

8-1-2014

Studying Eye Gaze of Children with Autism Spectrum Disorders in Interaction with a Social Robot

Huanghao Feng
University of Denver

Follow this and additional works at: <https://digitalcommons.du.edu/etd>



Part of the [Engineering Commons](#), and the [Psychology Commons](#)

Recommended Citation

Feng, Huanghao, "Studying Eye Gaze of Children with Autism Spectrum Disorders in Interaction with a Social Robot" (2014).
Electronic Theses and Dissertations. 193.
<https://digitalcommons.du.edu/etd/193>

This Thesis is brought to you for free and open access by the Graduate Studies at Digital Commons @ DU. It has been accepted for inclusion in Electronic Theses and Dissertations by an authorized administrator of Digital Commons @ DU. For more information, please contact jennifer.cox@du.edu, dig-commons@du.edu.

Studying Eye Gaze of Children with Autism Spectrum Disorders in Interaction with a Social Robot

A Thesis

Presented to

The Faculty of the Daniel Felix Ritchie School of Engineering and Computer Science

University of Denver

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

by

Huanghao Feng

August 2014

Advisor: Dr. Mohammad H. Mahoor

©Copyright by Huanghao Feng 2014

All Rights Reserved

Author: Huanghao Feng

Title: Studying Eye-Gaze Attention of Children with Autism Spectrum Disorders in Interaction with a Social Robot

Advisor: Dr. Mohammad H. Mahoor

Degree Date: August 2014

Abstract

Children with Autism Spectrum Disorders (ASDs) experience deficits in verbal and nonverbal communication skills including motor control, emotional facial expressions, and eye gaze attention. In this thesis, we focus on studying the feasibility and effectiveness of using a social robot, called NAO, at modeling and improving the social responses and behaviors of children with autism. In our investigation, we designed and developed two protocols to fulfill this mission. Since eye contact and gaze responses are important non-verbal cues in human's social communication and as the majority of individuals with ASD have difficulties regulating their gaze responses, in this thesis we have mostly focused on this area.

In Protocol 1 (eye gaze duration and shifting frequency are analyzed in this protocol), we designed two social games (i.e. NAO Spy and Find the Suspect) and recruited 21 subjects (i.e. 14 ASD and seven Typically Developing (TD) children) ages between 7-17 years old to interact with NAO. All sessions were recorded using cameras and the videos were used for analysis. In particular, we manually annotated the eye gaze direction of children (i.e. gaze averted '0' or gaze at robot '1') in every frame of the videos within two

social contexts (i.e. child speaking and child listening). Gaze fixation and gaze shifting frequency are analyzed, where both patterns are significantly improved or changed (more than half of the participants increased the eye contact duration time and decrease the eye gaze shifting during both games). The results confirms that the TD group has more gaze fixation as they are listening (71%) than while they are speaking (37%). However there is no significant difference between the average gaze fixations of ASD group.

Besides using the statistical measures (i.e. gaze fixation and shifting), we statistically modeled the gaze responses of both groups (TD and ASD) using Markov models (e.g. Hidden Markov Model (HMM) and Variable-order Markov Model (VMM)). Using Markov based modeling allows us to analyze the sequence of gaze direction of ASD and TD groups for two social conversational sessions (Child Speaking and Listening). The results of our experiments show that for the ‘Child Speaking’ segments, HMM can distinguish and recognize the differences of gaze patterns of TD and ASD groups accurately (79%). In addition, to evaluate the effect of history of eye gaze in the gaze responses, the VMM technique was employed to model the effects of different length of sequential data. The results of VMM demonstrate that, in general, the first order system (VMM with order $D=1$) can reliably represent the differences between the gaze patterns of TD and ASD group. Besides that, the experimental results confirm that VMM is more reliable and accurate for modeling the gaze responses of “Child Listening” sessions than the “Child Speaking” one.

Protocol 2 contains five sub-sessions targeted intervention of different social skills: verbal communication, joint attention, eye gaze attention, facial expressions recognition/imitation. The objective of this protocol is to provide intervention sessions based on the needs of children diagnosed with ASD. Therefore each participant attended in three times of baseline sessions for evaluate his/her existing social skill and behavioral response, when the study began. In this protocol the behavioral responses of every child is recorded in each intervention session where feedbacks are focused on improving their social skills if they lack one. For example if they are not good at recognizing facial expression, we give them feedback on how every facial expression looks like and ask them to recognize them correctly while we do not feedback on other social skills. Our experimental results show that customizing the human-robot interaction would improve the social skills of children with ASD.

Acknowledgments

I still remember the exciting moment that I saw NAO for the first time. I still remember how I got excited when Dr. Mahoor assigned me this project. In the past two years I have been working with these special children and I cannot stop thinking of their future life after ten years or even more. Without my late friend, Mr. Masoud Bahramisharif, and his help, I was not able to start this project. I could not accomplish the goals of this project without my dear mother's support. She means everything to me. Without Mr. S. Mohammad Mavadati's help I might not be able to finish this amazing work. I also want to thank my thesis committee members, Dr. Tim Sweeney, Dr. Matthew Rutherford, and Dr. Jun Zhang. I would thank them all in this acknowledgement.

I again would like to thank my advisor Dr. Mahoor for giving me this unique opportunity to work with NAO and supporting me through this research. Meanwhile, I want to thank the supports and help I received from Dr. Anibal Gutierrez from Florida International University and Ms. Mary Kastner and Ms. Sophie Silver from DU psychology department. I also want to appreciate all my lab mates who gave me lots of courage and motivation and to complete this research. Also many thanks to the University of Denver administrators and professors and staffs to have me here and let me study at DU. I learned a lot here. I feel lucky to be a student at DU.

Table of Contents

ABSTRACT.....	ii
ACKNOWLEDGEMENT	v
LIST OF TABLES	viii
LIST OF FIGURES	ix
<u>CHAPTER 1. INTRODUCTION</u>	1
1.1 AUTISM SPECTRUM DISORDERS (ASD)	1
1.2 SOCIALLY ASSISTIVE ROBOTICS	2
1.3 USING SOCIALLY ASSISTIVE ROBOTS FOR AUTISM THERAPY	5
1.4 THESIS CONTRIBUTIONS	8
1.5 ORGNIZATION	10
<u>CHAPTER 2. CONGNITIVE AND PHYSIOLOGICAL DIFFICULTIES</u>	12
2.1 AUTISM	12
2.2 AUTISM DIAGNOSTIC OBSERVATION SCHEDULE(ADOS)	15
2.3 EYE CONTACT AND GAZE DIRECTION	18
2.3.1 EYE CONTACT	19
2.3.2 JOINT ATTENTION	20
2.3.3 INTERVENTION FOR EYE CONTACT AND JOINT ATTENTION RESPONSES	21
<u>CHAPTER 3. HUMAN ROBOT INTERACTION IN AUTISM</u>	23
3.1 INTERACTIVE AND THERAPEUTIC ROBOTS DESIGNS FOR AUTISM	24
3.1.1 NON-HUMANOID ROBOTS	24
3.1.2 HUMANOID ROBOTS	25
3.2 DIFFERENT THERAPEUTIC APPROACHES FOR INDIVIDUALS WITH ASD	27
3.2.1 SELF-INITIATED INTERACTIONS	27
3.2.2 TURN-TAKING ACTIVITIES	29
3.2.3 EXPRESSION/EMOTION RECOGNITION AND IMITATION	30
3.2.4 JOINT/ EYE-GAZE ATTENTION	32
3.3 USING NAO IN AUTISM	35

<u>CHAPTER 4. METHODOLOGY AND EXPERIMENTAL RESULTS</u>	38
4.1 HARDWARE SETTING	39
4.1.1 NAO	39
4.1.2 CAPTURING SESSIONS AND ROOM DESIGN	42
4.2 PROTOCOL 1 (NAO SPY & FIND THE SUSPECTS)	45
4.2.1 PARTICIPANTS	46
4.2.2 PROTOCOL 1: GAME 1: NAO SYP (NS).....	46
4.2.3 PROTOCOL 1: GAME 2: FIND THE SUSPECTS (FTS)	47
4.3 PROTOCOL 1: EXPERIMENTAL RESULTS	51
4.3.1 EYE-GAZE FIXATION AND SHIFTING FREQUENCY ANALYSIS ...	53
4.3.2 EYE-GAZE PATTERN MODELING	62
4.3.2.1 HIDDEN MARKOV MODEL (HMM).....	62
4.3.2.2 EXPERIMENTAL RESULTS: HMM	65
4.3.2.3 VARIABLE-ORDER MARKOV MODEL (VMM)	65
4.3.2.4 EXPERIMENTAL RESULTS: VMM	67
4.3.2.5 DISCUSSION	69
4.4 PROTOCOL 2 (INTERVENTION SESSIONS).....	72
4.4.1 INTERVENTION SUB-SESSIONS	73
4.4.2 PROTOCOL 2: RESULTS	77
 <u>CHAPTER 5. CONCLUSION AND FUTURE RESEARCH DIRECTION</u>	87
 REFERENCE	93
 PUBLICATIONS.....	100

List of Tables

TABLE 2-1: Components of Autism Diagnostic Observation Schedule	16
TABLE 4-1: Time Duration Percentage in Game NS for ASD Group	55
TABLE 4-2: Eye Gaze Shifting Frequency in Game NS for ASD Group	57
TABLE 4-3: Time Duration Percentage in Game FTS for ASD Group	59
TABLE 4-4: Eye Gaze Shifting Frequency in Game FTS for ASD Group	59
TABLE 4-5: Classification Rates of HMM Algorithm for the TD and ASD Groups (Child Speaking and Listening Contets)	70
TABLE 4-6: Classification Rates of VMM Modeling for the “Child Speaking”	70
TABLE 4-7: Classification Rates of VMM Modeling for the “Child Listening”	71

List of Figures

Fig 2-1: Anatomy of Autism and Nervous system	14
Fig 4-1: NAO robot.....	40
Fig 4-2: Choregraphe user interface.....	41
Fig 4-3: Robot joints control pedal	42
Fig 4-4: Videos caregiver can see outside the room	43
Fig 4-5 a): Panorama view of the experimental room	44
Fig 4-5 b): Schematic robot based therapy session and video capturing setting	45
Fig 4-6: Kid is hugging NAO	47
Fig 4-7: Kid imitating “Tai Chi” activity	51
Fig 4-8: Overall mean percentage of eye gaze duration for all children with ASD after three sessions, dot line is the tendency of eye gaze duration	56
Fig 4-9: Overall mean frequency of eye gaze shifting for all children with ASD after three sessions, dot line is the tendency of eye gaze shifting	58
Fig 4-10: Overall mean percentage of eye gaze duration for all children with ASD after three sessions, dash-line is the tendency of eye gaze duration. The total percentage in each session greater than first game	61
Fig 4-11: Overall mean frequency of eye gaze shifting for all children with ASD after three sessions, dot line is the tendency of eye gaze shifting. Each sessions’s frequency is less than 0.0015 from the first game	61
Fig 4-12: Gaze labels and the corresponding gaze symbols (n = 4)	64

Fig 4-13: VMM maximum context length and eye gaze recognition	68
Fig 4-14: (Left) TD Group and (Right) ASD Group Gaze Classification Rate based on VMM algorithm.....	71
Fig 4-15: A) Kid is showing the lid; B) Kid is imitating the happy expression	76
Fig 4-16: A) Kid is following NAO's eye gaze and picking up one lid; B) Kid is pointing a specific lid which NAO is describing	77
Fig 4-17: Subject 19's behavior during baseline and intervention sessions, intervention sessions shows that T2.2 reached the peak	80
Fig 4-18: Subject 20's behavior during baseline and intervention sessions	81
Fig 4-19: Subject 21's behavior during baseline and intervention sessions, intervention sessions shows that T2.2 reached the peak	82
Fig 4-20: Subject 22's behavior during baseline and intervention sessions	83
Fig 4-21: Subject 23's behavior during baseline and intervention sessions, intervention sessions shows that T2.2 reached the peak	84
Fig 4-22: Subject 24's behavior during baseline and intervention sessions	85
Fig 4-23: Subject 25's behavior during baseline and intervention sessions, intervention sessions shows that T2.2 reached the peak	86

Chapter 1: Introduction

1.1 Autism Spectrum Disorders (ASD)

Autism is a general term used to describe a spectrum of complex developmental brain disorders causing qualitative impairments in social interaction and results in repetitive and stereotyped behaviors. Currently one in every 88 children in the United States are diagnosed with ASD and government statistics suggest the prevalence rate of ASD is increasing 10-17 percent annually [9]. Children with ASD experience deficits in appropriate verbal and nonverbal communication skills including motor control, emotional facial expressions, and eye gaze attention [10]. Currently, clinical work such as Applied Behavior Analysis (ABA) [11] [12] has focused on teaching individuals with ASD appropriate social skills in an effort to make them more successful in social situations [1]. With the concern of the growing number of children diagnosed with ASD, there is a high demand for finding alternative solutions such as innovative computer technologies and/or robotics to facilitate autism therapy. Therefore, research into how to design and use modern technology that would result in clinically robust methodologies for autism intervention is vital.

In social human interaction, non-verbal facial behaviors (e.g. facial expressions, gaze direction, and head pose orientation, etc.) convey important information between individuals. For instance, during an interactive conversation, the peer may regulate their

facial activities and gaze directions actively to indicate the interests or boredom. However, the majority of individuals with ASD show the lack of exploiting and understanding these cues to communicate with others. These limiting factors have made crucial difficulties for individuals with ASD to illustrate their emotions, feelings and also interact with other human beings. Studies have shown that individuals with autism are much interested to interact with machines (e.g. computers, iPad, robots, etc.) than humans [6]. In this regard, in the last decade several studies have been conducted to employ machines in therapy sessions and examine the behavioral responses of people with autism. These studies have assisted researchers to better understand, model and improve the social skills of individuals on the autism spectrum.

This thesis presents the methodology and results of a study that aimed to design a humanoid-robot therapy sessions for capturing, modeling and enhancing the social skills of children with Autism. In particular we mainly focus on gaze direction and joint attention modeling and analysis and investigate how the ASD and Typically Developing (TD) children employ their gaze for interacting with the robot. In the following section, we have a brief introduction of the existing assistive robots in the following section and how they have been used in autism applications.

1.2 Socially Assistive Robotics

Socially Assistive Robotics (SAR) can be considered as the intersection of *Assistive Robotics (AR)* and *Socially Interactive Robotics (SIR)*, which has referred to robots that assist human with physical deficits and also can provide certain terms of social interaction abilities [5]. SAR contains all properties of SIR described in [6], and also a few additional attributes such as: 1) user populations (different groups of users, i.e. elders; individuals with physical impairments; kids diagnosed with ASD; students); 2) social skills (i.e. speech ability; gestures movement); 3) objective tasks (i.e. tutoring; physical therapy; daily life assistance); 4) role of the robot (depends on the task the robot has been assigned for) [5].

Companion robots [7] is one type of SAR that are widely used for elderly people for health care supports. Research shows that this type of social robots can reduce stress and depression of individuals in elderly stage [8]. Service social robots are able to accomplish a variety of tasks for individuals with physical impairments [9]. Studies have shown that SAR can be used in therapy sessions for those individuals who suffer from cognitive and behavioral disorders (e.g. Autism). SAR provides an efficient helpful medium to teach certain types of skills to these groups of individuals [10] [11] [12].

Nowadays, there are very few companies that have been designing and producing socially assistive robots. The majority of existing SARs are not commercialized yet and because of being expensive and not well-designed user interfaces, they are mostly used for

the research purposes. Honda, Aldebaran Robotics and Hanson Robokind are the top leading companies that are currently producing humanoid robots.

Ideally socially assistive robots can have fully automated systems to detect and express social behaviors while interacting with humans. Some of the existing robot-human interfaces are semi-autonomous and they can recognize some basic biometrics (e.g. visual and audio commands of the user) and behavioral response. Besides, the majority of existing robots are very complicated to work with. Therefore in the last couple of years several companies have started to make these robots more user-friendly and responsive to both the user need and the potential caregiver commands [5].

Intelligent SARs aim to have the capability to recognize visual or audio commands, objects, and specific human gestures. Some of these robots have the ability of detect human face or basic facial expressions. For instance, ASIMO, a robot developed by Honda, it has a sensor for detecting the movements of multiple objects by using visual information captured from two cameras on its head. Plus its “eyes” can measure the distance of the objects from the robot [13]. Another example is from Aldebaran Robotics which designs small size humanoid robots, called NAO. NAO robot has two cameras attached that are used to capture single images and video sequences. This capturing module enables NAO to see the different sides of an object and recognize it for future use. Furthermore, NAO has a remarkable capability of recognizing faces and detecting moving objects.

Both of the aforementioned robots have speech recognition system. They can interpret voice commands to accomplish a certain set of tasks which have been pre-programmed in the system. NAO is able to identify words for running specific commands. However ASIMO is able to distinguish between voices and other sounds. This feature empowers ASIMO to perceive the direction of human's speaker or recognize other companion robots by tracking their voice [14]. These robots can also speak in many different languages. For example, NAO can speak in English, French, Chinese, Japanese and other languages up to more than ten languages. This feature gives the robot a great social communication functionality to interact with humans from all over the world.

1.3 Using Socially Assistive Robots for Autism Therapy

Socially assistive robots are emerging technologies in the field of robotics that aim to utilize social robots to increase engagement of users as communicating with robots, and elicit novel social behaviors through their interaction. One of the goal in SAR is to use social robots either individually or in conjunction with caregivers to improve social skills of individuals who have social behavioral deficits. One of the early applications of SAR is autism rehabilitation. As mentioned before, autism is a spectrum of complex developmental brain disorders causing qualitative impairments in social interaction. Children with ASD experience deficits in appropriate verbal and nonverbal communication skills including motor control, emotional facial expressions, and gaze regulation. These

skill deficits often pose problems in the individual's ability to establish and maintain social relationships and may lead to anxiety surrounding social contexts and behaviors [1]. Unfortunately there is no single accepted intervention, treatment, or known cure for individuals with ASD.

Recent research suggests that children with autism exhibit certain positive social behaviors when interacting with robots compared to their peers that do not interact with robots [2][3][4][5][6]. These positive behaviors include showing emotional facial expressions (e.g., smiling), gesture imitation, and eye gaze attention. Studies show that these behaviors are rare in children with autism but evidence suggests that robots trigger children to demonstrate such behaviors. These investigations propose that interaction with robots may be a promising approach for rehabilitation of children with ASD.

There are several research groups that investigated the response of children with autism to both humanoid robots and non-humanoid toy-like robots in the hope that these systems will be useful for understanding affective, communicative, and social differences seen in individuals with ASD (see Diehl et al., [6]), and to utilize robotic systems to develop novel interventions and enhance existing treatments for children with ASD [13] [14] [15]. Mazzei et al. [16], for example, designed the robot "FACE" to realistically show the details of human facial expressions. A combination of hardware, wearable devices, and software algorithms measured subject's affective states (e.g., eye gaze attention, facial expressions,

vital signals, skin temperature and EDA signals), were used for controlling the robot reactions and responses.

Reviewing the literature in SAR [5] [6] shows that there are surprisingly very few studies that used an autonomous robot to model, teach or practice the social skills of individuals with autism. Amongst, teaching how to regulate eye-gaze attention, perceiving and expressing emotional facial expressions are the most important ones. Designing robust interactive games and employing a reliable social robot that can sense users' socio-emotional behaviors and can respond to emotions through facial expressions or speech is an interesting area of research. In addition, the therapeutic applications of social robots impose conditions on the robot's requirements, feedback model and user interface. In other words, the robot that aims for autism therapy may not be directly used for depression treatment and hence every SAR application requires its own attention, research, and development

Only a few adaptive robot-based interaction settings have been designed and employed for communication with children with ASD. Proximity-based closed-loop robotic interaction [29], haptic interaction [30], and adaptive game interactions based on affective cues inferred from physiological signals [31] are some of these studies. Although all of these studies were conducted to analyze the functionality of robots for socially

interacting with individuals with ASD, these paradigms were limitedly explored and focused on their core deficits (i.e., Facial expression, eye gaze and joint attention skills).

Bekele and colleagues [32] studied the development and application of a humanoid robotic system capable of intelligently administering joint attention prompts and adaptively responding based on within system measurements of gaze and attention. They found out that preschool children with ASD have more frequent eye contact toward the humanoid robot agent, and also more accurate respond in joint attention stimulations. This suggests that robotic systems have the enhancements for successfully improve the coordinated attention in kids with ASD.

Considering the existing SAR system and the major social deficits that individuals with autism may have, we have designed and conducted robot-based therapeutic sessions that are focused on different aspects of social skills of children with autism. In this thesis we employed NAO which can be remotely controlled to communicate with the children. We conducted two different protocols to examine the social skills of children with autism and provide feedbacks to improve their behavioral responses. The contribution of our work has been introduced in Section 1.4 and the details of the game setting, experiments, modeling and analysis are provided in Chapter 4.

1.4 Thesis Contributions

The objective of this thesis is two-folded:

1. (Protocol 1) How and in what capacity socially assistive robot can help us to analyze and model the social behaviors (eye-gaze direction) of children with autism?

2. (Protocol 2) How to employ NAO to measure the social skill level of each participant. How to design social games to improve the social skills of children with ASD?

To answer these questions, we developed two protocols and executed them on a group of participants, the following steps for two protocols:

In Protocol 1:

- Designed two generic games: (NAO Spy & Find the Suspect).
- Recruited 14 high functioning children with ASD (age range 7 - 17) and 7

TD kids (same age range)

- Manually coded eye gaze direction of every frames (i.e. gaze at/averted).
- Analyzed gaze direction fixation and shifting.
- Utilized Markov models (HMM & VMM) to analyze gaze patterns of ASD

and TD participants.

In Protocol 2:

- Designed a detailed task-oriented games focused on one social skill at a time.
- Customized an intervention session for each subject.
- Conducted an online intervention setting and behavioral coding.

Protocol 1 is mostly focused on eye-gaze pattern analysis. Based on the manually coded data, we analyzed the gaze pattern of children with ASD and compared it with the gaze of the control TD group. We also used Markov modeling to evaluate how the gaze responses of TD and individuals with ASD are separable. In addition how the gaze patterns of individuals with ASD change during different sessions of interaction with NAO.

Protocol 2 focused on the behavioral responses of children with ASD for different social skills (e.g. facial expression recognition and imitation, following NAO's gaze direction and pointing, responding to the verbal questions, etc.). We have three baseline sessions that measure the initial social level of every child. Then through the intervention phase we provide some feedbacks and guidance to enhance their social skills while communicating with NAO. After each intervention session, we scored the child's responses and we decided whether s/he needs more intervention sessions or we can move on to other social skills that need an intervention. The details of both protocols have been explained in Chapter 4.

1.5 Organization

The thesis is organized as follows. Chapter 2 introduces the autism and some of the deficits that individuals with ASD have. Chapter 3 is a literature review of existing therapy and robot-based interaction studies for individuals with autism. Our robot-based therapy session and the details of the designed games and data collection in addition to the experimental results are introduced in Chapter 4. The thesis is concluded in Chapter 5.

Chapter 2: Cognitive and Physiological Difficulties in Autism

2.1 Autism

Individuals with autism spectrum disorder experience verbal and nonverbal communication impairments, including motor control, emotional facial expressions, and eye gaze attention. Oftentimes, individuals with high-functioning autism have deficits in different areas, such as (1) language delay, (2) difficulty in having empathy with their peer and understanding others emotions (i.e. facial expressions recognition.), and more remarkably (3) joint attention (i.e. eye contact and eye gaze attention). Autism is a disorder that appears in infancy [23]. Although there is no single accepted intervention, treatment, or known cure for ASDs, these individual will have more successful treatment if ASD is diagnosed in early stages. At the first glance at the individual with autism, you may not notice anything odd, however after trying to talk to her/him, you will understand something is definitely not right [77]. S/He may not make eye contact with you and avoid your gaze and fidget, rock her/his body and bang her/his head against the wall [77].

In early 1990s, researchers in the University of California at San Diego aimed to find out the connections between autism and nervous system (i.e. mirror neurons). Mirror neuron [77] is a neuron that is activated either when a human acts an action or observes the same action performed by others. As these neurons are involved with the abilities such as empathy and perception of other individual's intentions or emotions, they came up with

malfunctioning of mirror neuron in individuals with ASD [77]. There are several studies that focus on the neurological deficits of individuals with autism and studying on their brain activities. Figure 2-1 demonstrates the areas in the brain that causes the reduce mirror neuron activities in individuals with autism.

THE ANATOMY OF AUTISM

People with autism show reduced mirror neuron activity in the inferior frontal gyrus, a part of the brain's premotor cortex, perhaps explaining their inability to assess the intentions of others. Dysfunctions of mirror neurons in the insula and anterior cingulate cortex may cause related symptoms, such as the absence of empathy, and deficits in the angular gyrus may result in language difficulties. People with autism also have structural changes in the cerebellum and brain stem.

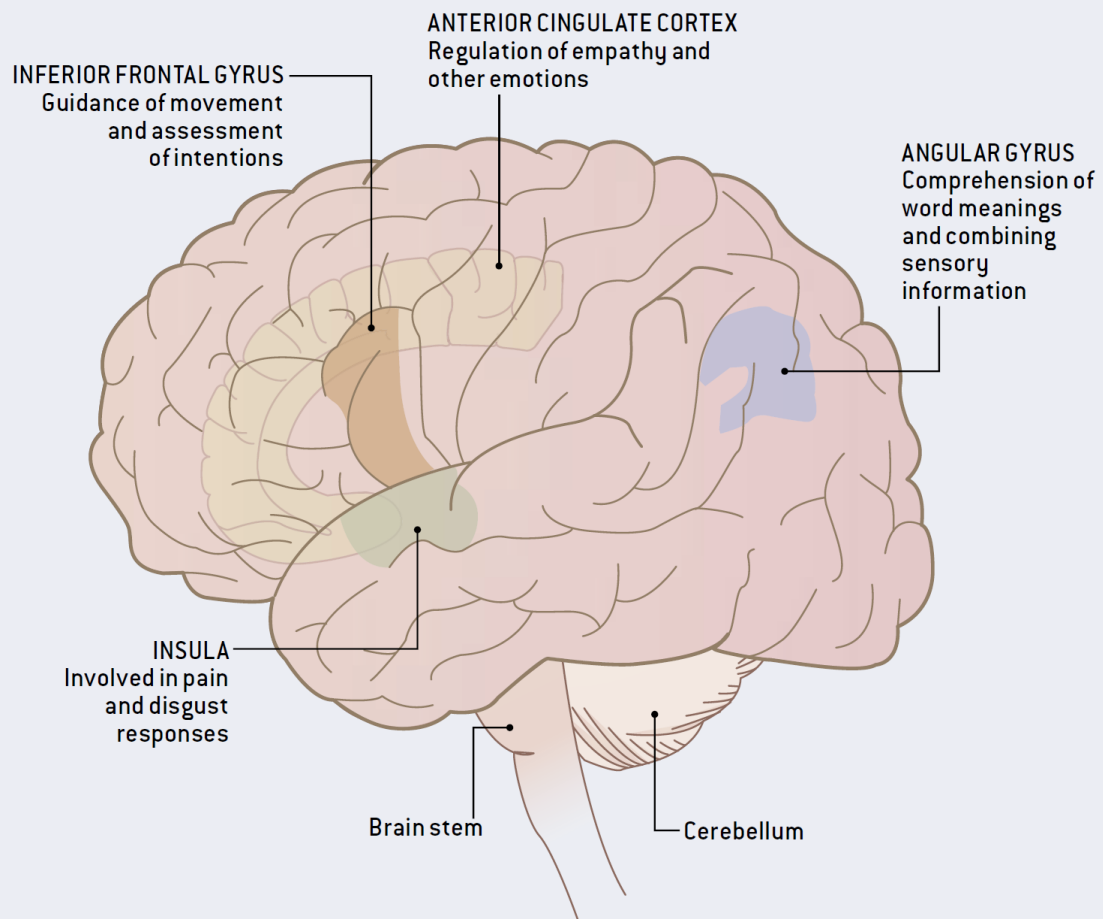


Fig 2-1: Anatomy of Autism and Nervous system [1]

Individuals with autism might also have several other unusual social developmental behaviors that may appear in infancy or childhood. For instance children with autism show less attention to social stimuli (e.g. facial expressions, joint attention), and respond less when calling their names. Compared with typically developing children, older children or adults with autism can read facial expressions less effectively and recognize emotions behind specific facial expressions or the tone of voice with difficulties [26]. In contrast to TD individuals, children with autism (i.e. high-functioning, Asperger syndrome) may be overwhelmed with social signals such as facial behaviors and expression and complexity of them and they suffer from interacting with other individuals, therefore they would prefer to be alone. That is why it would be difficult for individuals with autism to maintain social interaction with others [28].

In order to diagnose and assess the aspects and score the social skill level of an individual with autism, several protocols are available. One of the commercially available protocols is called Autism Diagnostic Observation Schedule (ADOS) [65] that consists of four modules and several structured tasks that are used to measure the social interaction levels of the subject and examiner. We are inspired by ADOS in designing our intervention protocols later described in Chapter 4. Hence, we briefly review ADOS in the next section.

2.2. Autism Diagnostic Observation Schedule (ADOS)

The Autism Diagnostic Observation Schedule (ADOS) is a standardized protocol for observing the social and communicative behaviors associated with autism. Eight tasks have contained in ADOS, as shown in the table below. 20-30 minutes are required for an examiner to complete the entire tasks [65].

Task	Target behaviors	Focused on Task
Construction task	Asking for help	Construction activity, Initiating conversation
Unstructured presentation of toys	Symbolic play Giving help to interviewer	Imagination
Drawing game	Taking turns in a structured task	Turn taking and joint attention
Demonstration task	Descriptive gesture and mime	Demonstrate a gesture and facial expressions
Poster task	Description of agents and actions	Language ability
Book task	Telling a sequential story	Language ability
Conversation	Reciprocal communication	Verbal Skill

Socio-emotional questions	Ability to use language to discuss socio-emotional topics	Verbal, Facial expression and joint attention skills
----------------------------------	---	--

TABLE 2-1: Components of Autism Diagnostic Observation Schedule

As shown in TABLE 2-1, each task contains one or few aspects of social skills including turn taking (refers to the process by which people in a conversation decide who is to speak next), joint attention, reading emotions etc. Right after the interview, examiner would provide a general ratings based on the observation in all the tasks which have been targeted to code.

ADOS contains four modules that are designed for specific age range and set of social developmental abilities. Examiner may use ‘Module 1’ if the child uses a little or no phrase speech however if s/he utilizes phrase speech but do not speak fluently ‘Module 2’ may be employed. Some examples of Modules 1 or 2 are responding to name, social smile, and bubble play. ‘Module 3’ is used for younger children who are verbally fluent and ‘Module 4’ is employed for adolescents and adults with fluent verbal skills. Modules 3 or 4 can include communication, and exhibition of empathy or comments on others' emotions. Considering these four modules, ADOS can provide scores regarding these four areas (1) Reciprocal social interaction, (2) Communication/language, (3) Stereotyped/restricted

behaviors, and (4) Mood and non-specific abnormal behaviors. In our study we employed ADOS and some tasks described in it for introducing new robot-based games and social interaction to children that will be explained in Chapter 4 section 2.3.

2.3 Eye Contact and Gaze Direction

Eye Contact sometimes referred to as eye-to-face gaze or gaze behaviors [67] that employs better verbal and nonverbal conversations [66]. In early developmental stages, children employ eye contact to regulate the face-to-face social interaction. Later, it coordinates the visual attention between another individual and object of an interest [68]. It has been shown that eye gaze regulation has important effects and influences in language and verbal information as well [69].

One of the earliest and most noticeable indicators of developmental delays and autism spectrum disorder is the deficit in dyadic (i.e. eye-to-face) and triadic eye gaze (i.e. joint attention directed at a third party or object) in social communicative behaviors [66]. As eye contact serves important social roles, and failure to emit this important signal may have significant drawbacks and implications in the educational, relational and social life of individuals with ASD. Given these potential negative outcomes, in this thesis we aim to utilize a humanoid robot to better model, analyze the gaze patterns of children with ASD and ultimately teach them the typical ways to regulate and control their gaze behaviors that

eventually would help them use the skills in their daily life. In the following section we will introduce and study the effects of eye contact and visual joint attention in more details.

2.3.1. Eye Contact

Early behavioral investigations about eye contact responses demonstrate that if children fail to orient toward the instructor, they would most probably fail to respond and learn a new concept [70, 71, and 72]. Beattie (1981), Lalljee and Cook (1972), Stephenson and Rutter (1974) [48 and 49] did some analysis on the effect of continuous gaze on the fluency of speech; they suggested that the conversation is more fluent when people can see each other facial behaviors. Besides, eye gaze would be necessary to make the subjects feel comfortable during a conversation. But it is also true that too much gaze reduces the quality of conversation.

The intimacy equilibrium model has been developed by Argyle and Dean (1965) [49] which is the most elaborate attempt to explain like-look relationship during interaction. The gaze significance has been worked by Adam Kendon (1976) [50], his statistic results showed that a subject spent less than 50% of his speaking time looking at his partner, but more than 50% of his listening time looking at the partner. Argyle (1974) [51] interpreted that the decrease time in gaze during speaking because of the speaker needs to think about what s/he is going to say without any distraction. On the other hand, the listener needs to

show his/her attention and try to collect more information during the listening time [49, 50, and 51].

As mentioned above, eye gaze is one of the important aspects of social communication and that is also an important deficits in autism. Hence, we decided to focus on this aspect for our research in Protocol 1. Games design and data analysis are based on this fact.

2.3.2. Joint Attention

Understanding and utilizing joint attention (i.e. directing gaze toward a third party or object) is one of the crucial deficits of individuals with autism. Joint attention is referred to as an important non-verbal cue to transfer the focus of an individual to another object, by using gaze fixations, pointing that may accompany by head indication. One of the early research in this area in 1975, aimed to study children's ability to follow eye gaze of others. This early research showed that adults can bring certain objects in the environment to get kid's attention using eye gaze [29]. By definition, joint attention seems to be necessary for functional speech, and deficits in this basic skill distinguish infants with ASD from typically developing children [30]. For instance, children with autism may stare at your finger as you are pointing to an object, and they consistently fail to aim to an object using their gaze direction, or pointing [26, 31, 30].

Insufficient eye gaze communication can be considered as some part of the joint attention deficit, and this type of disorder is considered as the noticeable symptom of children with autism [33].

2.3.3. Intervention for Eye Contact and Joint Attention Responses

Considering eye contact and joint attention deficits in individuals with ASD, several studies have been conducted to improve the eye contact responses of these individuals. The interventions were mostly based on a human-based therapy sessions. One group of therapy sessions utilized vocal or physical prompting for improving the eye contact responses. This category of intervention have shown to be successful for some group of individuals, but it has some possible disadvantages, including (1) some children may resist the use of physical prompts and therefore produce some interfering behaviors, and (2) using (vocal or instructional) prompts need more additional prompts to fade which results in slower skill acquisition.

In early 1980s, a new area of studies targeted the eye contact studies where they aimed to teach gaze regulation skills through various social-interactive strategies [73]. A group of procedures (including peer modeling, role playing, contingent imitation, time delay, and naturalistic behavior modification techniques) have been used and shown to moderately increase in a variety of social behaviors, including eye contact and joint attention [66]. In some recent studies some motivational variable and ‘*extrinsic*’ reinforces

(in the form of social praise, or edible items) have been utilized to encourage the children for more eye contact responses. Although it has been illustrated that this type of reinforcement may be a limiting factor for gaze shifting studies when teaching joint attention responses [74, 75] but it is still widely used in therapy sessions. In a recent study conducted in 2012, Plavnick and Ferreri proposed to use *mind training* to improve the social interactions of the children with ASD. This is another approach that has been shown can be applied for improving the eye contact responses of children with autism [76].

In this thesis we utilized a humanoid robot, called NAO, for designing and developing a robot-children interaction setting. This study aimed to improve the social communications of individuals with ASD and provide a platform to better model and analyze the gaze response of ASD and TD children. In order to detect, track, and improve the social interaction levels of individuals with ASD we used verbal commands, extrinsic reinforcements and social praises. We will explain more about the designed games and interventions in Chapter 4, but we will review the existing robots and human-robot interaction setting for autism intervention in Chapter 3.

Chapter 3: Human Robot Interaction in Autism

Children with ASD experience deficits in appropriate verbal and non-verbal communication skills including motor control, emotional facial expressions, eye-gaze attention, and joint attention. Many studies have been conducted to identify therapeutic methods that can benefit children with ASD [52]. However, only a few groups used humanoid robots for teaching or practicing social communication skills [53, 54, 55, 56, 57, 58, and 59].

For some of the social behaviors, such as eye contact, joint attention, facial expressions recognition, that are rarely seen in interactions of children ASD, several evidence suggest that robots can trigger them more effectively than human [78]. Researchers observed that individuals with ASD have more interest toward a robot therapeutic partner than a human. In most cases participants showed better speech and movement imitation compared with response to a human partner [79]. Although a recent case study [52] which was done by Ricks (2010) suggests that this approach might have clinical utility, still this area is obviously in its infancy. Studies have shown that positive feedback from the robot on the participants' performance is an effective way to encourage children with ASD to communicate more [52]. Other studies have also examined the use of affect recognition (e.g. emotional state, arousal level) based on psychophysiological responses to modify the behaviors during a robotic game. However, there is limited

information on the utility of humanoid robots' positive feedback in interventions for individuals with ASD.

3.1 Interactive and Therapeutic Robots Designs for Autism

Different types of robots have been used in autism research for various purposes. Some researchers have been attempting to utilize a realistic human appearance [56], while others have created robots with very mechanical forms [54], and others have developed robots with a cartoonish or animal form [58]. Generally speaking different categories of robot that have been used for autism research can be placed either into Non-Humanoid and Humanoid robots group [52], which will be explained in the following sections.

3.1.1 Non-Humanoid Robots

Non-humanoid robots are those robots which do not have the same body joint and facial appearance as human does. It contains those animal like, cartoonish, or non-human like appearances. These robots have been used by several researchers in the last two decades. This category of robots is generally easier to design and develop and less expensive, therefore, several of initial robot-human interaction for individuals with ASD was conducted by non-humanoid robots. The bubble-blowing robot at USC (while children approached it, the robot will nod head make voice or blow bubble from lower part of robot body), for instance, was not a human form robot and can be built simply [53]. Another non-humanoid robot used by researchers from University of Hertfordshire called Labo-1

[54], which can play tag games (tip you're it or tig), with children. (In the game, several children play with the robot together, the robot uses its heat sensor to approach kids as a type of interaction.) In Yale University, researchers were using a mobile robotic dinosaur named Pleo who can show emotions and desires by using its sounds and body movements [55]. Children in the clinic have been helped by Pleo's pet-like appearance, expressiveness, and versatility.

The reason why researchers using non-humanoid robots is that they found out that when children with ASD see humans, they usually will choose to avoid and not to interact with them. On the contrary, an animal shape or toy shape robot would be easier for kids to engage with and have a better interaction.

3.1.2 Humanoid Robots

Humanoid robots generally provide the human-like appearance and consist of body parts such as humanoid head, body and arms. Advanced humanoid robot would be able to move different parts of its body to walk or dance (NAO), some of the humanoid robot also has the capability to show facial expressions (e.g. ZENO). This type of robot unlike non-humanoid robot, they have the ability to accomplish more complicated social communication tasks than non-humanoid robot, but those tasks will be less complicated than human-human interaction. This capability can help us to design interaction sessions

and therapeutic sessions for children with autism and assist them with improving their social behaviors.

Robins from University of Hertfordshire, who is one of the pioneers which employed a study to evaluate the importance of robot's appearance for autism research. A doll-like robot called Robota were asked to interact with children with autism [56]. This example shows that children appeared to be more interested in interaction with less-human like robots. Researchers conclude that children with ASD would prefer a simple non-complexity and less details of human but still hold the humanoid form. So, a robot called KASPAR has been developed by Robins to fit this design criteria [57].

Similar conclusions have been made by researchers at the National Institute of Information and Communications Technology (NIICT) in Japan. They found out that when kids with ASD have interaction with their designed robot called Infanoid, the children tend to pay more attention on the mechanical parts of the robot's body than communicating with the robot itself [52]. A small soft snowman-shaped robot, called Keepon, was designed to represent as a simple, repeatable, mechanical robot regarding the reason mentioned above [58]. Keepon can express its emotions conveyed by shaking, rocking, and bobbing up and down which can be used as a super fun toy companion for kids with ASD. Another humanoid robot which has been designed by researchers at the University of Pisa, is known as FACE. The purpose of their project is to create a robot as realistic as possible to a human

face for evaluating how human react as the FACE displays different expressions [59]. (During the sessions, child (IQ around 85) with autism did not show any interest in FACE at the beginning. However, with verbal suggestion, kid replied the expression by using a word “damsel” which is from a fairy tale, though the FACE showing a sad expression on it.) This study suggested that by using FACE, it is possible to extend emotional recognition skill to children with autism.

In the last few years several different types of non-humanoid and humanoid robots have been used for autism therapeutics sessions that we will discuss about them in the next session.

3.2 Different Therapeutic approaches for Individuals with ASD.

As explained in Chapter 2, different individuals with autism might suffer from various types of social or developmental behavior. Therefore in order to have an effective therapeutic intervention setting we need to focus on various tasks and treatments. Bellow we will provide different intervention aspects that majority of children with ASD may suffer from.

3.2.1 Self-Initiated Interactions

The difficulty for initiating a social conversation or interaction is one of the impaired social skills of children with ASD. This problem may represent as difficulty for

conveying what they want, and why they want it. For example, when a child in early age who wants to urinate, he might have to ask for parent's help, rather than hold it there or let it be. Clinicians try to encourage those kids to ask to play certain toys and a reward will be given after they did it. Instead of human therapists, researcher extended this idea using robots to encourage the children to engage the robot proactively. The robot has built at USC [53] which has a large button on its back, and it was programmed to encourage social interaction with children. For example, the robot will nod its head and make a sound to encourage the kid to approach it; when the kid walk away, it will move its head down and make sad kind of sound to imply the child and ask him/her to come closer to the robot. If the child presses that button on the robot, it will blow bubbles and turn. In this study, one hundred minutes experiments have been recorded, three different conditions have been considered which are the time kids spent near 1) the wall; 2) the parent; and 3) behind the robot. Kids have been separated in two groups: '*Group A*' (children like the robot) and '*Group B*' (children do not like the robot) total number of eight children with ASD. The result shows that the Group A spent more than 60% of the time playing with the robot, and Group B spent more than 50% of the time showing the negative reaction (i.e. go away from robot, play with himself) from avoiding the robot. This study might not be very convincing because it is totally free play with the robot; the experimental settings haven't kept the same, and the limited numbers of participants. Also without control group like typically developing they could not compare the differences of ASD and TD children, within the

robot games. However, it shows the capability of encourage children to communication with robot, and lead the conversation [53].

3.2.2 Turn-Taking Activities

At the University of Hertfordshire and the University of California, researchers have built small mobile robots that focused on helping children with ASD in turn-taking behaviors [54 and 53]. It is easily to found out that children with ASD have a hard time to allow their conversation partner to participate. The researchers try to use these robots to help them become accustomed to waiting for responses after they say or do something. Labo-1 built by the University of Hertfordshire, which can play a game called tag with children. This game will forces them to alternate between engaging and avoiding the robot [54].

Labo-1 is a mobile platform that has an AI system resembled in a sturdy flat-topped buggy. Children have been allowed to freely play with Labo-1 as a teacher was deciding about the how to switch between different games/sessions considering children appears (i.e. difference reactions of children like tired or less interested into robot). From their initial trials, children were in overall happy to play with robot. At the beginning of the game, the robot showed several simple behavior patterns, such as going forward and backward. Kids showed positive response to these behaviors and enjoy to keep playing with Labo-1. Children were also enjoyed interacting with the robot while it used a feature called ‘*heat*

following behavior', they moved away from the robot and see if the robot can follow or not. There were five trials in total, three of them lasted around four minutes, and the remaining two had duration of approximately fourteen minutes. Researchers realized that the issues that may cause this difference might be related to the levels of the children's functioning. Since children are not in complete control the robot's actions, and children's response were totally different, some of them either walked or crawled around the room, some of them just simply lay on the floor to interact with robot only use arm movement [54].

During the interactions, it is obvious to notice that robot need more advanced behaviors to be developed and the scenario should have more control for data analysis and get more convincing results. Also the functioning level become another important element that need to be considered.

3.2.3 Expression/Emotion Recognition and Imitation

Another import difficulty of individuals with ASD is to recognize the expressions and emotions, besides appropriately imitating them. Studies show that kids with ASD have a hard time recognizing emotions and facial expressions. It would be difficult for them to deliver their emotions through their faces actions. Researchers pointed out that to kids with ASD, such emotion type information which contained faces or eye contact can result overwhelming or sensory overload. For example, a person could smile twice, and the child

with ASD might pick two entire different expressions from those two smiles. Robot can provide more constancy repeatable consistent behaviors than human does, and it would be a better way to teach children expressions and emotions.

KASPAR, a child-sized doll like robot which has a silicon-rubber face on it, developed by the University of Hertfordshire has been used to show bodily expressions by move head and arms. KASPAR was operated via wireless remotely, sessions are designed to allow the children to have free play interaction with robot. Some behaviors had been pre-programmed in the robot, those behaviors allows KASPAR to show several facial expressions, hand waving and drumming on the tambourine on its legs to express different emotions. During the interaction, three types of touch using the hands had been identified: grasping (different tension levels), stroking, and poking. The forces of touching can be detected by the tactile sensors equipped different places of KASPAR's arms, hands, face, and shoulders. By detecting different levels of touching, KASPAR would provide different movements or expression to tell the children the emotions or feeling of it. Emotion and facial expressions recognition could be taught via these outputs KASPAR given. The limitation of this study is very few numbers of children (five children in total) had participated in this study. Besides limited facial expressions (happiness, displeasure, surprise etc.) have employed in the robot system, and those expressions are hard to distinguish by the images they provided. There is no verbal communication between kids and robot, which is another weakness of this study [57].

FACE is a robot designed at the University of Pisa point to closely approximate a real human face and show the detail facial expressions. Children would be asked to imitate those expressions to practice their ability in facial expression recognition and imitation. Certain scenarios (i.e. 1) facial expression association: a) facial matching, b) emotion labeling; 2) emotion contextualization) would be given to kids and ask them to pick up an appropriate emotional expression for FACE to make. Several experiments have been implemented to help the children to generalize the information they learn from the therapy sessions. After practicing with FACE, the children were tested using the Childhood Autism Rating Scale and the results showed that while working with FACE, the ability of categories emotions and expressions for all kids (total number of 4 kids) have been improved. Also, researchers found out that those children can imitate facial expressions from FACE better than from humans, and it will be easier for therapist because of the automate repeatable of robots process. However, still very limited number of kids participated in the study that made the results somehow not quite untenable [59].

3.2.4 Joint/Eye-gaze Attention

One of the major deficiencies of individuals with ASD is the lack of continuous concentration on the same object [33]. Joint attention is a concept of remain focus on specific things. Helping children with ASD on this aspect, would also help them to achieve the success in learning other skills. Keep practicing joint attention would give them more

understandings of what others are aware of them, what they are aware of other and they both aware of same object.

For this purpose, researchers from National Institute of Information and Communications Technology in Japan developed the Keepon robot. For seeking the possible responses of using interpersonal communication, both ASD and TD kids have been recruited in the study. A silicon-rubber made yellow snowman like body covered above the mechanical parts of Keepon, with two eyes on the upper part of the robot, and a nose (microphone embedded) in between. Lower part which is the belly of Keepon can easily deform whenever it changes posture and when people touch it [58].

With four degrees of freedom (± 40 degrees of nodding, ± 180 degrees of shaking, ± 25 degrees of rocking, and bobbing with 15 mm stroke), Keepon is able to perform two action mode: Attentive action and Emotive action. In attentive mode, Keepon will orient its face/body to a certain object around it, two CCD cameras in its eyes would be able to making eye contact and joint attention with the target; in emotive mode, Keepon will still its attention in a certain direction, and rocks its body up and down or left and right to express its emotions like pleasure and excitement. In both modes, Keepon will also making little sounds to drag the attention of people around it or give some feedback when people touches it or grabs it [58].

There are two operation mode to control Keepon, either automatic mode or manual mode. In automation mode, locations of a human face, a toy with a predetermined color, and an optically moving region would be detected. An Attention Map are written inside of Keepon, it orients its body (eye gaze) to most salient point on the Attention Map; its emotional expression is determined by the type (face/toy/motion) and the saliency value of the point of interest. In manual mode, based on the onboard cameras and listens to the sounds captured by the onboard microphone, a person can easily control Keepon via a remote computer. The operator only need to click the interest on the panoramic map to displays emotional expression on Keepon [58].

After more than a year and half (over 500 child-sessions), this research provides some interesting results. Children who have autism and PDD, they usually have difficulty with communicating with others, which however were able to approach Keepon with security and curiosity, and had a good time with it. Some of the kids even learned how to share their pleasure with other people which extended the dyadic interaction to triadic interaction. Different children have different style with communicating with Keepon, based on those different reactions researchers might predict different personality of those children [58].

This study shows some promising conclusions, but still cannot provide statistic results to readers. Overall, the experiment settings have been considerate be thoughtful,

though the sessions kept in free play mode. Good amount of participants enrolled in the study, which makes the conclusion more convincing. The aim of the study has been fully illustrated during sessions, joint attention has ran through both action modes and provided a good feedback from the participants. Keepon's voice needs to be improved, not just making simple noise, but also have a complete conversation would be better. More statistic results needs to be analyzed in the future to compare both children with ASD and TD kids.

3.3 Using NAO in Autism

NAO is a multifunctional humanoid robot that was developed by Aldebaran Robotics and as it has capabilities such as making different gesture, moving different arm and leg movement and hear orientations, It has been used for different human-robot interaction sessions. In this section we will talk about the existing interactions sessions that were conducted by NAO and later in the next chapter we will explain about our therapy sessions and designed game based on NAO for children with ASD.

In University of Teknologi MARA, NAO was used to conduct seven interactions modules for interacting kids with autism. Each module lasts four minutes, and one minute break was provided between two sessions. Different interaction tasks have been contained in those modules (i.e. static interaction, joint attention, basic language skills). Frequency of child looking at robot and duration of each occurrence of interaction has been reported. After all, they concluded that those 7 modules can be applied to develop human-robot

integration therapy sessions for children with autism [80]. Same year, these researchers use 5 of those 7 modules did a case study, with the same setting, they recruited one high-functioning (with IQ 107) to complete those 5 tasks. They aimed to discover whether that child can provide a better exposure behavior with robot compared with the activity in the class. After running the five tasks for only one instance, they concluded that the child behavior have been improved significantly with robot than in the class, they also suggested that humanoid robot NAO can be used as a major platform to support and initiate interaction with children with ASD [81]. After this case study, they recruited other 5 children with ASD (low IQ, average around 50) and did the same experimental interaction sessions with them. For out of five children showed better performance during robot interaction compared with daily in-class performance [82]. Further research have been done by this group, they added emotion recognition module into the interaction sessions. Five body gesture emotions (hungry, happy, mad, scared, and hug/love) have been implemented in the program. Two boys have been enrolled in this study, and after finished the session, researchers pointed out that NAO has the potential capability to teach head and bod posture related to social emotions for children with autism, without provided any statistical analysis only based on observations [83]. This group has been initiated working with NAO for autism therapeutic session and implementing and compared different scenarios based on NAO. Reviewing the existing papers demonstrate that the number of participants and interaction sessions for these studies are very limited. They have used only

one session for each subject. Therefore they could not analyze the social responses of individuals with ASD statistically.

In our study we employ NAO since it has several functionalities that are embedded in it (e.g. text-to-speech, tactical sensor, face recognition, voice recognition etc.). This would help us to build a social communicative tasks for human-robot interaction. Based on the size of the robot and the friendly appearance of the robot we design, conduct and analyze the gaze related responses of ASD individual and compare it with TD control group. The details of our experiment and the results will be discussed in Chapter 4.

Chapter 4: Methodology and Experimental Results

As described in Chapter 2, individuals with ASD are interested in interacting in less complicated and easy-to-use social environments. Therefore one applicable approach is to use robot-based therapy sessions instead of conventional human-based ones. In this study we employed a humanoid robot, called NAO, and designed a set of interactive games that controlled the robot, to socially interact with a group of children with ASD. We captured audios and videos of the participants (14 ASD and 7 TD) to analyze their facial behavioral activities (e.g. gaze direction, facial expressions, etc.) and provide efficient feedback to improve their social skills.

Our experiments were conducted through two protocols:

Protocol 1: We designed two sets of games (e.g. “*NAO Spy*” (NS) and “*Find the Suspects*” (FTS)) which encouraged children to be involved in conversational contexts. This protocol focuses on different aspects of social interactions. As described in the following sections, this thesis particularly focused on annotation, analysis, and modeling of the gaze direction and comparing the gaze information of ASD versus typically developing children in interaction with NAO.

Protocol 2: This protocol we utilized the results and outcome of the first protocol to design therapeutic social games (base line and intervention sessions)

which are focused on social skills in more systematic ways. Specifically, we focused on five tasks (e.g. gaze direction and joint attention regulation, facial expression recognition and imitation, social conversation, etc.) and during the experiments the behavioral responses along with the visual responses and arousal levels (i.e. thermal conductance) are captured simultaneously. In this multidisciplinary study we recruited 7 ASDs (7-17 years old) and analyzed different social responses while they interacted with NAO.

The details of the designed video capturing systems, the aforementioned protocols and the experimental results are provided in the following sections.

4.1 Hardware setting

4.1.1 NAO

We used a humanoid robot called NAO developed by Aldebaran Robotics in France [ref]. NAO is 58 cm (23 inches) tall, with 25 degrees of freedom this robot can conduct most of the human behaviors. It also features an onboard multimedia system including, four microphones for voice recognition, and sound localization, two speakers for text-to-speech synthesis, and two HD cameras with maximum image resolution 1280×960 for online observation. As shown in Figure 4-1, these utilities are located in the middle of the forehead and the mouth area. NAO's computer vision module includes facial and shape recognition units.

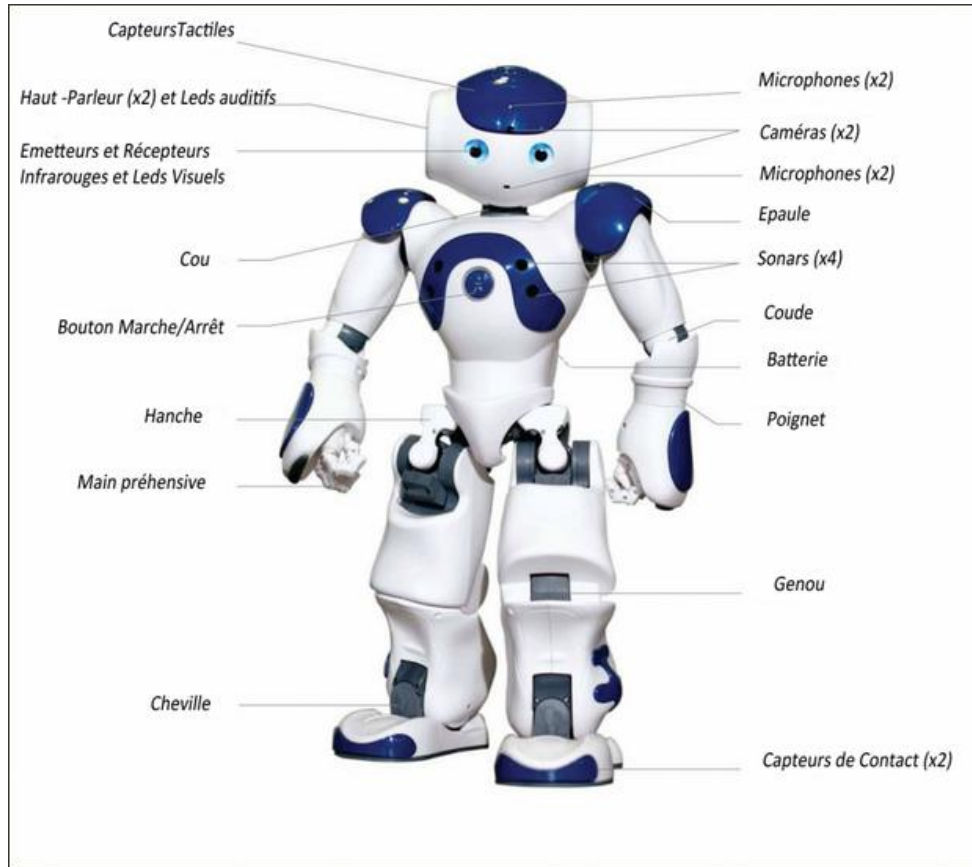


Fig 4-1: NAO robot

By using Choregraphe software (Shown in Figure 4-2), researcher can easily control NAO remotely. Inside the user interface we have access to NAO's cameras. It is also easy to control different joints of the robot (see Figure 4-3). This allows the operator to control and monitor the different activities of robots online.

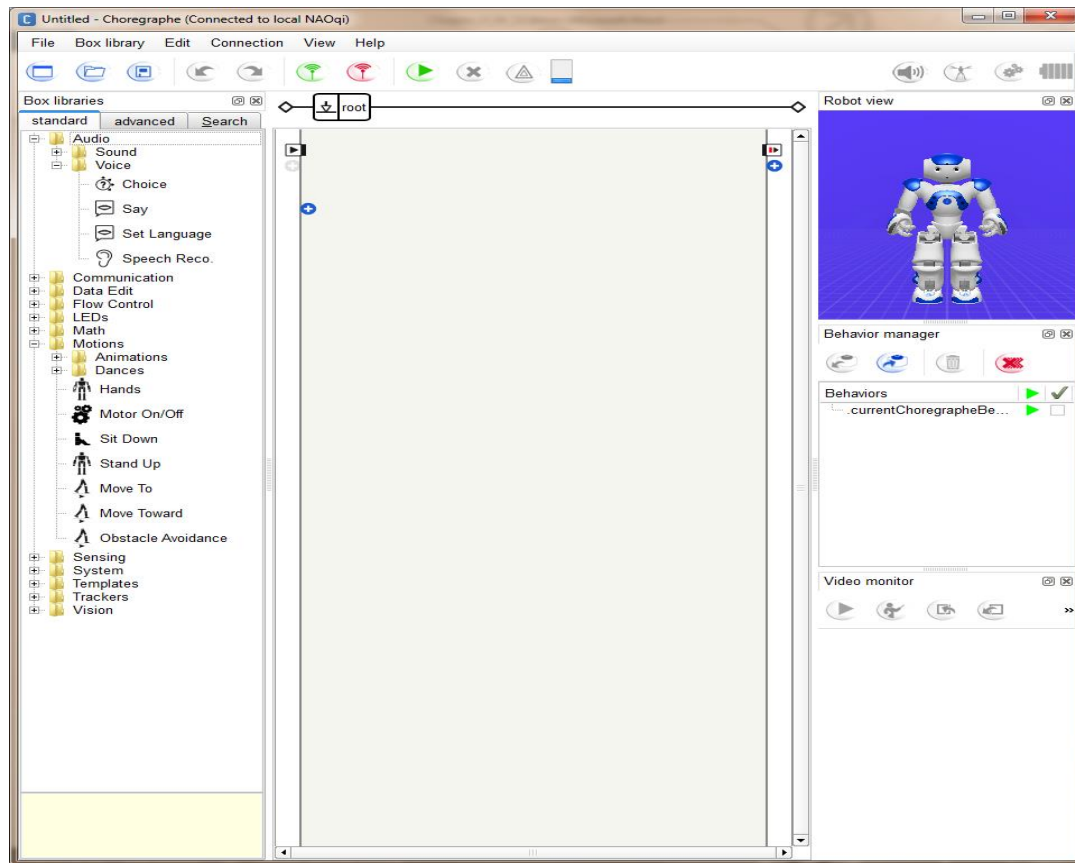


Fig 4-2: Choregraphe user interface

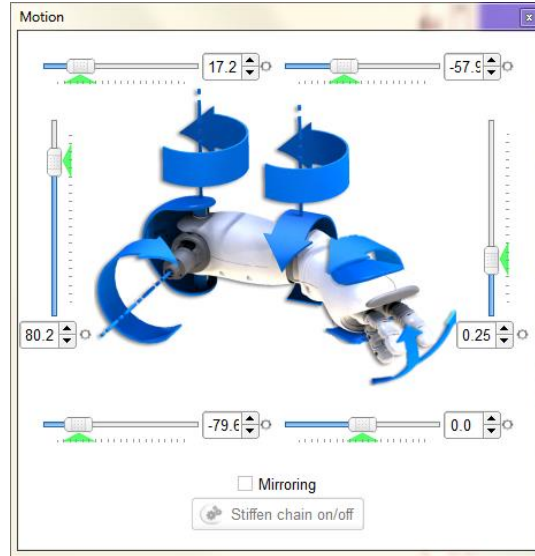


Fig 4-3: Robot joints control pedal

4.1.2 Capturing Sessions and Room Design

All the sessions were held in a $5m \times 5m \times 3m$ room with four surveillance cameras installed at each corner of the room. As shown in Figure 4-4, four additional HD web cameras also were installed in the room. One of them attached on top of the NAO's head. These additional cameras were installed later for post video processing using higher quality images. One of them on the ceiling for detecting the distance between subject and the robot; the other two located in front of the kid and at the right side of the subject for online observation. The room was also decorated with a mount of pictures pinned on the wall with different facial expressions.



Fig 4-4: Videos caregiver can see outside the room

As shown in Figure 4-5 a) & b), these surveillance cameras were connected to a recording system outside of the room connected to a LCD screen, which allowed the parents to watch their children as they participate in the study. Four HD cameras were connected to computers, which allowed researchers observe and record in the meantime. This gave the parents the opportunity to stop the session at any time if they felt the interactions were not appropriate for their child. Sessions were recorded for data analysis

and for future research. The height of the cameras was set to match the eye-gaze level of a normal sitting child about 3 feet high. A height adjustable TV table was located at one corner of the room for the robot to sit on. This allowed the robot to obtain the best quality of interaction with the children at eye level. In the following sections we will explain about the games and applied protocols.



Fig 4-5 a): Panorama view of the experimental room

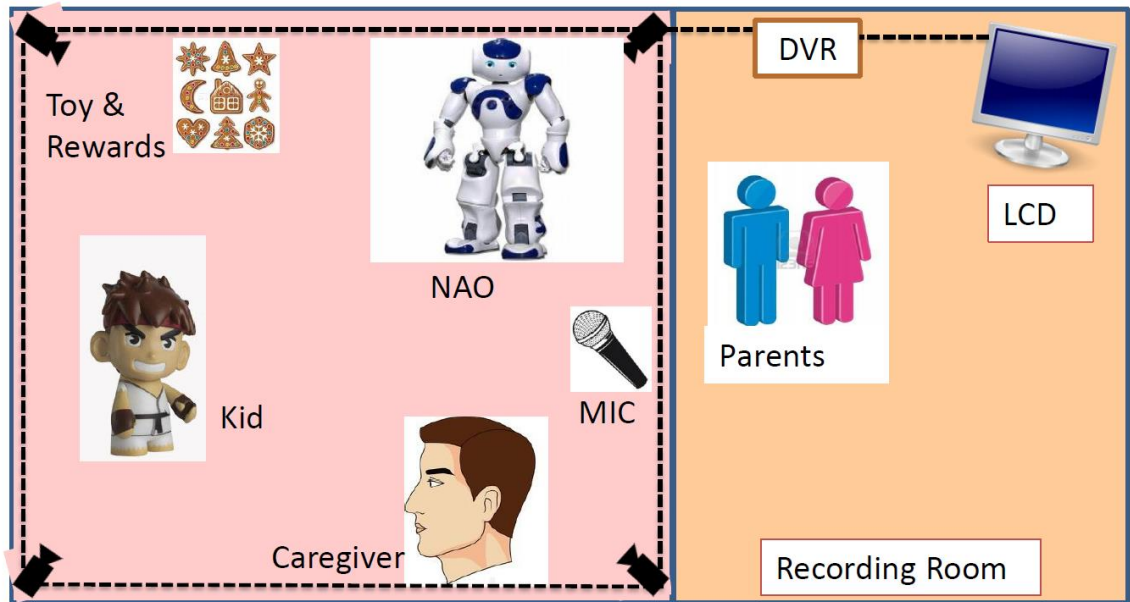


Figure 4-5 b): Schematic robot based therapy session and video capturing setting

4.2 Protocol 1(NAO Spy & Find the Suspects)

Protocol 1 is mostly focused designed to answer the following questions:

- How and in what sense a humanoid social robot like NAO can assist children with ASD to improve their social skills?
- How the gaze directions of TD and children with ASD are different while interaction with NAO in social contexts?
- What model and how accurately we can describe and differentiate the gaze direction of TD and individuals with ASD?

- How different are the gaze patterns of TD vs. children with ASD during the conversational contexts? (i.e. child listening, child speaking scenarios)

In order to answer these questions, two set of games (i.e. NAO Spy and Find the Suspects) have been designed to analyze and model the gaze directional patterns and responding capabilities of the children with ASD.

4.2.1 Participants

All of the participants were recruited from the Denver area, flyers were sent to families associated with the JFK partners [84], posted at autism treatment organizations, and local autism schools. Parents contacted the research assistant via emails or phone calls. The ages of the children ranged from 7 to 17 who had been diagnosed with high functioning ASD. All the parents were asked to provide documentation of ASD diagnosis to participate in the study. IRB approval was obtained and all parents signed informed consent form (children signed assent form). Fourteen children with ASD (13 verbal and 1 non-verbal) completed Protocol 1. Seven TD children have been recruited in this study also as a control group. TD subjects finished both Protocol 1 and Protocol 2.

4.2.2 Protocol 1: Game 1: NAO Spy (NS)

During this game, participants were given opportunities to engage in eye-gaze attention, joint attention, facial expression recognition and imitation, and body gesture

imitation skills during interactions with NAO. This game included five different activities, each designed to provide the opportunity to the participants to engage in various types of social behaviors.

- **Activity 1:** Participants were instructed on how to interact and communicate with NAO. They were asked to listen to NAO's instructions and provide responses to questions, such as "How are you today?" using simple words. During this activity NAO referred to the participant by their name and asked them for a hug in order to make the participants feel comfortable with the system (See Fig 4-6). The aim of this activity is try to build a friendly relationship between kids and robot at the beginning of each session.



Fig 4-6: Kid is hugging NAO

- **Activity 2:** Participants were asked to describe something fun they participated in recent weeks to NAO. In addition NAO also has been telling a short story (1-2 minute) during this segment. To have an interesting session and keep the children engaged, we used various types of stories during each session. These stories were basically funny, exciting, and easy to understand. Gaze directional patterns of children was annotated and analyzed to specially investigate the different eye gaze patterns of both TD and individuals with ASD during listening or speaking dyadic conversational contexts.
- **Activity 3:** During this activity participants were presented with three instructions by NAO. First, they were asked to collect all four small boxes that were in the room and line them up in front of NAO. Then NAO overtly looked at one of the boxes and asked the participant to open the box that NAO was looking at. Finally, NAO described a particular box and asked the participant to pick up the box and open it. The capabilities of TD and children with ASD for understanding the joint attention concept was targeted in this activity.
- **Activity 4:** This activity was a continuation of activity 3 and after opening the box the participant was asked by NAO to describe the facial expression in the box (an image of a facial expression was attached to a small beanbag inside the box) and was also asked to imitate the facial expression. Participants were rewarded with a hand high-five from NAO and a piece of candy for describing

and imitating the facial expression. Next, NAO asked the participant to look at the pictures hanging on the wall and find the picture with the same facial expression as the one they had just found in the box. Participants were again rewarded with a high-five and a LEGO ® mini-figure for completing the activity. Because of the weaknesses of expression/emotion sensitive, this activity would be able to practice the expression recognition and imitation by using the same set of still images with different facial expressions.

- **Activity 5:** During this activity participants were asked to imitate a standing and balancing movement on one leg demonstrated by NAO. This activity was originally designed to analyze how children with ASD can imitate different body postures. In addition it was considered as an entertaining activity for keeping children interested and excited.

4.2.3 Protocol 1: Game 2: Find the Suspects (FTS)

In this game, participants were given the opportunity to engage in some of the same skills as the previous game but in the context of slightly different activities. This game included four different activities each designed to present participants with the opportunity to engage in different types of behaviors.

- **Activity 1 (focusing on eye-gaze patterns):** During this activity, NAO asked the participants a few simple questions (e.g. “what is your name?”, “what is

your favorite color?”, “How old are you?”) and waited for them to answer and then touch one of the pressure sensors on his head to continue. Once the participant answered and touched the sensor, NAO asked the same questions again but asked the participant to look at him in the eye when answering the question.

- **Activity 2 (focused on joint attention and facial expression recognition/imitation):** During this activity, NAO described an expression and asked the child to look for the described picture that was hanged on the wall. After the child found the picture, he/she was asked to imitate that facial expression.
- **Activity 3 (entertainment):** Participants were asked to complete a puzzle of facial expression and were given a candy reward for successful completion. Designing this part was based on researchers’ personal experiences. Also gathering the ideas from the participants. Completing facial expression puzzles could also be a good way to have a better understanding of different components of human face and also different posture of those parts in different facial expressions.
- **Activity 4 (gesture imitation):** Tai Chi (gesture imitation/motor control): During this activity NAO asked participants to take part in a short physical imitation activity called “Tai Chi.” During this activity, NAO demonstrated a

few poses involving his arms and legs and asked the participants to imitate his poses (see Figure 4-7).



Fig 4-7: Kid imitate the “Tai Chi” activity

4.3 Protocol 1: Experimental Results

In the experimental section we report the results of the eye gaze analysis, based on the human-robot therapy sessions for individuals with ASD. We employ the gaze patterns of TD children as the control group to better investigate the gaze differences between TD and children with ASD.

From Chapter 2 we can conclude that during social interaction, eye contact and gaze direction regulation is one of the significant elements to send important information to others [47]. In this context, there are two types of eye gaze features that can be analyzed: 1) Gaze fixation; 2) Gaze shifting.

Gaze fixation can provide an important nonverbal cue about how the listener and speaker are involved in the conversation. For instance frequent eye contact may display the concentration or high level of interests in the topic of the conversation. Besides, some observations [49, 50, and 51] illustrated that speaker might look up and down while thinking and eye gaze would come back to listener while his talking. On the other hand, listener would spend most of the time looking at the speaker no matter what the speaker is thinking or talking [50]. Both participants in the conversation would be able to understand whether it is an attractive topic and whether they can understand each other via eye contact. On the contrary, if one participant eye gaze shifts a lot it might be a sign of less interest in the topic or has been distracted by other facts [50].

In our studies, in order to investigate the gaze pattern of participants we manually coded their gaze direction using '0' vs '1' coding (i.e. gaze averted (0), gaze at NAO (1)). For the Protocol 1 in our experiment, Activity 1 and 2 in 'NAO Spy' and Activity 1 in 'Find the Suspect' have been used for measuring gaze fixation and shifting patterns. More than 453,000 frames total (about 15,120 seconds, $M = 120s$ in each session) were manually

coded offline using the Continuous Measurement System (CMS) [60]. The gaze labels were the used to model the gaze patterns of both ASD and TD children.

The experimental section for Protocol 1 can be classified into two categories:

- Investigating gaze direction of ASD and TD children while they were interacting with the robot. The *Gaze Fixation* is a percentile index which measures the participant's gaze duration as s/he was looking at the robot. The *Gaze Shifting* on the other hand, track the rate of gaze direction shifting by counting the number gaze direction switching (i.e. looking at-looking away and vice versa) w.r.t. NAO.
- Modeling the gaze direction using Markov modeling (i.e. Hidden and Variable-order Markov models) to mathematically represent the differences of gaze patterns of TD and individuals with ASD.

The rest of this chapter will talk about it in more details.

4.3.1 Eye-Gaze Fixation and Shifting Frequency Analysis

Gaze Fixation and shifting are typically used to convey the level of interest, and attention of the listener/speaker throughout the dyadic conversation. In this part, we aim to investigate how TD and individuals with ASD utilize gaze direction while interacting with NAO. For instance typically we used gaze direction as we are thinking, or talking; and while we want to show being interested in it. However for individuals with ASD, we are

not able to read this pattern while interacting with NAO. The following Figures and Tables would describe it statistically.

Table 1 shows the gaze fixation percentage of individuals with ASD for all sessions in the NS game (i.e. counts the number of the frames with label 1 divided by total number of frames multiply by 100).

$$\text{Gaze Fixation (\%)} = \frac{\text{Number of frames (labeled as 1)}}{\text{Total number of frames in video}} \times 100 . \quad (1)$$

From Table 1 it is easy to observe that 10 kids out of 15 shows the increasing eye gaze fixation time after three game sessions, which represented the eye contact between those kids and robot have been improved. It can also be seen from those Mini Figure in the last column in the table. Through the overall Mean Percentage row, we are able to conclude that eye gaze fixation has increased 7% and 15% than first session and second session. We can also find out this tendency from Figure 4-8 shown below.

Duration Table	Number of Session			
Subject#	<i>Session 1</i>	<i>Session 2</i>	<i>Session 3</i>	<i>Mini Figure</i>
SN002	0.69	0.66	0.62	
SN003	0.43	0.32	0.78	
SN004	0.7	0.5	0.53	
SN005	0.56	0.58	0.63	
SN006	0.52	0.41	0.56	
SN008	0.52	0.39	0.56	
SN010	0.52	0.43	0.61	
SN011	0.6	0.7	0.78	
SN012	0.43	0.75	0.94	
SN013	0.54	0.42	0.54	
SN014	0.64	0.65	0.54	
SN015	0.33	0.15	NA	
SN016	0.71	0.68	0.75	
SN017	0.71	0.47	0.62	
SN018	0.61	0.23	0.64	
Mean Percentage	0.57	0.49	0.64	

TABLE 4-1: Time Duration Percentage in Game NS for ASD Group

As it shows in Figure 4-8, there is a valley point between first session and third session although it still shows the improvement overall. The reason behind this can be various. The most possible and reasonable explanation we can come up with is for the very first time of visiting the humanoid robot, most of the kids have shown a high concentration during the session, they wanted to satisfy their curiosity of the robot for the first in their life. As they notice the next couple times, the game sessions are almost same, they might paid less attention, that cause the valley point in second session.

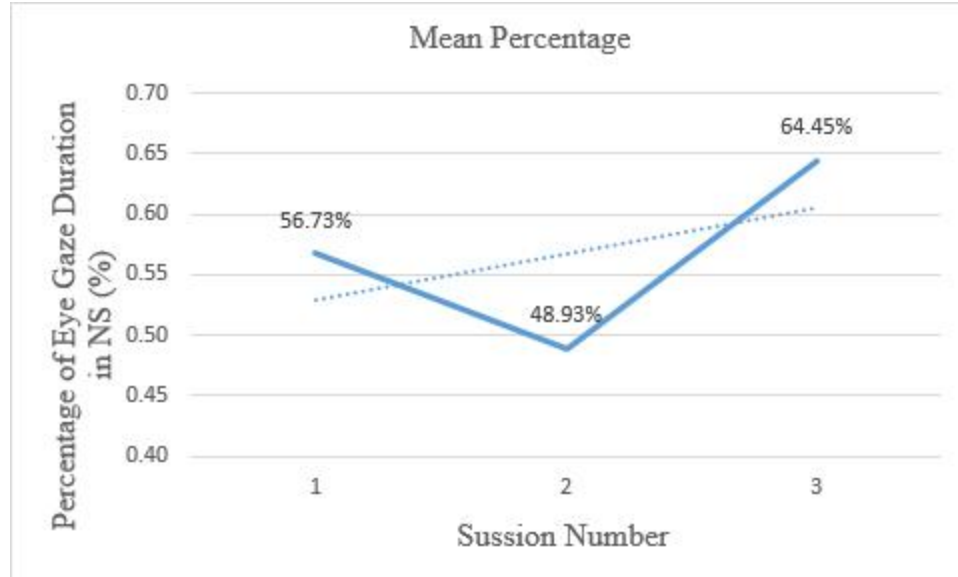


Fig 4-8: Overall mean percentage of eye gaze duration for all children with ASD after three sessions, dot line is the tendency of eye gaze duration.

Gaze Fixation: Eye gaze shifting frequency has been calculated in Table 4-2.

$$\text{Shifting Frequency (\%)} = \frac{\text{Number of } 1 \rightarrow 0 + \text{Number of } 0 \rightarrow 1}{\text{Entire FrameTotal number of frames in videos}} \times 100. \quad (2)$$

Table 4-2 shows the gaze shifting of participants for every 3 session. These data show decreases in the number of gaze shifts across the three sessions. 11 kids out of 15 shows the decreasing eye gaze shifting frequency after three sessions. From Mini Figure column it is quite obvious that 9 kids out of 11 ones who shows the significant change in

third session compared with previous sessions. Besides, 7 kids shows both improvement in increasing eye gaze fixation and in decreasing eye gaze shifting.

Figure 4-9 illustrates the decreases in gaze shifts across the three sessions. From the tending line, it is easy to find out that the eye gaze shifting frequency is decreasing from session to session. As shown in Figure 2 the third session shifting frequency only have half of the amount of the shifting frequency in first session.

















Frequency Table	Number of Session			
Subject#	<i>Session 1</i>	<i>Session 2</i>	<i>Session 3</i>	<i>Mini Figure</i>
SN002	0.0147	0.0112	0.0086	
SN003	0.0047	0.0031	0.0055	
SN004	0.0075	0.0081	0.0067	
SN005	0.0109	0.0076	0.0097	
SN006	0.0099	0.007	0.0069	
SN008	0.0072	0.0072	0.0069	
SN010	0.0152	0.0132	0.0092	
SN011	0.0082	0.0073	0.0052	
SN012	0.0033	0.0076	0.002	
SN013	0.0077	0.0048	0.0043	
SN014	0.0081	0.0063	0.0044	
SN015	0.0308	0.0192	NA	
SN016	0.0249	0.0243	0.0296	
SN017	0.0242	0.0166	0.03	
SN018	0.0232	0.009	0.0309	
Mean Frequency	0.013367	0.010167	0.006309	

TABLE 4-2: Eye Gaze Shifting Frequency in Game NS for ASD Group

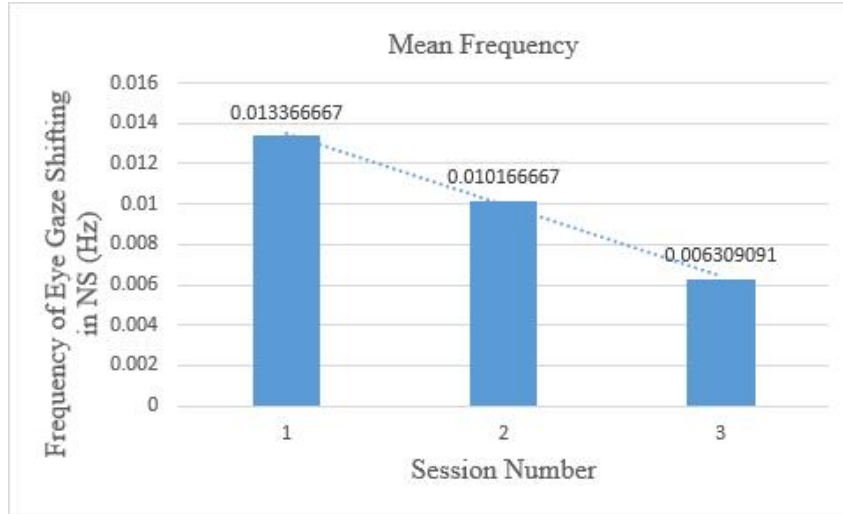


Fig 4-9: Overall mean frequency of eye gaze shifting for all children with ASD after three sessions, dot line is the tendency of eye gaze shifting.

By using the same measurement as NS into FTS, 8 kids out of 13 shows the increasing duration time of eye gaze fixation, and 7 out of 13 shows the decreasing eye gaze shifting frequency in Table 4-3 and Table 4-4 as shown below.

Duration Table	Number of Session			
Subject #	<i>Session 1</i>	<i>Session 2</i>	<i>Session 3</i>	<i>Mini Figure</i>
SN002	0.57	0.75	0.81	
SN003	0.43	0.97	0.92	
SN005	0.46	0.57	NA	
SN006	0.61	0.76	0.76	
SN010	0.79	0.84	0.71	
SN011	0.82	0.99	0.8	
SN012	0.62	0.98	0.99	
SN013	0.34	0.52	0.61	
SN014	0.08	0.54	0.5	
SN015	0.58	0.41	NA	
SN016	0.69	NA	NA	
SN017	0.54	0.68	0.49	
SN018	0.51	0.37	0.54	
Mean Percentage	0.541538	0.733	0.7625	

TABLE 4-3: Time Duration Percentage in Game FTS for ASD Group

Frequency Table	Number of Session			
Subject #	<i>Session 1</i>	<i>Session 2</i>	<i>Session 3</i>	<i>Mini Figure</i>
SN002	0.0124	0.0079	0.0069	
SN003	0.0051	0.0023	0.004	
SN005	0.0116	0.0108	Dropped	
SN006	0.0091	0.0064	0.0046	
SN010	0.0102	0.0116	0.0082	
SN011	0.0048	0.0051	0.0048	
SN012	0.0034	0.0096	0.0025	
SN013	0.0041	0.0053	0.0058	
SN014	0.0027	0.0059	0.0031	
SN015	0.0291	0.007	NA	
SN016	0.0047	NA	NA	
SN017	0.0134	0.0162	0.0206	
SN018	0.0108	0.0162	0.0149	
Mean Frequency	0.009338	0.00719	0.004988	

TABLE 4-4: Eye Gaze Shifting Frequency in Game FTS for ASD Group

Figure 4-10 illustrates the positive change of the eye-gaze duration percentage across the three sessions for all the participants. Comparing the first session and third session, these data also showed increases in eye-gaze duration. Compare with first game the total percentage of duration time has increased.

Figure 4-11 shows the decrease of the eye-gaze shifting frequency for all three sessions. Comparing this game with NS game, eye-gaze shifting frequency amount are decreased overall. Results from both games suggested that game based interaction session has the advantages of improve social behaviors such as eye contact. The result also tells us that most of the kids becoming more focusing on topics with robot, and less distracted from other environment affects. As we know, better eye contact represents a better conversation, after two games; we assume that kids with ASD will build better communication skill in the future. However, social context is not only eye contact but also other factors like reading facial expressions, joint attention, and fluent speech. The results of our experiments for the first protocol have been published in [86-89]. This is the purpose of designing protocol 2 which can help us dig even deeper of this study.

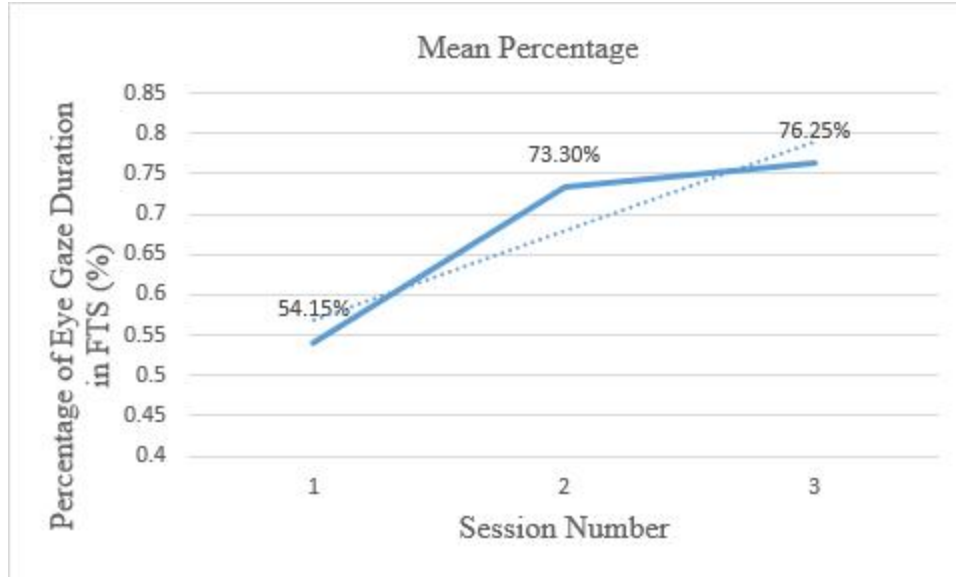


Fig 4-10: Overall mean percentage of eye gaze duration for all children with ASD after three sessions, dash-line is the tendency of eye gaze duration. The total percentage in each session is greater than first game.

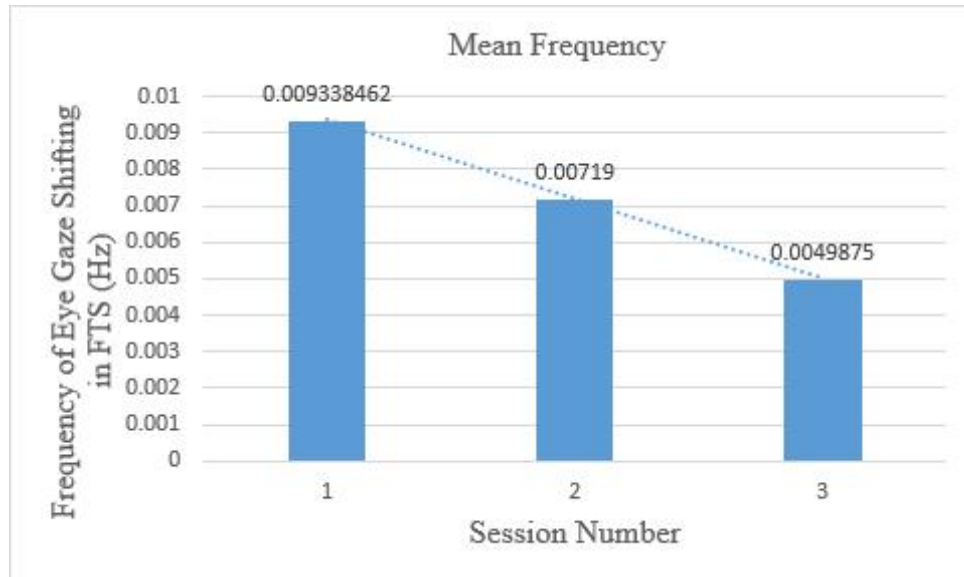


Fig 4- 11: Overall mean frequency of eye gaze shifting for all children with ASD after three sessions, dot line is the tendency of eye gaze shifting. Each session's frequency is less than 0.0015 from the first game.

4.3.2 Eye-Gaze Pattern Modeling

4.3.2.1 Hidden Markov Model (HMM)

Hidden Markov Model a powerful stochastic state model to represent and classify time-sequential data. The learning ability of the HMM has inspired several researchers to apply it for different computer vision and machine learning applications such as speech recognition [61] and Human activity recognition [62].

The discrete HMM is a statistical Markov model, with set of unobservable (hidden) states that at each time t , emits one output from a group of observable symbols. An HMM can be represented by N states, $Q = \{q_1, q_2, \dots, q_N\}$, and state transition probability $A = [a_{ij}] \in R^{N \times N}$ can be formulated as:

$$a_{ij} = Pr(s_{t+1} = q_j | s_t = q_i), \quad 1 \leq i, j \leq N, \quad (3)$$

$$\text{s.t.} \quad a_{ij} \geq 0, \sum_{j=1}^N a_{ij} = 1$$

Note that in Equation 3, it has been assumed the Markov chain is homogeneous so that the transition probability matrix A does not depend on time. The original state of the system at $t = 0$ is q_0 and the initial state probability is defined by. Moreover, the HMM has a set of M observable symbols $V = \{v_1, v_2, \dots, v_M\}$ that is emitted from one of the hidden states. Emission probability (i.e. output probability) is probability of producing an output symbol v_k being in the state q_j (See Equation 4).

$$b_j(k) = \Pr(v_k | s_t = q_j), \quad 1 \leq j \leq N, 1 \leq k \leq M, 1 \leq t \leq T \quad (4)$$

Each HMM $\lambda = (A, B, \pi)$, is characterized by state transition probability matrix (A), emission probability (B), and initial state probability (π). Hidden Markov modeling is capable of analyzing sequence of data with length T ($t = 1, 2, 3, \dots, T$). to use HMM for classification applications, two phases need to be accomplished[]:

Learning: Given the structure of HMM model and set of output sequences $O = \{O_1, O_2, \dots, O_T\}$, estimate the HMM parameters $\lambda = (A, B, \pi)$.

Evaluation: Given HMM parameters, specify the probability of an observed output sequence $O = \{O_1, O_2, \dots, O_T\}$.

In order to employ HMM for classifying a sequence of eye gaze patterns, first a finite set of symbols needs to be defined ($O_i | i = 1, 2, \dots, n$). As explained before, the gaze direction feature has been annotated in every frame of the captured videos and was coded

as either gaze at (1) or gaze averted (0). In our experiments we selected n consecutive binary-coded frames to assigned $2n$ symbols. In other words to come up with 16 unique symbols (i.e. $\{‘0000’, \dots, ‘1111’\}$ or $\{O_1, \dots, O_{16}\}$) we acquired the gaze information of four ($n = 4$) consecutive frames. Figure 4-12 demonstrates a sequence of gaze direction labels and the corresponding symbols.

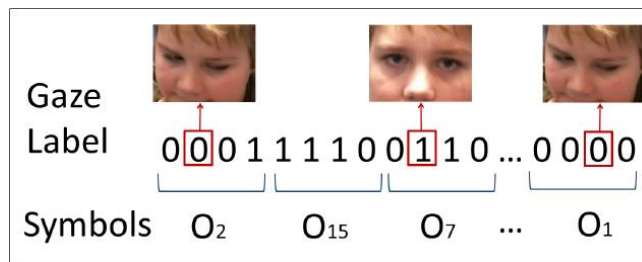


Fig 4-12: Gaze labels and the corresponding gaze symbols (n = 4).

To model and analyze the eye gaze patterns in ASD and TD categories, we selected a sequence of five symbols $I = \{I_1; \dots; I_5\}$ as $\{I_i \in O_j | 1 \leq i \leq 5, 1 \leq j \leq 16\}$. In the experiments we used sequence of observed symbols for both training and classifying eye gaze sequential data into one of $C = 2$ (TD vs ASD) classes. In order to learn HMM parameters of a category, a set of data that belongs to that class (TD or ASD class), was employed to optimize the HMM parameters (i.e. $\lambda_i(A_i, B_i, \lambda_i)$) (Learning Phase). Furthermore, to recognize the category of the observed sequence of symbols, I , the $\{\Pr(\lambda_i|I) | i = 1, 2\}$ was calculated and the observed sequence was assigned to a class which had the highest likelihood probability (Evaluation Phase).

4.3.2.2 Experimental Results: HMM

As explained previously, we utilized NAO to interact with children in a series of conversational games. To specifically analyze the eye gaze of participants in social contexts, we extracted the video segments which corresponded to “*Child listening*” and “*Child Speaking*”. Thereafter, we specified the sequence of eye gaze symbols for training two HMMs and utilizing the learned models to categorize a given test sequence into one of TD or ASD classes.

In this study we have analyzed the eye gaze data of 21 subjects. Our evaluation was based on the Leave-One- Subject-Out (LOSO) cross validation technique, in which the hidden Markov models were trained for both classes using 20 participants. We then tested against the excluded subject and repeated the same approach for all 21 subjects. In addition we reported F1-score which combine both precision (fraction of retrieved instances that are relevant) and recall (fraction of relevant instances that are retrieved) factors in a single measure ($F1 = 2 \frac{\text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$). Our experiment shows that HMM can reliably discriminate between the gaze patterns of TD and ASD groups, in the child speaking segments (Accuracy 79% and F1-score 0.88).

4.3.2.3 Variable-order Markov Model (VMM)

One common approach for analyzing and classifying a sequence of discrete data is to employ the first order Markov model (a memoryless state machine). Another potential alternative is to focus on general-purpose prediction algorithms, utilizing Variable-order Markov Models over finite alphabet Σ [63].

Let us assume Σ is a finite set of alphabets. Given a training sequence $q^n = q_1 q_2 \dots q_n$, where $q_i \in \Sigma$, the goal of VMM is to learn a model P , which can provide a probability assignment for any sequence of symbols. Mathematically speaking, for any context $s \in \Sigma$, and a symbol $\sigma \in \Sigma$, the model generates a conditional probability distribution for $\hat{P}(\sigma|s)$.

In VMM the prediction stage utilizes average log-loss $L(\hat{P}, x_1^T)$ of $\hat{P}(\cdot|\cdot)$ with respect to the test sequence $x^T = x_1 x_2 \dots x_T$.

$$L(\hat{P}, x_1^T) = -\frac{1}{T} \sum_{i=1}^T \log_2 \hat{P}(x_i | x_1, \dots, x_{i-1}) \quad (5)$$

There are several different VMM methods that can be applied for classifying sequential data. In a study conducted in 2004 [63], six prominent VMM algorithms have been compared. Out of these six approaches, prediction by partial match (PPM) algorithm, gave the best accuracy and reliability. In our experiments we also employed PPM algorithm that is introduced below.

Prediction by partial match, is a finite-context statistical modeling technique that can be interpreted as blending several fixed-order context models of order k [64]. In other words PPM can be considered as a combination of fixed-order context models with different values of k , ranging from zero to a pre-determined maximum (D). For each model, PPM keeps track of the length- k sequence of all characters that have been observed so far in the training sequence.

In addition PPM handles the zero frequency problem using escape mechanism [63]. In the escape mechanism, the goal is to determine the probability of unseen sub-sequence of symbols after the context s has been seen in the training sequence. In other words for each context of length $k \leq D$ the probability of $\Pr(\widehat{P}_k(Esc|s))$ is allocated for all symbols that have not appeared after the context s . For more details see [63].

4.3.2.4 Experimental Results: VMM

In order to model the eye gaze pattern using the VMM, we define the alphabets by looking at four consecutive eye gaze labels ('0000' to '1111'). Therefore 16 alphabets have been defined and represented by $\Sigma = \{A, B, \dots, P\}$.

Considering the proposed PPM algorithm, it may perceive that PPM's performance should always improve when the maximum context length (D) is increased. Although, increasing maximum context length specifies more predictions, it also causes the longer

contexts have greater chance of not giving rise to any prediction at all. This results the escape mechanism to be used more frequently. Figure 4-13 shows the accuracy of the gaze direction recognition for different value of D. The figure demonstrates that in this study the proposed VMM eye gaze modeling with $D = 1$, outperforms other orders of VMM ($D = \{2, \dots, 6\}$) which validates the above discussion.

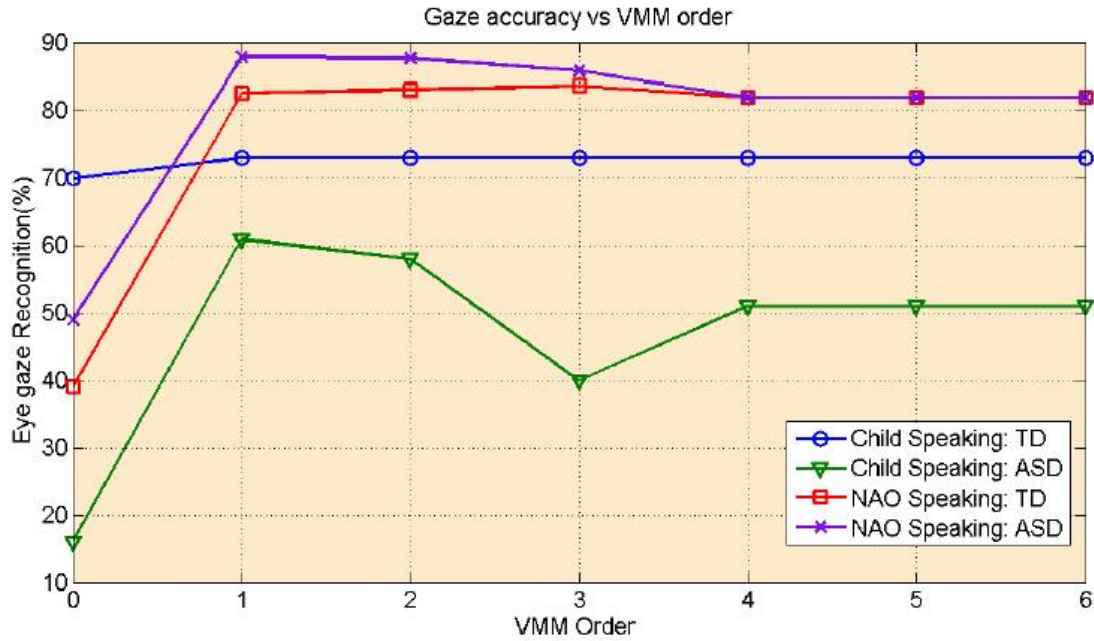


Fig 4-13: VMM maximum context length and eye gaze recognition.

LOSO cross validation technique has been applied in the experiments to train VMMs for the ASD and TD groups. Classification of a test sequence was conducted through a log-likelihood approach, which specifies the chosen test sequence's probability of occurrence, for a given model ($\sum, D = 1$). This yields a likelihood value for every model

(TD and ASD) and the test sequence will assign to a model which has the higher value. This procedure has been repeated for every sequence. The experimental results show that VMM is capable of modeling the gaze pattern of both TD and ASD groups with a high reliability (F1-Score ≥ 0.75). The Results of these experiments have been reported and discussed in [88,89]

4.3.2.5 Discussion

Markov-based modeling is a powerful approach for analyzing a sequence of data. In our study we aim to investigate and analyze the dynamics of gaze patterns of both groups of children (TD and ASD). We used HMM and VMM algorithm to analyze the gaze responses of the participants in two social communicational contexts (i.e. child speaking and child listening).

We first analyzed and modeled the gaze direction of children using HMM. Therefore two separate hidden Markov models that can discriminate between the gaze responses of TD child and peers with autism were learned then for a give sequence of eye gaze labels we aim to recognize the class (TD vs ASD). As shown in Table 4-5, HMM can classify the gaze responses of TD children (in the listening context) with accuracy of 65%. Moreover, HMM can recognize the gaze patterns of children with ASD while they are speaking with high accuracy (82%). This results demonstrate that there are some

differences between the temporal pattern of ASD and TD groups in the two contexts (speaking vs listening). Therefore in the next phase we employed VMM method to characterize the gaze patterns of children reliably.

	Context	TD	ASD
Accuracy (%)	Child Listening	65.73%	41.53%
	Child Speaking	31.58%	82.03%

TABLE 4-5: Classification rates of HMM algorithm for the TD and ASD groups (Child Speaking and Listening contexts)

As explained in section 4.3.2.3, for a given sequence of data, VMM algorithm may go through different orders (length of sequence of data), and automatically find the one that has the most similar pattern with the training set. In order to evaluate how the gaze patterns of children with Autism and TD group are different, we compare various orders of VMM (D=0 to D=5). The results demonstrate that VMM with order one (D=1), can represent the gaze patterns of TD and ASD with the best accuracy (See Table 4-6 and 4-7).

Context: Child Speaking		D=0	D=1	D=2	D=3	D=4	D=5
Accuracy (%)	TD	78.95	73.68	73.68	73.68	73.68	73.68
	ASD	16.81	61.74	58.84	40.87	51.59	51.88

TABLE 4-6: Classification rates of VMM modeling for the “Child Speaking” context

Context: Child Listening		D=0	D=1	D=2	D=3	D=4	D=5
Accuracy (%)	TD	39.89	82.58	83.15	83.71	82.02	82.02
	ASD	49.64	87.81	87.66	85.78	82.80	82.95

TABLE 4-7: Classification rates of VMM modeling for the “Child Listening” context

In addition the results shown in Figure 4-14 illustrates that, increasing the order of VMM, will improve the reliability of modeling the gaze pattern of TD and ASD groups for the listening context. However changing the VMM order within the range of 0-5 does not have significant effects on the gaze model accuracy. Besides that the results shows that VMM with order one, recognizes the gaze patters of TD and ASD well, but it cannot model the gaze pattern of ASD group for the speaking context.

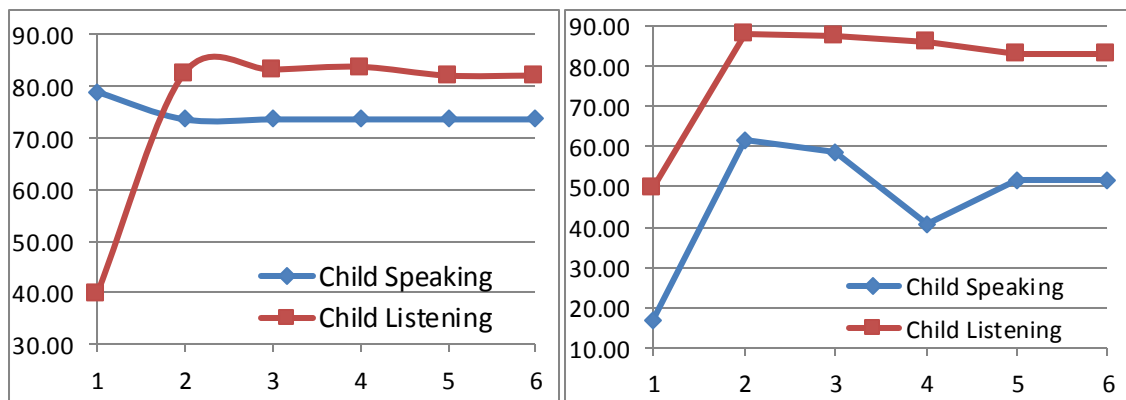


Fig. 4-14 (Left) TD Group and (Right) ASD Group Gaze Classification Rate based on VMM algorithm

Considering these results, we conclude that the eye gaze patterns of both TD and ASD group have memory. We also have seen that, for the context of “Child Listening” VMM can model the difference between the gaze patterns of TD and ASD group well therefore it came up with high accuracy (over 80%). However for the context of “Child Speaking”, there are lots of misclassification for both groups and VMM failed to classify the gaze responses of ASD group reliably.

4.4 Protocol 2 (Intervention Sessions)

In Protocol 1, we focused on the analysis of the eye-gaze pattern of children during interaction with NAO. Based on the results it is reasonable to say that game based sessions are able to affect eye-gaze behaviors of children in a social context. However, since children on the spectrum may have wide range of deficits in social environments, we designed a robot based therapy session that can specifically focus on different social skills independently. This encouraged us to design a new set of games and intervention sessions, which are mainly focused on verbal and non-verbal communication skills.

The objective of this protocol is to find out a quantitative analysis solution for this question: ‘How good social skills of children with ASD can be improved using a socially assistive robot in therapy intervention sessions?’ To answer this question, we designed our

protocol as described below. One visit session has been divided into five small sub-sessions with different aimed tasks injected in each one of them. For each participant, three baseline sessions are executed at the beginning of the project. Data are collected the same way as Protocol 1, and based on the result in baseline sessions, project examiner are made the decision of what task should go for intervention sessions. Each intervention session would also repeated for three times, one intervention session at a time, for example, if subject #1 is having intervention for sub-session #1, then other sub-sessions are keep in the baseline settings. After three sessions of intervention for one sub-session, based on the results, examiner decide whether keep doing the same intervention or move on to another sub-session which has low rates from the baseline data. At the end of each sub-session, NAO has a high-five with subjects to cheer them up as a reward. Few minutes break has given between each sub-session. During the break, participants may allow to go out of the experiment room and hang out with their families. Candies are given at the end of the final session.

4.4.1 Intervention Sub-sessions

In Protocol 2 practicing few specific behaviors such as basic question understanding, joint attention, emotional facial expressions recognition, were assigned into different sub-sessions. After certain amount of visiting sessions, multiple social skills were improved as predicted.

In the first sub-session, NAO asked several questions including basic personal information question such as name, age, family members etc. We collected correct answers to these questions from conversation with parents. Also some entertainment questions or activities are also pre-programmed in the session. To different subjects, a certain interaction was designed in it. Those specific interaction would not count in the measurement in analyzing the results.

Two tasks have contained in the second sub-session which are joint attention and facial expression recognition (with given options). Five lids with different facial expressions attached at back, lying on floor at both sides of the table. NAO would point one side at a time randomly and asked the child to bring one lid and show it to the robot (See Figure 4-15(A)). Then a given optional question was asked from NAO: “What is the facial expression at the back of the lid? Is it sad, angry, happy or neutral?” Children should answer the question for 5 rounds. Order of the expressions are totally random because the back side of those lids are face down toward floor, children cannot see the expressions after they pick them up.

Eye-gaze attention has targeted in the third sub-session. Children should be able to follow NAO’s eye directions and head positions. There are three lids located at three marks on the table, two of them are at the very edge of both sides of the table, and the other one at the center of the front edge of the table. That helped the children to distinguish those

directions from NAO's head movement easily without confusion. Children were supposed to follow the eye-gaze directions and pick one lid up then move on to the next position back and forth for couple of times (See Figure 4-16(A)). At the first five turns NAO would move the head in a normal speed, then the examiner boosts up the speed a little just for entertaining kid and having a fun time. That speed up turns is not included for future analyses.

Sub-session four is the combination of sub-sessions two and plus a facial expression imitation module after the facial expression recognition question. However, different from previous sub-session, this time the expression options would not be provided in the question. Children were supposed to recognize expressions and speak them out themselves. The recognition rate and imitation rate are measured in post analysis (see Figure 4-15(B)).



Fig 4-15: A) Kid is showing the lid; B) Kid is imitating the happy expression

The purpose of the last session is to teach/practice kids pointing skill. Three lids were put on the table as session #3, children are supposed to point each lid with certain color introduced by NAO. For example, NAO says: “Can you point to the yellow lid?” This behavior was repeated 5 times at the beginning, then examiner just to entertain the child and have a fun time for few more pointing behavior in a faster speed. Those speed up turns are not included for future analyses (see Figure 4-16(B)).

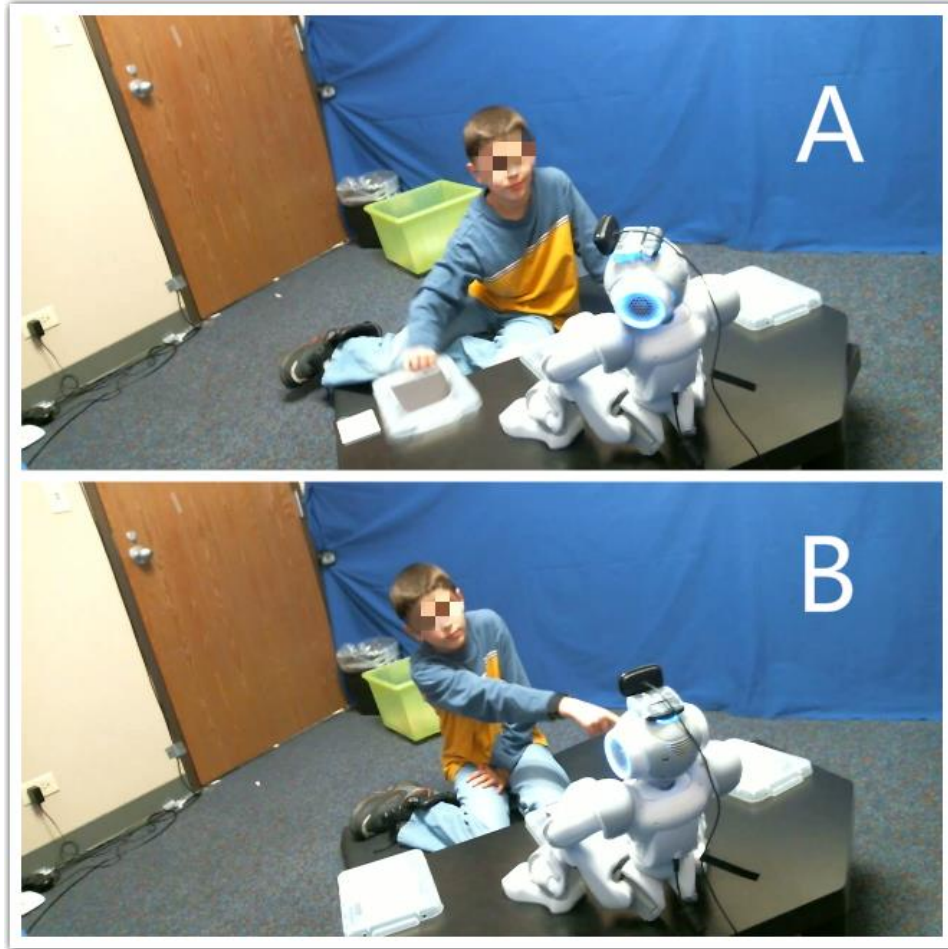


Fig 4-16: A) Kid is following NAO's eye gaze and picking up one lid; B) Kid is pointing a specific lid which NAO is describing.

4.4.2 Protocol 2: Results

In Protocol 2, five sub-sessions which contains seven types of tasks have been defined and coded. These tasks are (T1, T2.1, T2.2, T3, T4.1, T4.2, and T5) that are described below:

Sub-session 1:

- **T1:** Answering five questions: questions are related to name, color, number of siblings, age, and time go to bed.

Sub-session 2:

- **T2.1:** Following NOA's Pointing: NAO points to a box and the child supposed to bring back the lid that NAO is pointing to (for five times).
- **T2.2:** Recognizing Facial Expressions: participants recognize facial expressions on the other side of the lid, as NAO provides option list of those facial expressions (e.g. sad, angry, happy or neutral).

Sub-session 3:

- **T3:** Following Eye Gaze of NAO: Joint attention toward a lid, NAO looked 5 times in different directions randomly during the session.

Sub-session 4:

- **T4.1:** Recognizing Facial Expressions: participants recognize facial expressions on the other side of the lid (similar to T2.2), without providing list of options for the

facial expressions (from session to session, 4 basic expressions have kept but on different faces).

- **T4.2:** Imitating Facial Expressions: NAO asked subjects to imitate that expression that is shown on the picture shown on the lid.

Sub-session 5:

- **T5:** Pointing Lids: NAO description lid (based on a color), participants supposed to point to that lid (randomly describe five different lids).

As shown in Figures 4-17 to 4-23, as we expected different individuals with ASD have different social responses for various social situations. As it is illustrated in following Figures, the baseline characteristic of children would vary sharply between different children. As a general comparison, for sub-session 2 (T2.2.) the baseline of some of the subjects (e.g. SN019, SN020, SN023, and SN025) which is related to facial expression recognition task, is around 40% to 60%. The results statistically show that this group of children can just recognize very few basic facial expressions. For SN020, SN021, SN023, and SN025 they were having a hard time with task T4.2. This demonstrated they have some issues for facial muscle controlling to imitate different prototypic facial expression. One of the subjects in our study (SN024) has lot of difficulty for responding to different designed games, and his scores for different games were low

for almost all of the social behaviors. In order to see how every ASD child responded to different games in Protocol 2, please see Figure 4-17 to Figure 4-23.

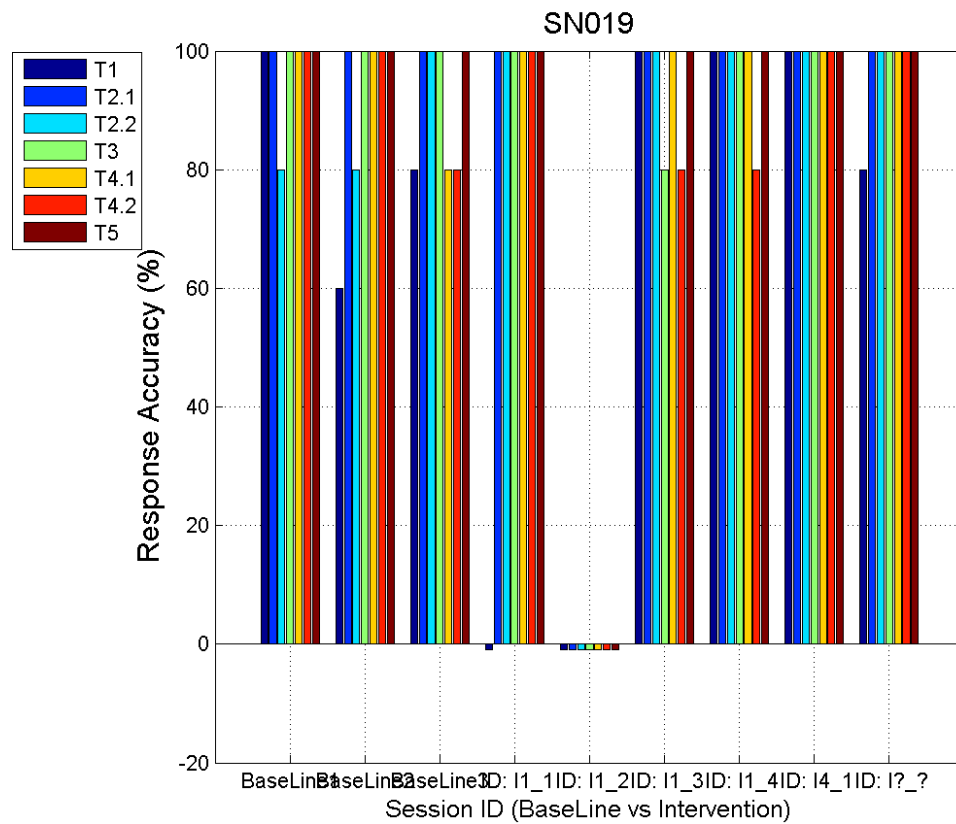


Fig 4-17: Subject 19's behavior during baseline and intervention sessions, intervention sessions shows that T2.2 reached the peak

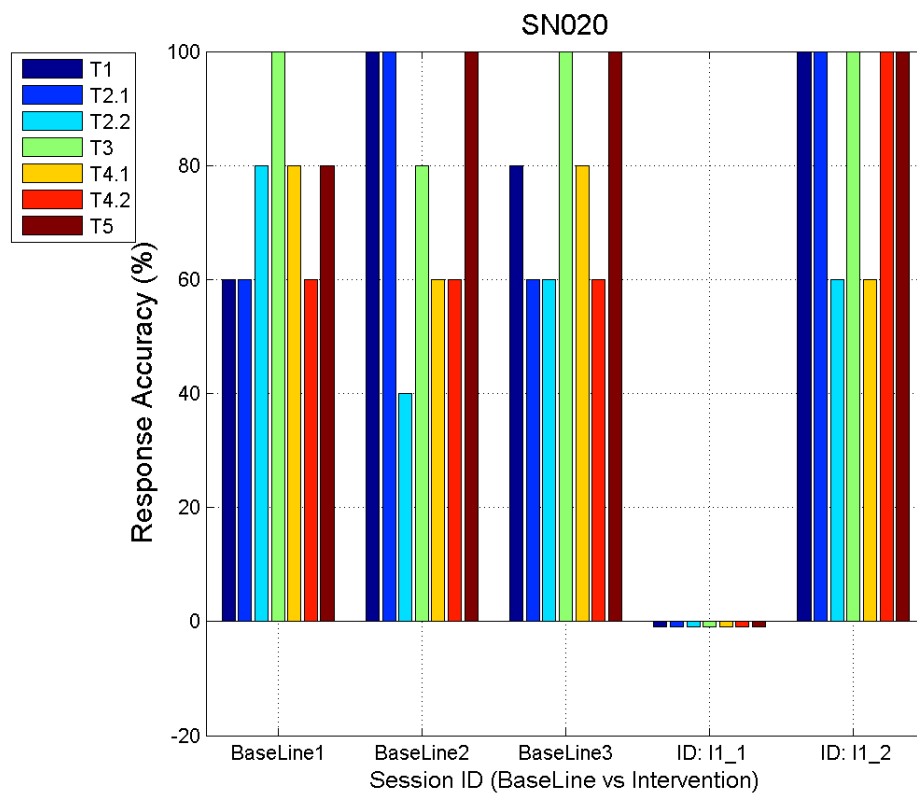


Fig 4-18: Subject 20's behavior during baseline and intervention sessions

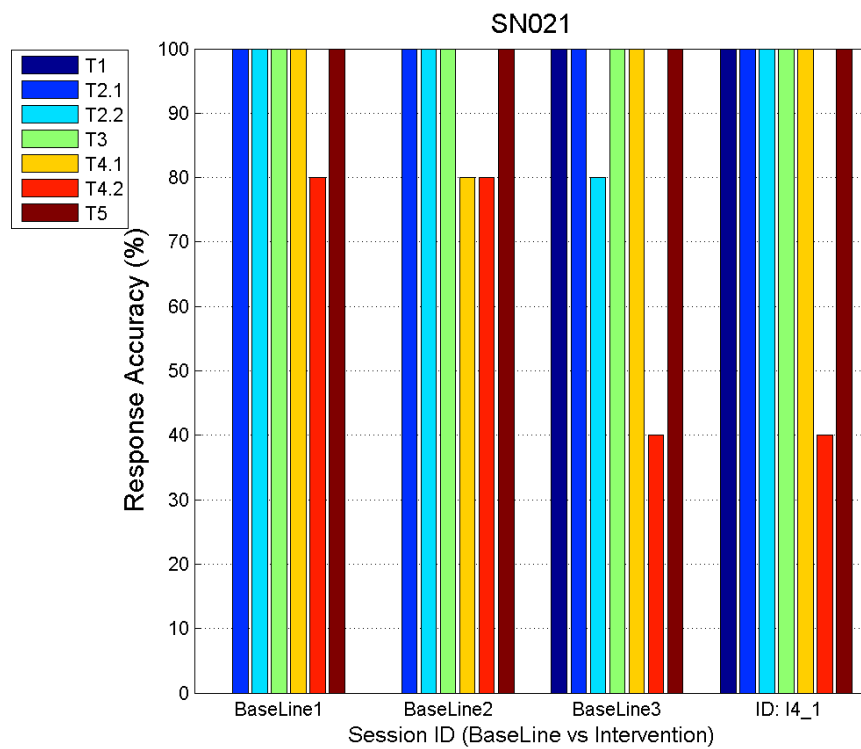


Fig 4-19: Subject 21's behavior during baseline and intervention sessions, intervention sessions shows that T2.2 reached the peak

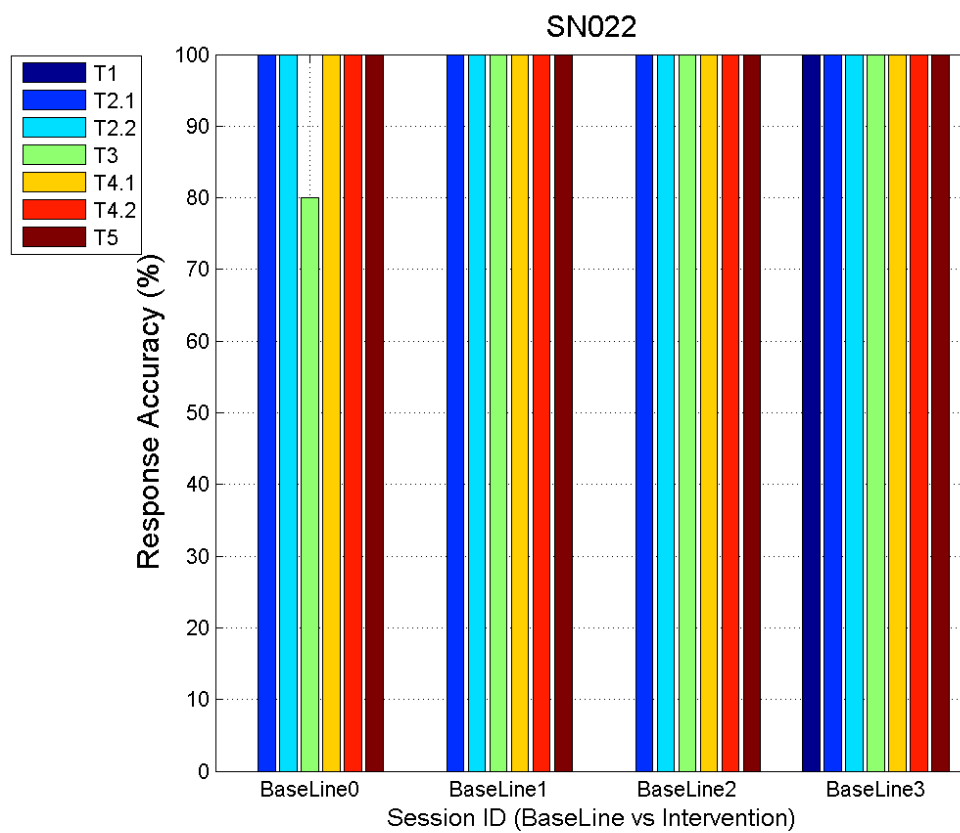


Fig 4-20: Subject 22's behavior during baseline and intervention sessions

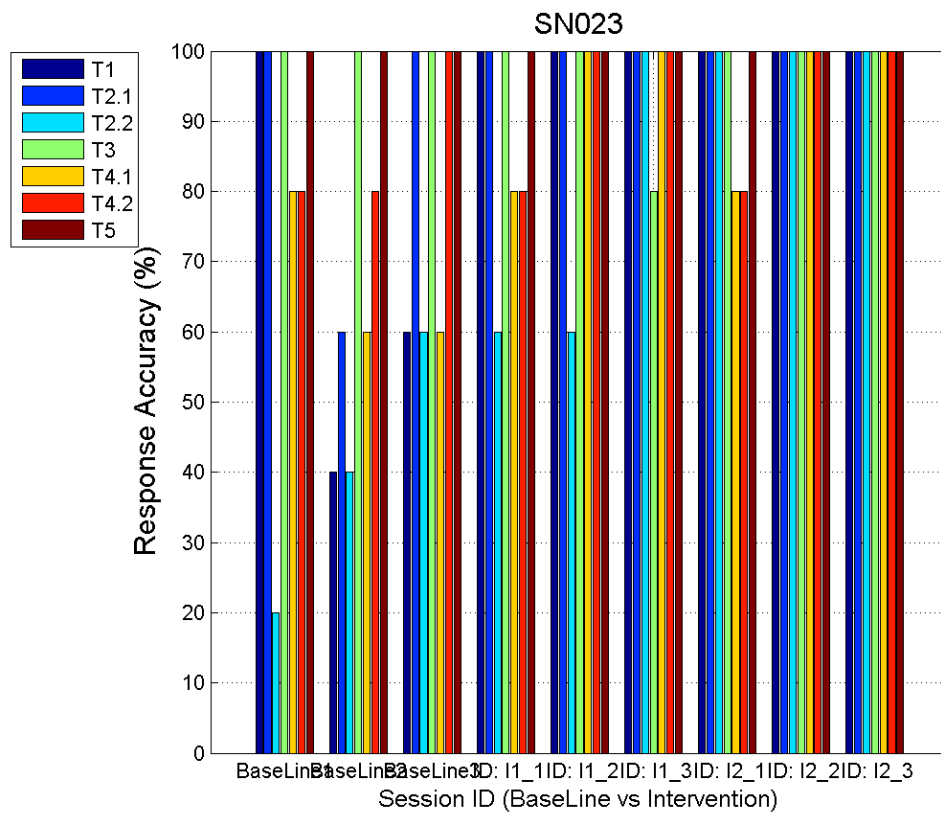


Fig 4-21: Subject 23's behavior during baseline and intervention sessions, intervention sessions shows that T2.2 reached the peak

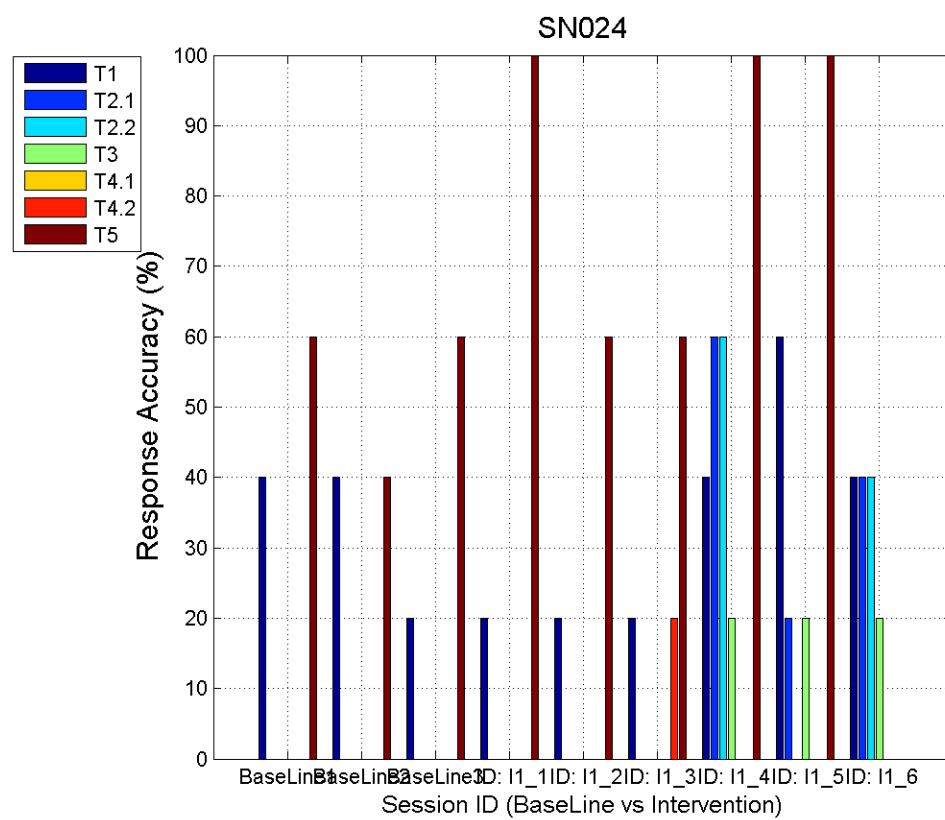


Fig 4-22: Subject 24's behavior during baseline and intervention sessions

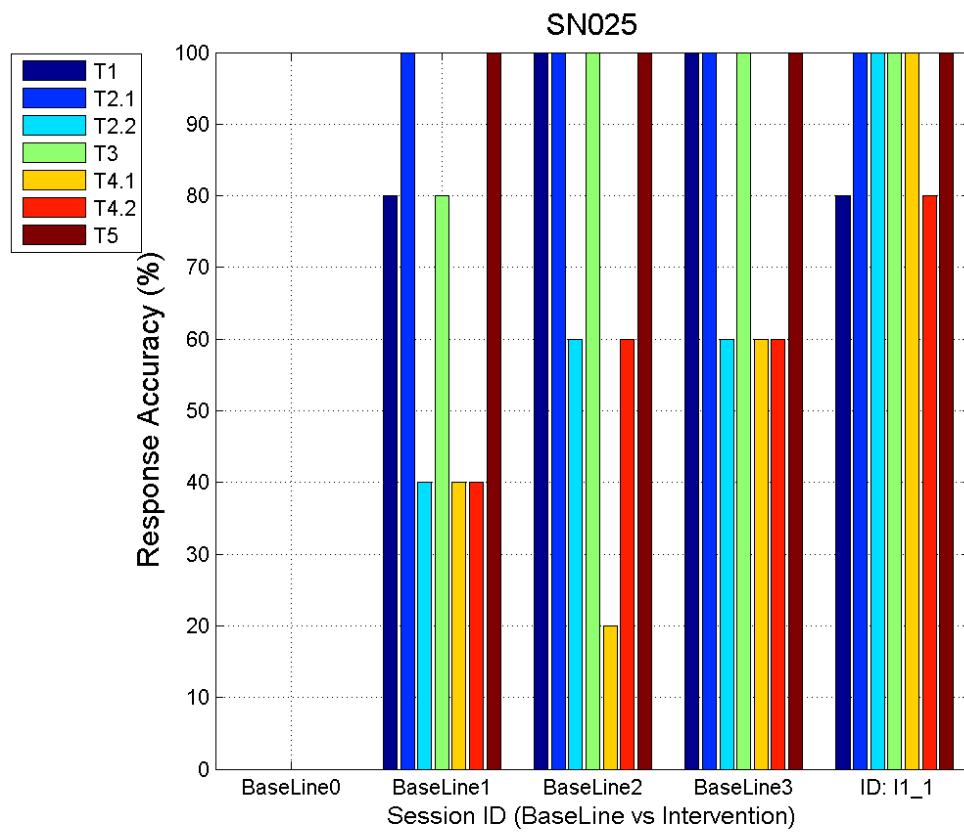


Fig 4-23: Subject 25's behavior during baseline and intervention sessions, intervention sessions shows that T2.2 reached the peak

Chapter 5: Conclusion and Future Research Direction

The purpose of this study was to use humanoid robot into autism therapy sessions, and have a fluent human-robot interaction with children with ASD. Different deficits of autism are considered in this thesis such as basic communication skills, joint attention, eye contact, and facial expression recognition and imitation. During interaction sessions these social behaviors were practiced several times for children with ASD and some promising results are achieved.

The spotlight of this thesis was to focus on eye contact between a robot and children and analyze different eye gaze patterns of children using Markov-based computer algorithms such as HMM and VMM. More than just eye gaze, we also designed interaction therapy sessions including practicing different social behaviors deficits for kids with ASD and these designs show the plausibility of using humanoid robots into autism therapy intervention. One humanoid robot, NAO, was employed in this study. Based on the robot's abilities (i.e. test-to-speech, voice recognition, etc.), two protocols were developed. The results of Protocol 1 reveal that participants with ASD can learn to interact with the humanoid robot and engage in the social-communicative behavior (i.e., making and maintaining eye contact). Results also show that the children with ASD improved the levels of eye contact across sessions with the humanoid robot suggesting that participants were

engaged in the robot-interaction (in 71% of the sessions the fixation time during the communication part is over 50%). Across the three sessions of protocol 1, participants showed improvements in their gaze attention towards the robot (67% of the participants have increased eye contact fixation time) suggesting they paid attention to and looked at the robot in a similar way (more focus on look into other's eyes and less shift eyes around) that one would expect they would look at a human clinician. This is particularly important with regards to the '*Find the Suspect*' (FTS) game where the robot asked participants to look at NAO's eyes while they were answering the questions.

The gaze duration data from the FTS game showed the most robust increases in gaze duration suggesting that participants may have been following the robot's directions and improving their gaze durations as a result. This is the most promising and exciting finding because it shows that the robot can be useful in producing robust changes in clinically significant behaviors for the autism population. There is no available study that compare the gaze responses of such group of children on the autism spectrum.

Also in Protocol 1, a new approach has been presented to model and classify the eye gaze behaviors of TD and children with ASD while socially interacting with a humanoid robot (NAO). As Markov-based models are powerful techniques to learn and classify the dynamics of sequential data, in our study we utilized and compared Hidden and Variable-order Markov Models. Our experimental results demonstrate that both HMM and VMM are capable of representing the differences between eye gaze directions of TD

and individuals with autism. In particular, a first order Hidden Markov model recognized eye gaze patterns of children in the "child speaking" session with an accuracy of 79% (F1-score 0.88). This verifies that the gaze patterns of children is memoryless as they are speaking. Besides, VMM (i.e. first order $D=1$) can discriminate between the TD and children in ASD group's eye gaze patterns with an accuracy of 87% (F1-score 0.92) while children are listening to NAO. The VMM results confirm that as children are listening, their gaze patterns can be represented more accurately by a model with memory. The results validate different characteristics of eye gaze patterns of ASD and TD children in two distinct social contexts (i.e. child speaking and child listening).

In Protocol 2, we added more social behaviors into the interventions (i.e. personal information understanding, pointing response, facial expressions recognition and imitation). Different intervention sessions would be given to improve the weakness of specific social behaviors to specific kids. Most of the participants passed the first task (answer questions) in a quick way (one set of intervention session), only one kid was not able to finish this task after two more sets of interventions. Almost every participant could pass the joint attention task, but for facial expressions recognition task, the neutral face was recognized as sad face for couple times. Imitating facial expressions was a big task for some of the kids during the sessions. We observed that the children with ASD suffered a hard time in imitating different facial expressions. Some of them might recognize the expression but not be able to show it on their own face in a proper way. By given feedback

in the intervention sessions, they did try to move their facial muscles to make specific expressions and to learn those facial expressions.

These preliminary findings support the use of humanoid robots as possible therapeutic agents for individuals with ASD. These results show that participants were engaged with the robot and directed their attention to the robot during a long period of the sessions. Most children enjoyed engaging with the robot (from their parents words, the kids keep asking them when will they play with the robot again, and based on the observation, almost every kid showed a good interaction and behavior during each session). Data taken from an exit survey completed by the parents of every individuals with ASD showed that most children demonstrated improvements with eye contact and joint attention after completing the study.

Overall, robot-based therapy sessions for improving the social behaviors and especially the gaze behavioral responses of children with ASD is a new research topic. Our ongoing project aims to study and model the eye contact and gaze response of children in more details. We encourage other interested researchers to investigate the efficiency of robot to teach behaviors to individuals with autism and hopefully they can use the learned abilities outside of the experimental and research settings.

One interesting area that can be studied in the future is to use robots to jointly interact with care-givers to help individuals with autism. This study will allow a direct comparison of the interaction to human and humanoid interactive partners. These results

will serve as an important basis to significantly advance the emerging field of robot-assisted therapy.

For future research direction, we suggest to analyze the collected facial expression videos to find out how accurately y children with ASD can imitate basic facial expressions. Also audio analysis would be an interesting topic to discover. Based on our observation, we noticed that children with ASD sometimes use only one tone when they speak. Also some of them might have an unexpected high volume peak during the conversation.

New games can be designed as well. We learnt that our designed games seem too easy to some of our partisans. Different games with different difficulty levels will be very useful in the future. The reason behind this is to keep children with ASD more challenged during the sessions so they would not board after a few intervention sessions.

For intervention sessions, it is also good to add more activities in between. We suggest to start with human-human interaction sessions where one examiner do all the interaction with participants. Then follow up with robot-human interaction sessions and at the end of all sessions switch to human-human interaction sessions. By using the same analysis method, it is easy to find out how much improvement the children would gain by interacting with a robot and whether it affects the human-human interaction session. Then by comparing the human-based interactions with the last human sessions, one would easily tell the differences.

Online real time facial expression recognition can be implemented on NAO during the intervention sessions. With some efforts, NAO can be programmed to recognize how correctly children can exhibit facial expressions. Then NAO can provide feedback and instruct them how to show correct facial expressions.

References:

1. Wolff, S., & Chess, S. A behavioral study of schizophrenic children. *Acta Psychiatrica of Child Psychology and Psychiatry*, 1964, 40, 438-466.
2. Pierno, A. C., Mari, M., Lusher, D., & Castiello, U. 200, "Robotic movement elicits visuomotor priming in children with autism." *Neuropsychologia*, 46, 448-454.
3. Tang, K., Diehl, J.J., Villano, M., Wier, K., Thomas, B., Shea, N., et al. 2011, "Enhancing empirically-supported treatments for autism spectrum disorders: A case study using an interactive robot." Presented at the international meeting for autism research.
4. Vallano, M., Crowell, C. R., Wier, K., Tang, K., Thomas, B., Shea, N., et al. 2011, "DOMER: A wizard of oz interface for using interactive robots to scaffold social skills for children with autism spectrum disorders." In *Proceedings of the ACM/IEEE international conference on human-robot interaction*, pp. 279-280. New York, NY: ACM press.
5. D. Feil-Seifer., & M. J. Mataric. *Defining Socially Assistive Robotics*. In *proceedings of the 2005 IEEE 9th International Conference on Rehabilitation Robotics*, 2005, Chicago, IL, USA, 465-468.
6. T. Fong, I. Nourbakhsh, and K. Dautenhahn. A survey of socially interactive robots. *Robotics and Autonomous Systems*, 42(3-4):143-166, 2003.
7. K. Wada, T. Shibata, T. Saito, and K. Tanie. Analysis of factors that bring mental effects to elderly people in robot assisted activity. In *proceedings of the international Conference on Intelligent Robots and Systems*, volume 2, pages 1152-1157, Lausanne, Switzerland, October 2002.
8. N. Edwards and A. Beck. Animal-assisted therapy and nutrition in alzheimer's disease. *Western Journal of Nursing Research*, 24(6):697-712, October 2002.
9. H. Huttenrauch. Fetch-and carry with CERO: Observations from a long-term user study with a service robot. In *proceedings of the International workshop on Robot and Human Interactive Communication*, Pages 158-163, Berlin, Germany, September 2002.
10. K. Dautenhahn, I. Werry, J. Rae, P. Dickerson, P. Stribling, and B. Ogden. Robotic playmates: Analysing interactive competencies of children with autism playing with a mobile robot. In K. Dautenhahn, A. Bond, L. Canamero, and B. Edmonds, editors, *Socially Intelligent Agents: Creating Relationships with Computers and Robots*, pages 117-124. Dordrecht: Kluwer Academic Publishers, 2002.
11. F. Michaud & S. Caron. Roball, the rolling robot. *Autonomous Robots*, 12(2):211,-222, March 2002.
12. F. Michaud and C. Thberge-Turmel. Mobile robotic toys and autism. In K. Dautenhahn, A. Bond, L. Canamero, and B. Edmonds, editors, *Socially Intelligent*

- Agents: Creating Relationships with Computers and Robots, pages 125-132. Kluwer Academic Publishers, 2002.
13. Obringer, L. Ann and S. Jonathan. "Honda ASIMO Robot". How Stuff Works. Retrieved 15 July 2011.
 14. American Psychiatric Association. Diagnostic and statistical manual of mental disorders: DSM-IV. 4 ed. Washington, DC: American Psychiatric Association; 2000. ISBN 0-89042-025-4. OCLC 768475353. Diagnostic criteria for 299.00 Autistic Disorder.
 15. B. Robins, K. Dautenhahn, and J. Dubowski, Does appearance matter in the interaction of children with autism with a humanoid robot? *Interaction Studies*, vol. 7, no. 3, pp. 479-512, 2006.
 16. K. Dautenhahn and I. Werry. Towards interactive robots in autism therapy: background, motivation and challenges. *Pragmatics & Cognition*, 12:1-35 (2004).
 17. P. De Silva, K. Tadano, A. Saito, S. Lambacher, and M. Higashi, Therapeutic-assisted robot for children with autism, *The 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, Oct. 11-15, 2009, St. Louis, USA.
 18. D. Ricks, M. Colton, Trends and Considerations in Robot-Assisted Autism Therapy, *IEEE International Conference on Robotics and Automation*, May 2010.
 19. G. Pioggia, R. Iglizzi, M. L. Sica, M. Ferro, F. Muratori, A. Ahluwalia, and D. De Rossi, Exploring emotional and imitational android-based interactions in autistic spectrum disorders, *Journal of CyberTherapy & Rehabilitation*, vol. 1, issue 1, Spring 2008, pp. 49-61.
 20. H. Kozima, and C. Nakagawa. Interactive robots as facilitators of children's social development. *Mobile Robots towards New Applications*, Edited by Aleksandar Lazinica, pp. 784, Germany, December 2006 (2005).
 21. M. Blow, K. Dautenhahn, A. Appleby, C. L. Nehaniv, and D. Lee, The art of designing robot faces – dimensions for human-robot interaction, *Human Robot Interaction*, Salt Lake City, UT, 2006.
 22. Bender, L. Childhood schizophrenia. Clinical study of one hundred schizophrenic children. *American Journal of Orthopsychiatry*, 1947, 17, 40-56.
 23. Lotter, V. Epidemiology of autistic conditions in young children. I. Prevalence. *Social Psychiatry*, 1966, 20, 111-121.
 24. Rutter, M. Psychotic disorders in early childhood. In A. Coppen & A. Walk (Ed.), *Recent developments in schizophrenia*. British Journal of Psychiatry Special Publication. Ashford, Kent: Headley Bros., 1967.
 25. Vrono, M. Schizophrenia in childhood and adolescence. *International Journal of Mental Health*, 1974, 2, 7-116.
 26. Popper, K. *The logic of scientific discovery* (2nd Ed.). London: Hutchinson, 1959.

27. Popper, K. *Conjectures and regulations: The growth of scientific knowledge* (4th Ed.). London: Routledge and Kegan Paul, 1972.
28. Bartak, L., Rutter, M., & Cox, A. A comparative study of infantile autism and specific developmental receptive language disorder. I. The Children. *British Journal of Psychiatry*, 1975, 126, 127-145.
29. Churchill, D. W., & Bryson, C. Q. Looking and approach behavior of psychotic and normal children as a function of adult attention and preoccupation. *Comparative Psychiatry*, 1972, 13, 171-177.
30. Hutt, C., & Vaizey, M. J. Differential effects of group density on social behavior. *Nature* (London), 1966, 209, 1371-1372.
31. Sorosky, A. D., Omitz, E. M., Brown, N. B., & Ritvo, E. R. Systematic observations of autistic behavior. *Archives of General Psychiatry*, 1968, 18, 439-449.
32. Kanner, L. Autistic disturbances of affective contact. *Nervous Child*, 1943, 2, 217-250.
33. Rutter, M. *Diagnosis and definition of childhood autism. Autism: A reappraisal of concepts and treatment* New York: Plenum, 1987.
34. Creak, M. Schizophrenia syndrome in childhood: Progress report of a working party. *Cerebral palsy Bulletin*, 1961, 3, 501-504.
35. Bender, L. Childhood schizophrenia. Clinical study of one hundred schizophrenic children. *American Journal of Orthopsychiatry*, 1947, 17, 40-56.
36. Lotter, V. Epidemiology of autistic conditions in young children. I. Prevalence. *Social Psychiatry*, 1966, 20, 111-121.
37. Rutter, M. Psychotic disorders in early childhood. In A. Coppen & A. Walk (Ed.), *Recent developments in schizophrenia*. British Journal of Psychiatry Special Publication. Ashford, Kent: Headley Bros., 1967.
38. Vrono, M. Schizophrenia in childhood and adolescence. *International Journal of Mental Health*, 1974, 2, 7-116.
39. Popper, K. *The logic of scientific discovery* (2nd Ed.). London: Hutchinson, 1959.
40. Popper, K. *Conjectures and regulations: The growth of scientific knowledge* (4th Ed.). London: Routledge and Kegan Paul, 1972.
41. Bartak, L., Rutter, M., & Cox, A. A comparative study of infantile autism and specific developmental receptive language disorder. I. The Children. *British Journal of Psychiatry*, 1975, 126, 127-145.
42. Churchill, D. W., & Bryson, C. Q. Looking and approach behavior of psychotic and normal children as a function of adult attention and preoccupation. *Comparative Psychiatry*, 1972, 13, 171-177.
43. Hutt, C., & Vaizey, M. J. Differential effects of group density on social behavior. *Nature* (London), 1966, 209, 1371-1372.

44. Sorosky, A. D., Omitz, E. M., Brown, N. B., & Ritvo, E. R. Systematic observations of autistic behavior. *Archives of General Psychiatry*, 1968, 18, 439-449.
45. Wolff, S., & Chess, S. A behavioral study of schizophrenic children. *Acta Psychiatrica of Child Psychology and Psychiatry*, 1964, 40, 438-466.
46. Marchant, R., Howlin, P., Yule, W., & Rutter, M. Graded change in the treatment of behavior of autistic children. *Journal of Child Psychology and Psychiatry*, 1974, 15, 211-227.
47. V. Hugot, *Eye Gaze Analysis in Human-Human Interactions*, Master of Science Thesis, Stockholm, Sweden, 2007.
48. G. W. Beattie, A further Investigation of the Cognitive Interference Hypothesis of Gaze Patterns during Conversation. *British Journal of Social and Clinical Psychology*, 21 (p. 243-248).
49. M. Argyle, M. Lalljee, M. Cook, The Effects of Visibility on Interaction in a Dayd, *Human Relations*, 21 (p. 3-17).
50. A. Kendon, Some Functions of Gaze Direction in Social Interaction, *Acta Psychologica*, 26 (p. 22-63).
51. M. Argyle, L. Lefebvre, M. Cook, The Meaning of Five Patterns of Gaze, *European Journal of Social Psychology*, 4 (p. 125-136).
52. D. J. Ricks, M. B. Colton, Trends and Considerations in Robot-Assisted Autism Therapy, *IEEE International Conference on Robotics and Automation Anchorage Convention District*, 2010, May 3-8, Anchorage, Alaska, USA
53. D. J. Feil-Seifer and M. Mataric. Robot-Assisted Therapy for Children with Autism Spectrum Disorders. *Procs. Conf. on Interaction Design for Children: Children with Special Needs*, Chicago, USA (2008).
54. K. Dautenhahn and I. Werry. Toward interactive robots in autism therapy: Background, motivation and challenges. *Pragmatics & Cognition*, 12: 1-35 (2004).
55. B. Scassellati. Personal communication, Nov. 2008.
56. B. Robins, K. Dautenhahn, R. Te Boekhorst, and A. Billard, Robotic assistants in therapy and education of children with autism: can a small humanoid robot help encourage social interaction skills? *Univ Access Inf Soc* (2004) 4: 105-120 July 8, 2005.
57. B. Robins, K. Dautenhahn, and J. Dubowski, "Does appearance matter in the interaction of children with autism with a humanoid robot?" *Interaction studies*, vol.7, no. 3, pp. 479-512, 2006.
58. H. Kozima, C. Nakagawa, and Y. Yasuda. "Interactive robots for communication-care: a case-study in autism therapy." *Procs. IEEE Int. Symp. On Robot and Human Interactive Communication (ROMAN05)*, Nashville, USA, (2005).
59. G. Poggia, R. Igiozzi, M. Ferro, A. Ahluwalia, F. Muratori, and D. D. Rossi, An Android for Enhancing Social Skills and Emotion Recognition in People With

- Autism. IEEE Transactions on Neural Systems and Rehabilitation Engineering, Vol. 13, No. 4, December 2005
60. <http://measurement.psy.miami.edu/cms.phtml> , 2013. Continuous Measurement System (CMS)
 61. Bing Hwang Juang and Laurence Rabiner, Hidden Markov Models for speech recognition. *Technometrics*, 33(3):251-272, 1991.
 62. J. Yamato, J. Ohya, and K. Ishii, Recognizing human action in time-sequential images using hidden markov model. In *Computer Vision and Pattern Recognition, 1992 Proceedings CVPR'92, 1992 IEEE Computer Society Conference on* , pages 379-385. IEEE, 1992
 63. R. Begleiter, R. El-Yaniv, and G. Yona. On prediction using variable order markov models. *J. Artif. Intell. Res.(JAIR)*, 22:385-421, 2004.
 64. J. G. Cleary and W. J. Teahan. Unbounded length contexts for ppm. *The computer Journal*, 40(2 and 3): 67-75, 1997.
 65. C. Lord, M. Rutter S. Goode, J. Heemsbergen, H. Jordan, L. Mawhood and E. Schopler, Autism Diagnostic Observation Schedule: A Standardized Observation of Communicative and Social Behavior, *Journal of Autism and Developmental Disorders*, Vol. 9, No.2, 1989
 66. Carbone, Vincent J., et al. "Teaching eye contact to children with autism: A conceptual analysis and single case study." *Education and Treatment of Children* 36.2 (2013): 139-159.
 67. Mirenda, Patricia L., Anne M. Donnellan, and David E. Yoder. "Gaze behavior: A new look at an old problem." *Journal of Autism and Developmental Disorders* 13.4 (1983): 397-409.
 68. Arnold, Randy J. Semple, Ivan Beale, Claire M. Fletcher-Flinn, Angela. "Eye contact in children's social interactions: What is normal behaviour?." *Journal of Intellectual and Developmental Disability* 25.3 (2000): 207-216.
 69. Podrouzek, Wayne, and David Furrow. "Preschoolers' use of eye contact while speaking: The influence of sex, age, and conversational partner." *Journal of Psycholinguistic Research* 17.2 (1988): 89-98.
 70. Foxx, Ronald M. "Attention training: The use of overcorrection avoidance to increase the eye contact of autistic and retarded children." *Journal of Applied Behavior Analysis* 10.3 (1977): 489-499.
 71. Helgeson, David C., et al. "Eye-contact skill training for adolescents with developmental disabilities and severe behavior problems." *Education & Training in Mental Retardation* (1989).
 72. Greer, D. R., & Ross, D. E. (2007). *Verbal behavior analysis*. New York, NY: Pearson Education.

73. Tiegerman, Ellenmorris, and Louis H. Primavera. "Imitating the autistic child: Facilitating communicative gaze behavior." *Journal of autism and developmental disorders* 14.1 (1984): 27-38.
74. Jones, E. A., & Carr, E. G. (2004). Joint attention in children with autism: Theory and intervention. *Focus on Autism & Other Developmental Disabilities*, 19, 13–26.
75. Whalen, C., & Schreibman, L. (2003). Joint attention training for children with autism using behavior modification procedures. *Journal of Child Psychology and Psychiatry*, 44, 456–468.
76. Plavnick, J. B., & Ferreri, S. J. (2012). Collateral effects of mind training for children with autism. *Research in Autism Spectrum Disorders*, 6, 1366-1376.
77. Ramachandran, Vilayanur S., and Lindsay M. Oberman. "Broken mirrors: A theory of autism." *Scientific American* 295.5 (2006): 62-69
78. Dautenhahn, K. & Werry, I. (2000). Issues of Robot-Human Interaction Dynamics in the Rehabilitation of Children with Autism. *Proc. From Animals To Animats, The Sixth International Conference on the Simulation of Adaptive Behavior (SAB2000)*, Paris, France (Sep, 2000)
79. Iain Werry & Kerstin Dautenhahn & willian Harwin, Investigating a Robot as a Therapy Partner for Children with Autism,
80. S. Shamsuddin, H. Yussor, L. I. Ismail, F. A. Hanapiah, S. Mohamed and H. A. Piah, Initial response of Autism Children in Human-robot interaction therapy with humanoid robot NAO, 2012 IEEE 8th International Colloquium on Signal Processing and its Applications
81. S. Shamsuddin, H. Yussor, L. I. Ismail, F. A. Hanapiah, S. Mohamed, H. A. Hanapiah and Nur Ismarrubie Zahan, Initial response in HRI- a case study on evaluation of child with autism spectrum disorders interacting with a humanoid robot NAO, International Symposium on Robotics and Intelligent Sensors 2012 (IRIS 2012)
82. S. Shamsuddin, H. Yussor, L. I. Ismail, F. A. Hanapiah, S. Mohamed, H. A. Hanapiah and Nur Ismarrubie Zahan, Humanoid Robot NAO interacting with autistic children of moderately impaired intelligence to augment communication skills, International Symposium on Robotics and Intelligent Sensors 2012 (IRIS 2012)
83. S. Shamsuddin, H. Yussor, M. A. Miskam, M. A. c. Hamid, N. A. Malik, H. Hashim, F. A. Hanapiah and L. I. Ismail, Humanoid robot NAO as HRI mediator to teach emotions using game-centered approach for children with autism,

84. <http://www.jfkpartners.org/>
85. H. Feng, M.Mahoor, A.Gutierrez, M. Kustner, J. Zhang, "Using Social Robots at Improving Eye Gaze Attention of Children with Autism Spectrum Disorders", Proceeding of INFAR 2013 May, Donostia, San Sebastian, Basque Country, Spain, 2013.
86. H. Feng, A.Gutierrez, "Using Social Robots to Improve Directed Eye Gaze of Children with Autism Spectrum Disorders", Presented in TARC July 2013 San Marcos, TX, USA, 2013
87. H. Feng, M.Mahoor, A.Gutierrez, M. Kustner, J. Zhang, "Using Social Robots to Improve Directed Eye Gaze of Children with Autism Spectrum", Proceeding of ICHI September 2013 Philadelphia, PA, USA, 2013.
88. S. M. Mavadati, H. Feng, S. Silver, A.Gutierrez, M.Mahoor, "Children-Robot Interaction: Eye Gaze Analysis of Children with Autism during Social Interactions", International Meeting for Autism Research (IMFAR), Atlanta, 2014.
89. S. M. Mavadati, H. Feng, A.Gutierrez, M.Mahoor, "Modeling Eye Gaze of Children with Autism During a Robot-based Therapy Setting", Submitted to Proceeding of IEEE EMBS 2014, Chicago, IL, USA, 2014.

Publications:

- [a] H. Feng, M.Mahoor, A.Gutierrez, M. Kustner, J. Zhang, “Using Social Robots at Improving Eye Gaze Attention of Children with Autism Spectrum Disorders”, *Proceeding of INFAR 2013 May*, Donostia, San Sebastian, Basque Country, Spain, 2013.
- [b] H. Feng, A.Gutierrez, “Using Social Robots to Improve Directed Eye Gaze of Children with Autism Spectrum Disorders”, *Presented in TARC July 2013 San Marcos, TX, USA*, 2013
- [c] H. Feng, M.Mahoor, A.Gutierrez, M. Kustner, J. Zhang, “Using Social Robots to Improve Directed Eye Gaze of Children with Autism Spectrum”, *Proceeding of ICHI September 2013 Philadelphia, PA, USA*, 2013.
- [d] S. M. Mavadati, H. Feng, S. Silver, A.Gutierrez, M.Mahoor, “Children-Robot Interaction: Eye Gaze Analysis of Children with Autism during Social Interactions”, *International Meeting for Autism Research (IMFAR)*, Atlanta, 2014.
- [e] S. M. Mavadati, H. Feng, A.Gutierrez, M.Mahoor, “Modeling Eye Gaze of Children with Autism During a Robot-based Therapy Setting”, *Submitted to Proceeding of IEEE EMBS 2014*, Chicago, IL, USA, 2014.