kurzius-Spencer, J. Baio, and M. Yeargin-Allsopp, "Prevalence and characteristics of autism spectrum disorder among 4-year-old children in the autism and developmental disabilities monitoring network," *J. Develop. Behav. Pediatrics*, vol. 37, no. 1, pp. 1–8, 2016.
- [2] M. Stewart, D. Forbes, D. W. Austin, and J. A. McGillivray, "Parenting a child with an autism spectrum disorder: A review of parent mental health and its relationship to a trauma-based conceptualisation," *Adv. Mental Health*, vol. 15, no. 1, pp. 4–14, 2017.
- [3] R. L. Cachia, D. W. Moore, and A. Anderson, "Mindfulness in individuals with autism spectrum disorder: A systematic review and narrative analysis," *Rev. J. Autism Develop. Disorders*, vol. 3, no. 2, pp. 165–178, 2016.
- [4] P. Mundy, "Joint attention and social-emotional approach behavior in children with autism," *Develop. Psychopathol.*, vol. 7, no. 1, pp. 63–82, 1995.
- [5] P. Mundy and A. Gomes, "Individual differences in joint attention skill development in the second year," *Infant Behav. Develop.*, vol. 21, no. 3, pp. 469–482, 1998.
- [6] L. Billeci, A. Narzisi, A. Tonacci, B. Sbriscia-Fioretti, L. Serasini, F. Fulceri, F. Apicella, F. Sicca, S. Calderoni, and F. Muratori, "An integrated EEG and eye-tracking approach for the study of responding and initiating joint attention in autism spectrum disorders," *Sci. Rep.*, vol. 7, no. 1, 2007, Art. no. 13560.
- [7] S. M. Eack, D. P. Greenwald, S. S. Hogarty, A. L. Bahorik, M. Y. Litschge, C. A. Mazefsky, and N. J. Minshew, "Cognitive enhancement therapy for adults with autism spectrum disorder: Results of an 18-month feasibility study," *J. Autism Develop. Disorders*, vol. 43, no. 12, pp. 2866–2877, 2013.
- [8] M. Begum, R. W. Serna, and H. A. Yanco, "Are robots ready to deliver autism interventions? A comprehensive review," *Int. J. Social Robot.*, vol. 8, no. 2, pp. 157–181, 2016.
- [9] E. S. Kim, R. Paul, F. Shic, B. Scassellati, "Bridging the research gap: Making HRI useful to individuals with autism," *J. Hum.-Robot Interact.*, vol. 1, no. 1, pp. 26–54, 2012.
- [10] I. Fujimoto, T. Matsumoto, P. R. S. De Silva, M. Kobayashi, and M. Higashi, "Mimicking and evaluating human motion to improve the imitation skill of children with autism through a robot," *Int. J. Social Robot.*, vol. 3, no. 4, pp. 349–357, Nov. 2011.
- [11] E. S. Kim, L. D. Berkovits, E. P. Bernier, D. Leyzberg, F. Shic, R. Paul, and B. Scassellati, "Social robots as embedded reinforcers of social behavior in children with autism," *J. Autism Develop. Disorders*, vol. 43, no. 5, pp. 1038–1049, 2013.
- [12] J. Wainer, F. Amirabdollahian, K. Dautenhahn, and B. Robins, "Using the humanoid robot KASPAR to autonomously play triadic games and facilitate collaborative play among children with autism," *IEEE Trans. Auton. Mental Develop.*, vol. 6, no. 3, pp. 183–199, Sep. 2014.
- [13] Y. Feng, M. Chu, W. Wei, and Q. Jia, "Engagement evaluation for autism intervention by robots based on dynamic Bayesian network and expert elicitation," *IEEE Access*, vol. 5, pp. 19494–19504, 2017.
- [14] P. Pennisi, A. Tonacci, G. Tartarico, L. Billeci, L. Ruta, S. Gangemi, and G. Pioggia, "Autism and social robotics: A systematic review," *Autism Res.*, vol. 9, no. 2, pp. 165–183, 2016.
- [15] B. Robins, K. Dautenhahn, and J. Dubowski, "Does appearance matter in the interaction of children with autism with a humanoid robot?" *Interact. Stud.*, vol. 7, no. 3, pp. 479–512, 2006.
- [16] B. Scassellati, "Quantitative metrics of social response for autism diagnosis," in *Proc. IEEE Int. Workshop Robot Hum. Interact. Commun.*, Roman, Italy, Aug. 2005, pp. 585–590.
- [17] J. Feil-Seifer and M. Mataric, "Using proxemics to evaluate human-robot interaction," in *Proc. 5th ACM/IEEE Int. Conf. Hum.-Robot Interact.*, Mar. 2010, pp. 143–144.
- [18] J. J. Diehl, M. Villano, C. R. Crowell, and L. M. Schmitt, "The clinical use of robots for individuals with Autism Spectrum Disorders: A critical review," *Res. Autism Spectr. Disorders*, vol. 6, no. 1, pp. 249–262, 2012.
- [19] Y. Feng, Q. Jia, G. Chen, and C. Li, "Control architecture design for intelligent robot assisted intervention for children with autism," in *Proc. IEEE 11th Conf. Ind. Electron. Appl. (ICIEA)*, Jun. 2016, pp. 1345–1349.
- [20] H. Kozima, C. Nakagawa, and M. P. Michalowski, "Keepon," *Int. J. Social Robot.*, vol. 1, no. 1, pp. 3–18, 2009.
- [21] J. Greczek, E. Kaszubski, A. Atrash, and M. Mataric, "Graded cueing feedback in robot-mediated imitation practice for children with autism spectrum disorders," in *Proc. 23rd IEEE Int. Symp. Robot Hum. Interact. Commun.*, Roman, Italy, Aug. 2014, pp. 561–566.
- [22] K. Dautenhahn, C. L. Nehaniv, M. L. Walters, B. Robins, H. Kose-Bagci, N. A. Mirza, and M. Blow, "KASPAR—A minimally expressive humanoid robot for human-robot interaction research," *Appl. Bionics Biomech.*, vol. 6, nos. 3–4, pp. 369–379, 2009.
- [23] H. Kozima, Y. Yasuda, and C. Nakagawa, "Children-robot interaction: A pilot study in autism therapy," *Prog. Brain Res.*, vol. 164, pp. 385–400, 2007.
- [24] S. Tariq, S. Baber, A. Ashfaq, Y. Ayaz, M. Naveed, and S. Mohsin, "Interactive therapy approach through collaborative physical play between a socially assistive humanoid robot and children with autism spectrum disorder," in *Proc. Int. Conf. Social Robot.*, 2016, pp. 561–570.
- [25] L. Damiano and P. Dumouchel, "Anthropomorphism in human-robot co-evolution," *Frontiers Psychol.*, vol. 9, p. 468, Mar. 2018.
- [26] B. Scassellati, M. Mataric, and H. Admoni, "Robots for use in autism research," *Annu. Rev. Biomed. Eng.*, vol. 14, nos. 3–4, pp. 275–294, 2012.

- [27] A. R. Taheri, A. Meghdari, H. R. PourEtemad, S. L. Holderread, and M. Alemi, "Clinical application of humanoid robots in playing imitation games for autistic children in Iran," *Procedia-Social Behav. Sci.*, vol. 176, pp. 898–906, Feb. 2015.
- [28] H. Akechi, Y. Kikuchi, Y. Tojo, K. Hakarino, and T. Hasegawa, "Mind perception and moral judgment in autism," *Autism Res.*, vol. 11, no. 9, pp. 1239–1244, 2018.
- [29] Y. Funabiki and T. Shiwa, "Weakness of visual working memory in autism," *Autism Res.*, vol. 11, no. 9, pp. 1245–1252, 2018.
- [30] H. Kumazaki, Z. Warren, A. Swanson, Y. Yoshikawa, Y. Matsumoto, H. Ishiguro, N. Sarkar, Y. Minabe, and M. Kikuchi, "Impressions of humanness for android robot may represent an endophenotype for autism spectrum disorders," *J. Autism Dev. Disorders*, vol. 48, no. 2, pp. 632–634, 2018.
- [31] J. Lee, H. Takehashi, C. Nagai, G. Obinata, and D. Stefanov, "Which robot features can stimulate better responses from children with autism in robot-assisted therapy?" *Int. J. Adv. Robot. Syst.*, vol. 9, no. 3, p. 72, 2012.
- [32] O. Rudovic, J. Lee, L. Mascarell-Maricic, B. W. Schuller, and R. W. Picard, "Measuring engagement in robot-assisted autism therapy: A cross-cultural study," *Frontiers Robot. AI*, vol. 5, p. 36, Jul. 2017.
- [33] F. Petric, K. Hrvatinic, A. Babic, L. Malovan, D. Miklic, Z. Kovačić, M. Cepanec, J. Stošić, and S. Šimleša, "Four tasks of a robot-assisted autism spectrum disorder diagnostic protocol: First clinical tests," in *Proc. IEEE Global Hum. Technol. Conf. (GHTC)*, Oct. 2014, pp. 510–517.
- [34] F. Petric, D. Miklic, and Z. Kovačić, "POMDP-based coding of child-robot interaction within a robot-assisted ASD diagnostic protocol," *Int. J. Hum. Robot.*, vol. 15, no. 2, 2018, Art. no. 1850011.
- [35] Z. Zheng, H. Zhao, A. R. Swanson, A. S. Weitlauf, Z. E. Warren, and N. Sarkar, "Design, development, and evaluation of a noninvasive autonomous robot-mediated joint attention intervention system for young children with ASD," *IEEE Trans. Human-Mach. Syst.*, vol. 48, no. 2, pp. 125–135, Apr. 2018.
- [36] P. Mundy, P. Doehring, and A. Hogan, "A preliminary manual for the abridged early social communication scales (ESCS)," Univ. Miami, Coral Gables, FL, USA, Tech. Rep., 2003.
- [37] C. Lord and M. Rutter, "Autism diagnostic observation schedule-ADOS-2," Autism Diagnostic Observ. Schedule Manual, WPS, Los Angeles, CA, USA, Tech. Rep., 1999.
- [38] M. E. Libby, S. Bancroft, W. H. Ahearn, and J. S. Weiss, "A comparison of most-to-least and least-to-most prompting on the acquisition of solitary play skills," *Behav. Anal. Pract.*, vol. 1, no. 1, pp. 37–43, 2008.
- [39] D. J. Feil-Seifer and M. J. Mataric, "A simon-says robot providing autonomous imitation feedback using graded cueing," in *Proc. Int. Meeting Autism Res. (IMFAR)*, 2012.
- [40] E. T. Bekele, U. Lahiri, A. R. Swanson, J. A. Crittendon, Z. E. Warren, and N. Sarkar, "A step towards developing adaptive robot-mediated intervention architecture (ARIA) for children with autism," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 21, no. 2, pp. 289–299, Mar. 2013.
- [41] N. Naseer, K.-S. Hong, M. J. Khan, and M. R. Bhutta, "Classification of prefrontal and motor cortex activities for development of three-class fNIRS-BCI," in *Proc. 20th Annu. Meeting Org. Hum. Brain Mapping (OHMB)*, Hamburg, Germany, 2014.
- [42] Z. Zheng, A. R. Swanson, A. S. Weitlauf, Z. E. Warren, N. Sarkar, and E. M. Young, "Robot-mediated imitation skill training for children with autism," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 6, pp. 682–691, Jun. 2016.
- [43] B. Huskens, J. Gillesen, R. Didden, E. Barakova, and R. Verschuur, "Promoting question-asking in school-aged children with autism spectrum disorders: Effectiveness of a robot intervention compared to a human-trainer intervention," *Develop. Neurorehabil.*, vol. 16, no. 5, pp. 345–356, 2013.



**SARA ALI** received the M.Sc. degree in robotics from Middlesex University London. She is currently pursuing the Ph.D. degree in robotics and intelligent machine engineering with the National University of Sciences and Technology (NUST), Pakistan, where she is also an Assistant Professor with the School of Mechanical and Manufacturing Engineering (SMME). Her research interests include human-robot interaction and brain-computer interface (BCI).



(HRI), therapeutic robotics (TRs), and brain-computer interface (BCI).

**FAISAL MEHMOOD** received the B.Eng. degree in electrical engineering from the COMSATS Institute of Information Technology, in 2016 and the M.Eng. degree in robotics and intelligent machine engineering from the National University of Sciences and Technology (NUST), Islamabad, Pakistan, in 2018, where he is currently a Lab and Research Engineer with the School of Mechanical and Manufacturing Engineering (SMME). His research interests include human-robot interaction



a focus on the SME sector. He has led several large projects funded by Innovate U.K., the Digital Research and Development Fund for the Arts, and the European Research Council.

**DARREN DANCEY** received the Ph.D. degree in artificial neural networks. He is the interim Head of the School for Computing, Mathematics and Digital Technology, Manchester Metropolitan University. His substantive post is the Head of Division: Computer Science and Information Systems. He has a research background in artificial intelligence. In recent years, he has concentrated on creating collaborations and knowledge exchange between universities and industry with



human-machine interaction, humanoids, and motion planning.

**YASAR AYAZ** received the Ph.D. degree in robotics and machine intelligence from Tohoku University, Sendai, Japan, in 2009, where he is a specially appointed Associate Professor with the Graduate School of Engineering. He is the Chairman and the Central Project Director of the Pakistan's National Center of Artificial Intelligence (NCAI) and also the Head of the Department of Robotics and Intelligent Machine Engineering, NUST, Pakistan. His research interests include



hybrid brain-computer interfacing, rehabilitation, and machine learning.

**MUHAMMAD JAWAD KHAN** received the B.E. and M.S. degrees in mechatronics engineering from Air University, Pakistan, in 2007 and 2010, respectively, and the Ph.D. degree in mechanical engineering from Pusan National University, South Korea, in 2018. He was involved in hybrid EEG-fNIRS-based BCI for rehabilitation during his Ph.D. He is currently an Assistant Professor with the School of Mechanical and Manufacturing Engineering, NUST, Pakistan. His research interests include hybrid brain-computer interfacing, rehabilitation, and machine learning.



Editorial Board Member of three SCI(E) indexed journals. His research interests include robotic rehabilitation, biorobotics, neurorobotics, and machine learning.

**NOMAN NASEER** received the bachelor's, master's, and Ph.D. degrees in mechatronics engineering. He is currently the Head of the Neurorobotics Research Group, Air University, Islamabad, and the CEO of Alpha Automation (PVT Ltd.). He is also an Adjunct Professor with the Ho Chi Minh City University of Technology, Vietnam. He has published more than 70 peer-reviewed papers. He has served as a Reviewer of above 60 SCI(E) indexed journals and is serving as an



**RITA DE CASSIA AMADEU** graduated in psychology from Universidade Paulista, Brazil. She is a Teacher with a demonstrated experience in education and health care industry and an Expert in analytical skills, youth development, mental health, neurotechnology, and research. She is a Neuropsychology Researcher with Seoul National University, South Korea, and an Artificial Intelligence Researcher with NUST, Pakistan. Her current research interests include neurodegenerative

diseases, artificial intelligence, forensic science, and criminology. Diverse topics in psychology, brain and behavioral sciences, quantum physics, neurology, human rights, social work, and anthropology are subjects of her true passion.



**HALEEMA SADIA** received the M.A. degree in special education, mental retardation from the Department of Special Education, in 2008, and the M.Ed. degree in autism (children) from the University of Birmingham, U.K., in 2012. She is currently the Director of the Autism Resource Centre, Islamabad, Pakistan. She was with the Human Development Research Foundation (HDRF), London, U.K. and the Oasis Trust, Pakistan.



**RAHEEL NAWAZ** is the Director of Digital Technology Solutions and a Reader in Analytics and Digital Education, Manchester Metropolitan University (MMU). He has founded and/or headed several research units specializing in artificial intelligence, digital transformations, data science, digital education, and apprenticeships in higher education. He has led numerous funded research projects in U.K., EU, South Asia, and Middle East. He holds adjunct or honorary positions with several

research, higher education, and policy organizations, both in U.K. and overseas. He regularly makes media appearances and speaks on a range of topics, including digital technologies, artificial intelligence, digital literacy, and higher education. Before becoming a full-time academic, he served in various senior leadership positions in the private higher and further education sector and was an Army Officer before that.

...