

Development of Experimental Protocol for Analyzing Arousal, Attention, and Language of Children with Communication Disorders Using Wearable Sensors during Child-Robot Interaction

Asif Mahmud¹, Arshia Khan²

^{1,2}Department of Computer Science University of Minnesota, Duluth, MN, United States
Email address: ¹mahmu037 @ d.umn.edu, ²akhan @ d.umn.edu

Abstract— Communication disorder refers to any disability that affects one's ability to properly understand and apply speech and language to get involved in communication with others effectively. Children suffering from any of these disorders often fall behind in academic achievement, proper social interaction, and smooth daily life. According to prior research, early diagnosis and therapeutic intervention are often effective measures to treat these disorders which eventually lead to academic success and a better social and personal life. The application of humanoid robots through child-robot interaction has become widely popular recently as a part of these therapeutic interventions. In this project, we have designed and developed an experimental protocol as a part of a long-standing goal to analyze the arousal, attention, language, and non-verbal social communication behavior of this group of children using wearable sensors while interacting with a humanoid robot compared to a human therapist. The development of the protocol consists of the research questions and study endpoints identification, interaction module design, study procedure design, data banking process, and policy identification, study population identification and statistical considerations, privacy and confidentiality considerations, and generating consent forms and recruitment materials. We programmed five interaction modules with the humanoid robot NAO and defined data collection protocol for electrodermal activity, heart rate variability, and eye gaze tracking sensors. Electrodermal activity and heart rate variability are measures of one's arousal and stress level while an eye-gaze tracking sensor can be used to track a child's attention. We have also designed a modified version of the protocol so that the study can be performed online through video conferencing in response to Covid-19 pandemic situation. Moreover, potential study extensions, for instance, employing the same experiment to understand the social communication behavior of children with Autism Spectrum Disorder, have been discussed. In the future, this comprehensive experimental setup can readily be used to involve these groups of children to participate in the pilot study. The collected data from the experiments is planned to be analyzed to have a better understanding of the social interactions between humanoid robots and children with communication disorders for therapeutic intervention.

Keywords— Communication Disorders, Autism spectrum disorder, Humanoid Robot, NAO, Child-Robot Interaction, Electrodermal Activity, Heart Rate Variability, Eye Gaze Tracker, Therapeutic Intervention.

I. INTRODUCTION

According to a supplement to the 2012 National Health Interview Survey, published by the Centers for Disease Control

and Prevention: "Nearly 8% of children aged 3–17 years had a communication disorder during the past 12 months" and "Approximately 55% of children aged 3–17 years who had any communication disorder received an intervention service during the past 12 months." The researchers acknowledged that early diagnosis and therapeutic intervention are often effective measures to treat communication disorders which eventually lead to academic success and a better social and personal life. According to the Centers for Disease Control and Prevention, in 2018 one in every 59 children is diagnosed with autism spectrum disorder (ASD) as well. Children with ASD often suffer from social communication disorder. In recent years, humanoid robots are being used as an experimental therapeutic tool to treat these children with communication-related disorders. In this project, we have designed and developed an experimental protocol as a part of our long-standing goal to better understand the social interactions between humanoid robots and children with these disorders for therapeutic intervention. This will help future researchers, as well as healthcare professionals, better understand how the verbal and non-verbal expressive language characteristics of these children are correlated with their physiological responses such as electrodermal activity and heart rate variability. The results from the study can be used to analyze how these physiological responses of these children differ while being engaged with a human partner vs being engaged with a humanoid robot during different interaction modules. The goal is to find the partner who elicits the most communication attempts in this population to increase therapeutic outcomes across a variety of communication-related disorders.

The development of the protocol comprises of the research questions and study endpoints identification, interaction module design, study procedure design, data banking process, and policy identification, study population identification and statistical considerations, privacy and confidentiality considerations, and generating consent forms and recruitment materials. We have developed five interaction modules with the humanoid robot NAO and defined data collection protocol for electro- dermal activity, heart rate variability, and eye gaze tracking sensors. Electrodermal activity and heart rate variability are measures of one's arousal and stress level while an eye-gaze tracking sensor can be used to track a child's

attention. We have also designed a modified version of the protocol so that the study can be performed online through video conferencing in response to Covid-19 pandemic situation when in person study poses health risk. A potential extension of the study could be performing the same experiment on children with Autism Spectrum Disorder. It will help researchers better understand their social communication behavior and how effective humanoid robots such as NAO are as a therapeutic tool for this group of children. Also, the study can be extended to involve typically developed children who will serve as a reference to which we can compare and have a better understanding of the behavior of children having these disorders.

This paper is organized in the following way: in section II, we are going to shed some light into the background of our work. In section III, we will present the methods of the development of the experimental protocol. Then we will discuss the potential study endpoints in section IV. The modified online version of the proposed study will be discussed in section V. Then in section VI, potential study extensions will be discussed. And in the last section, we will discuss the conclusion and highlight future works regarding the study.

II. BACKGROUND

A. Communication Disorder

Communication disorder refers to any disability that affects one's ability to properly understand and apply speech and language in order to get involved in communication with others effectively. A wide variety of disabilities related to comprehending, detecting, expressing and applying language and speech falls under the range of communication disorders. The disorders can range from simple sound substitution to the inability to understand or use one's native language. These deficits can be present in an individual's speech, language, hearing or any combination thereof.

B. Autism Spectrum Disorder

Autism spectrum disorder is a developmental disorder and it is considered a spectrum disorder because its complications and complexities have a wide variety. Every person having this disorder might be completely unique. People might be diagnosed with this disorder at any age. But generally children are affected by this disorder at their early age and gradually the symptoms become clearer. This is the reason it is called a developmental disorder.

Leo Kanner was the first person to come up with the term 'Autism'. In his paper [1], he ran experiment on 11 children, who lacked social skills, loved to stay alone and do repetitive tasks. Basically, he termed this condition as 'Early Infantile Autism'. But as it has already been said that it is a spectrum disorder, the symptoms are not limited to the above mentioned ones only. They might have trouble making eye contact, being far away from peers and engaging with older or younger, switching contexts irrelevantly during a conversation, trouble expressing or understanding emotions, expressions and body languages, trying to avoid listening or looking at people, having difficulty to carry on interactive conversations, rarely

showing responsiveness to different activities and so on. A person having the disorder might have any subset of the above mentioned symptoms or have a few new traits that is completely different from others. So far, there is no direct treatment that can completely cure this disorder. And no one also knows the reason why people get affected by this. People of any age, gender, ethnicity and culture might get diagnosed with this disease. Though cannot be completely curable, different therapies and psychological treatments are available to make the person overcome the effects of the disorder.

C. NAO

NAO is a humanoid robot. It was developed by Aldebaran Robotics, a French robotics company. But later SoftBank Group, a Japanese multinational company acquired the french robotics company. And then they rebranded it as SoftBank Robotics. It is widely used all over the world for different educational and research purposes. This is a programmable robot that can be used in many fields including healthcare, entertainment and so on. NAO's approachable appearance, moderate size, adequate sensors, and advanced programmable computational power have made it a wonderful tool for being used in human-robot interaction by researchers and students. Its size is almost equal to a toddler which is very adorable. Kids find it more like a toy instead of an intimidating robot. Its conversational skills and humanlike attitudes make it even more attainable by kids. These are the primary reasons why Nao is engaged in interacting with children having communication disorders.

D. Electrodermal Activity

According to Critchley et al. (2013) [2], "Electrodermal activity (EDA) reflects the output of integrated attentional and effective and motivational processes within the central nervous system acting on the body". EDA is used as an indicator of one's responsiveness. Skin conductance/resistance varies based on the state of the sweat glands in the skin. And our sweat glands are controlled by the sympathetic nervous system. So any physiological or psychological arousal triggers change in the state of the sweat glands of our skin. And it changes the skin conductance/resistance. So it can be inferred that physiological or psychological response triggers change in the skin conductance/resistance. That's the reason EDA is used as measure of human responsiveness or arousal or stress level.



Fig. 1. NAO Robot



Fig. 2. Empatica E4 EDA & HRV Sensor

E. Heart Rate variability

Heart rate variability (HRV) is a measure of the irregularity between the intervals of two successive heartbeats. For instance, say a person has a heartbeat of 60 beats per minute. It doesn't necessarily mean his heart beats once each second. There might be variability of time between two consecutive beats. For example, the gap between the first two beats maybe 0.5 seconds and the gap between the next two consecutive ones maybe 1.5 seconds. This phenomenon is controlled by the autonomic nervous system of the human body. Whenever a person is relaxed or in baseline condition he will have a high HRV but factors such as stress reduce one's HRV. So HRV can be used as a measure of one's stress level.

F. Eye Tracking Device

Tobii Pro Glasses 2 is a wearable eye tracker. With the glasses on, we can keep track of what a person is looking at in real-time. With the help of Tobii Pro Lab software, a person's gaze data can be observed and analyzed to have a better insight into a person's visual and cognitive attention.



Fig. 3. Tobii Pro Glasses 2

G. Mean length of utterance (MLU)

Mean length of utterance (MLU) is an index of language development in children based on the average length of verbal speech or utterances in their spontaneous speech. It is usually calculated by counting morphemes rather than words.

H. Type token ratio (TTR)

A language analysis method where grammatical parts of speech are measured and classified for vocabulary diversity. The formula for TTR is total number of different words divided by total number of words. A high TTR indicates a high degree of lexical variation while a low TTR indicates the opposite. It is also a measure of language development in young children.

I. Socially Assistive Robot (SAR)

Robots that provide social assistance to human users with the goal of creating close and effective interaction with another human.

J. cHRI

Child-robot interaction (cHRI) is a field of study concerning interactions between children and robots.

K. PHI

Protected Health information (PHI) is defined as any information that could be used to identify an individual.

L. Related Works

In recent years, researcher have been examining a therapy technique: using social assistive robots in the realm of social communication disorders. The goal of socially assistive robotics is to create close and effective interactions with human users for the purpose of giving assistance and achieving measurable progress in rehabilitation and learning (Feil-Seifer & Mataric, 2005 [3]). There have been numerous studies looking at the social aspect of human-robot interaction (HRI) with the use of robots: Keepon, Paro, NAO, KASPAR and Milo. Ones that are humanoid in appearance include KASPAR, Milo and NAO have been used in studies to evaluate the social skills of children, including those with autism spectrum disorder (ASD). Other robots with a non-humanoid appearance have also been used in cHRI studies, including Keepon and Paro (Dumochel & Damiano, 2017 [4]). Prior research has shown children who have social communication disorder, such as autism spectrum disorder (ASD), showed an attentional preference towards a humanoid-robot, such as NAO, during various tasks. Not only did they show increased attention spans during the duration of the session, but also clinical improvements in areas of deficits, such as joint attention and social orientation (Kumazaki et al. 2018 [5]; Shamsuddin et al. 2012 [6]; Warren et al. 2015 [7]; Zheng et al. 2013 [8]).

Research concerning cHRI has ranged in scope and area so far. Kennedy et al. (2017) [9] studied nonverbal immediacy on recall of a story. The goal of their research was to establish "guidelines for social behaviour in robot tutoring scenarios." Measures such as gaze, gestures, facial expression, vocal intonation, proximity and touch informed their research. Druga et al. (2017) [10] did not use a humanoid robot, but rather an artificial agent to study how 3-10 year old children perceived the agents (e.g. friendly or not). These were mostly in the form of smart speakers, such as an Amazon Alexa device and showed the children involved were not afraid of the robotic voice. Another study programmed NAO with lesson modules for children with autism to practice their social interaction skills. The study found that a significant amount of eye contact was exchanged between the child and the robot throughout the modules, giving a good indicator towards the positivity of robot based intervention for children (Miskam et al. 2013 [11]). It was also found that the humanoid robot NAO can possibly help kids with ASD to improve social and relational abilities through activities where the child had to guess the emotional gestures expressed by NAO (Conti et al. 2019 [12]; Miskam et al. 2014 [13]).

The use of a humanoid robot is supported by other studies, as Druga et al. (2017) [10] cite work from other researchers which found that children interact with robots much the same as they interact with people. Studies have also shown that children treat robots as communicators and interact with a robot much as they would with a peer in respect to attention, language use and pragmatics. Therefore, implementing a

socially assistive humanoid-robot as a positive addition, rather than a therapist, to rehabilitation sessions can be used for increasing interest and attention for positive outcomes (Feil-Seifer & Mataric, 2009 [14]). There have been studies that have looked at expressive language use by children with a robot, measuring TTR and MLU (Breazeal et al. 2016 [15]; Kory, J. 2014 [16]; Westlund et al. 2017 [17]; Westlund et al. 2019 [18]). However, these studies used a robot that did not have a humanoid appearance.

In addition to the pragmatic and expressive measures of language, electrodermal activity is another parameter that can be used to measure the interaction in cHRI. EDA to date has been used to measure the hyper and hypo arousal states of human participants in various studies by responding to the internal physiological states of the user through skin conductance. According to Krupa, et al (2016) [19], the electrical properties of the skin exhibit certain variations in reaction to psychological processes, of which can be measured in terms of conductance and resistance. This activity is related to the sympathetic branch of the autonomic nervous system (ANS) and the variations of conductance and resistance in relation to these properties can be defined as Electrodermal activity (EDA) (Krupa, et al. 2016 [19]). With this, heart rate variability (HRV) can also be measured based on sympathetic and parasympathetic influences, providing important information to cardiac activity during situational demands. In combination with the EDA, the two measures play a key role in emotion recognition, of which is a key factor lacking in children who have social communication disorders (Krupa et al. 2016 [19]).

Studies have shown, increased arousal measured with the EDA coincides with prolonged viewing, a concomitant of preferential interaction (Bal et al. 2010) [20]. Literature has also found children with social communication disorders exhibit increased arousal states when appearing calm to the viewer, but exhibit decreased arousal states when viewing a human face (Krupa et al. 2016 [19]; Stagg et al. 2013 [21]). Leite et al. (2013) [22] presented a case study examining the preferences of a child with ASD while interacting with a non-humanoid robot by measures using EDA activity. The researchers found the correlation between motivation, engagement, and attention to be statistically significant in relation to the EDA measures. Therefore, a current gap in literature is determining how well the arousal rates from EDA & HRV correlate with nonverbal social communication behaviors, such as facial expressions, gestures, tone of voice, and proxemics and verbal language characteristics such as MLU, and TTR. This correlation will provide insights into how children with various communication disorders internally feel while interacting with a humanoid robot versus a human therapist, as well as, whether there is an increase in communication attempts, either verbally or nonverbally to enhance therapy outcomes.

III. METHODS

A. Study Design

This is a quantitative experimental design planned to examine the effects of different social communication partners,

such as NAO versus a human therapist, on the expressive language output and physiological measures of a child who has a communication disorder. The child will be asked to participate in three social games including Simon Says, guess the animal game, and reading a story with first a human therapist, and then again with the NAO robot. The presentation of partner interactions is purposely fixed to minimize participation variance across conditions. The human partner will be presented first to allow the participant to become acquainted in the area and with the activities before introducing a novel partner. Despite the human partner having a strict role, they have the ability to discontinue faster if the child becomes distressed. However, the human partner will strictly stick to the script identical to the robot to minimize any variance in partner input. If redirection is needed, a student investigator will step in and provide the necessary assistance. Since the purpose of this study is merely exploratory, statistical significance will be less of an issue and the likelihood that the outcomes of the dependent variables will be due to chance is noted.

The course of the experiment is designed to take place in two parts. It will consist of the participants and legal guardians attending part 1, orientation and part 2, study intervention where the participant interacts with the human therapist, followed by interacting with the humanoid robot, NAO.

- 1) *Part One:* The participants along with their legal guardian will attend an orientation meeting at the designated site with the principal investigator(s) and student investigator(s). This session will consist of:
 - The participant meeting the robot.
 - The legal guardian signing the consent forms.
 - The student investigator(s) and principal investigator(s) explain what to expect while interacting with the robot.
 - The legal guardian choosing the day to come with their child for the experiment. They will have the option to choose to accommodate the participants' schedules.
 - Pre-survey, completed by the legal guardian regarding the child's social communication status, general anxiousness, and sociodemographics.

The timeline of part 1 is described in Table I.

TABLE I. Protocol Part 1 Timeline

Activity	Time (Minutes)	Elapsed Time (Minutes)
Introduction of the student and principal investigators on the team	1-2	1-2
Introduction of the robot	3-4	4-6
Explanation of what to expect for the part two of the study	2-3	6-9
Student investigators will explain the consent form and legal guardians will provide final signature	2-3	8-12
Legal guardians will choose the predetermined time slot and day of when they can perform the experiment	3-5	11-17
Legal guardians will complete brief survey	6-8	17-25
Remaining time will be left for questions	3-5	20-30
Total Time	20-30	20-30

- 2) *Part Two:* On the experiment day, participants will arrive

and wait in the designated site until their time. When their turn arrives, the student investigator(s) will take the participant and their legal guardian to the experiment room to begin the study. A student investigator will be controlling NAO throughout the experiment. Another student investigator will interact with the child as a human therapist. The primary investigator/another student investigator will preside over the session and guide everyone throughout the process.

First, the participant will be seated in a four-legged stationary chair. The chair used will depend on the size of the child. Their feet must be flat on the floor to ensure safety. Then eye gaze tracking glasses, EDA and HRV sensor will be placed on the participant followed by a calming period to calibrate the sensors and gather baseline data.

Participants will then interact with a human therapist in a series of three activities:

- Simon Says (Participant follows verbal directions given by the human therapist and gives verbal directions back to the human therapist)
- Guess the Animal (Participant will guess the name of the animal when presented with an audio recording and slight imitation of animal movements)
- Storytelling (Participant listens to a short story told by the human therapist and asks/answers questions related to the story)

After interacting with a human therapist, once baseline for the sensors is reset, participants will move onto interactions with a humanoid robot, NAO, in the series of the same three activities again. But this time NAO will play the exact same role as the human therapist.

After this, the participant will take part in a brief survey and thus the study will be completed.

Each activity is elaborately described in the section III-C (Interaction Modules).

The timeline of part 2 is described in Table II.

TABLE II. Protocol Part 2 Timeline

Activity	Time (Minutes)	Elapsed Time (Minutes)
Place and Calibrate Tobii Pro Eye Glasses	1-3	1-3
Place EDA and HRV sensor	1-3	2-6
Rest to obtain baseline data	3	5-9
Playing Game: Activity One (Simon Says) with Human Therapist	2-4	7-13
Playing Game: Activity Two (Guess the Animal Game) with Human Therapist	2-4	9-17
Playing Game: Activity Three (Storytelling Activity) with Human Therapist	2-4	11-21
Rest: Reobtain baseline	3	14-24
Introduce NAO into the therapy room	3	17-27
Playing Game: Activity One (Simon Says) with NAO	2-4	19-31
Playing Game: Activity Two (Guess the Animal Game) with NAO	2-4	21-35
Playing Game: Activity Three (Storytelling Activity) with NAO	2-4	23-39
End of activities: Brief survey and answer participants' questions	4	27-43
Total Time	27-43	27-43

B. Data Collection Protocol for the In-Person Study

1) *Eye Gaze Tracking Data:* Eye gaze tracking data will be collected with Tobii Pro Glasses 2 (also referred to as Tobii Glasses). And data acquisition will be done using Tobii Pro software. The child will wear the glasses. The glasses will gather the graphical data of the participant's gaze with no PHI throughout the session and save to a memory chip within a part of the glasses. Once the session has concluded the researchers will plug in the data to a computer and view the eye gaze data. If a child rejects the glasses due to noncompliance, procedures will include taking subjective data, the approximate positioning of eye direction, such as towards the communication partner (either human therapist or NAO) or away from the partner.

In the beginning of the study, the Tobii Glasses will be placed on the participants to calibrate the tracking sensors on the glasses to the child's gaze. This step will take approximately 1-3 minutes and includes:

- Determining fit of the glasses for the child and adjusting accordingly.
- Once the proper fit has been determined, the student investigator will ask the child to look at a dot on a small calibration card. This will be the extent of calibration.
- If the participant rejects or becomes uncooperative with the Tobii glasses, the glasses will be removed immediately and the secondary outcome will focus on a subjective measure of eye contact.

The cooperative participants will wear the glasses throughout the entire interaction session with NAO and the human therapist.

2) *EDA and HRV Data:* Electrodermal activity (EDA) and heart rate variability (HRV) data will be acquired with Empatica E4 wristband, a wearable wristband placed on a single location of wrist of choice by participants. The data will be uploaded to the secure cloud platform of Empatica server (<https://www.empatica.com/>) during the sessions with no PHI. Post experiment, the data will be downloaded and removed from the server immediately.

At the beginning of the study, after placing and calibrating the Tobii Glasses, a student investigator will place the Empatica E4 wristband on the participant's wrist. To obtain a baseline EDA and HRV measurement, the participant will continue sitting in the chair in the therapy room for approximately 3 minutes. This process is done by:

- Skin cleaning using a cleaning wipe on the wrist only.
- Placement of the wristband containing the EDA and HRV sensor.
- The child will be asked to sit still in the chair once the wristband is placed. The child will not sit longer than 3 minutes.

After the participant and the human therapist are done with their series of three activities, the participant will again be asked to sit calmly for 3 minutes to reset the baseline EDA and HRV measurement.

The wristband needs to be worn by the participant throughout the whole interaction session.

3) *Verbal or Non-Verbal Communication Data:* The whole

session will be video-recorded using a VALT (Video, Audio, Learning Tool)-based System, which is a secure, credential controlled video/audio recording and reviewing system installed at the designated site. In case of absence of such a system, an external camera might be used to record the sessions with proper data security and confidentiality caution. The recorded sessions will be reviewed to analyze the verbal language data (MLU, TTR, etc.) and nonverbal social communication behaviors including but not limited to facial expressions, gestures, tone of voice, and proxemics, etc.

There would be no long-term follow-up data.

C. Interaction Modules

The participant will engage in a series of three activities with a human partner. Then the child will engage in the same series of three activities with the robot partner, NAO. The three activities are described below from the perspective of child-robot interaction. During child-adult interaction, NAO's role will be played by a human therapist.

1) *Simon Says Game*: Simon Says (or Simple Simon Says) is a children's game for three or more players. One player takes the role of "Simon" and issues instructions (usually physical actions such as "jump in the air" or "stick out your tongue") to the other players, which should be followed only when prefaced with the phrase "Simon says". Players are eliminated from the game by either following instructions that are not immediately preceded by the phrase, or by failing to follow an instruction which does include the phrase "Simon says". It is the ability to distinguish between genuine and fake commands, rather than physical ability, that usually matters in the game; in most cases, the action just needs to be attempted. We have implemented a simple remake of the actual Simon Says game. Instead of having multiple people it will have only two participants- the participant and NAO. The other variation from the typical game is that there will be turn taking, where the participant and NAO are continuously going to take turns being Simon (giving directions) for five rounds. At the onset of the game, NAO will ask the participant to play the game with him. If the participant accepts to play, then NAO will describe different activities it can perform to the participant. These activities will be-

- Arms Up- Lifting both arms up, straight out in front of the body, parallel to the ground.
- Point to nose
- Move right leg
- Move left leg
- Raise your left arm
- Raise your right arm
- Nod your head

The participant will command NAO to perform any of the above tasks with or without saying Simon says in the beginning. NAO will respond accordingly. After the participant's turn is over, then it will be NAO's turn to be Simon. And like the actual game, NAO will check the participant's attention by asking the participant to perform any

of the following activities with or without saying "Simon Says" at first-

- Arms Up- Lifting both arms up, straight out in front of the body, parallel to the ground.
- Point to nose
- Move right leg
- Move left leg
- Raise your left arm
- Raise your right arm
- Nod your head

NAO will only start its sentence with "Simon Says" in every alternate command. For instance, the first command will start with "Simon Says" but the second one will not. After each command, the participant will either perform the said task and give some approval prompt indicating that they did it or they will tell NAO that they forgot to say "Simon Says". After hearing the prompt from the participant based on its actual command, NAO will either congratulate them or say that it tricked them. Then the participant once again will be Simon and this will continue for another four iterations.

2) *Guess The Animal Game*: Next, the participant and NAO will participate in the second activity, "guess the animal game". The game will work in the following way: In each cycle, Nao will think of an animal randomly from the following 10 animals:

- Cat
- Dog
- Elephant
- Cow
- Horse
- Monkey
- Gorilla
- Sheep
- Mouse
- Lion

Then NAO will play the sound of that animal, slightly imitate some movement of the animal and ask the participant to guess the animal. If the participant guesses correctly, NAO will move to the next cycle and think of another animal from the list. If the participant fails to guess the animal correctly, NAO will give some verbal hints about the animal. If this time the participant guesses correctly, NAO will move to the next cycle as usually. If the participant fails again, then NAO will reveal the name of the animal and move to the next cycle. There will be a total of five cycles of this game.

3) *Storytelling Activity*: The final activity will consist of NAO reading a short story, followed by asking questions to the participant. NAO will read only from the script and will not include any additional comments. NAO will tell the child, "I am going to read you a story please sit down and listen". NAO will then read the passage "Reading Time", taken from the *Easy Stories for Language* book by Patty R. Schuchardt. The story lasts about 24 seconds to read and is as follows:

"Father is reading a story to his children. He is reading about a little baby pig. The baby pig's name is Pilly. Pilly lives on a farm. Bob, Karen, and Sally listen very closely to their father. They are enjoying the story about the pig."

The children start to yawn. Father knows it's time to end the story. He says that Pilly needs to go to sleep. "Just like children, pigs need to sleep", he says. The children hop into bed and go to sleep."

NAO will then ask the participant the following questions from the book:

- "What is Father doing?"
- "Tell me what the story is about."
- "Where does Pilly live?"
- "What is your favorite bedtime story?"

The researchers will not be looking at a right or specific answer from the child in response to the questions. Instead the questions will be reviewed securely accessing the recorded sessions and will be used for measuring participant's responses to the robot regarding mean length utterance (how long the sentences are), eye contact, nonverbal social communication behaviors (including but not limited to, facial expressions, gestures, tone of voice, and proxemics/space between the child and interaction partner), and the physiological differences, in terms of electrodermal activity and heart rate variability, in the social interaction.

Due to the age and nature of the population, a redirection protocol will be implemented. For each of the activities, with both NAO and the human therapist, the participant will have three chances to be redirected back to the activity if he or she becomes distracted. Redirection will consist of the human therapist giving a verbal statement. For example, "It looks like you are distracted, can you guess the name of the animal?". If

the participant reaches the limit, the activity stops and moves on to the next activity.

D. Interaction Module Development with NAO

Nao is a bipedal robot having a close resemblance to a human toddler. In this project, we used the version 6 of NAO which was released in 2018. All the specifications and programming aspects described in this section are based on this version. It has 25 degrees of freedom that enable it to move different parts of its body having a weight of 5.5 kg. It has a dimension of 574 * 275 * 311 mm.

It has the following sensors: ¹

- Eight force sensitive resistors in it's two feet to walk, move and keep balance
- A 3-axis gyrometers with an angular speed of 500/s to locate himself in space
- A 3-axis accelerometer with an acceleration of 2g to locate himself in the space
- Two ultrasonic sensors to detect objects and shapes
- Touch sensors on head, hands and feet to perceive the environment
- Joint position sensors

It has the following interaction devices: ²

- Two loudspeakers in it's ear
- Four microphones on the head

¹http://www.bx.psu.edu/~thanh/naoqi/family/robots/index_robots.html

²http://www.bx.psu.edu/~thanh/naoqi/family/robots/index_robots.html

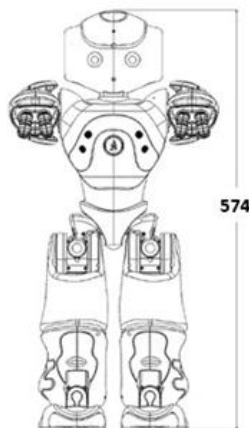


Fig. 4. Dimension of NAO

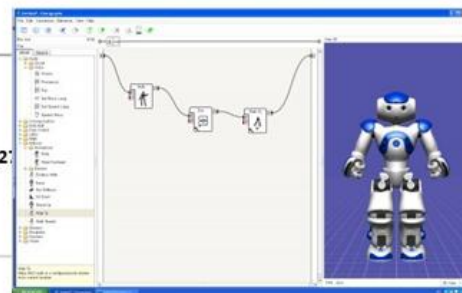
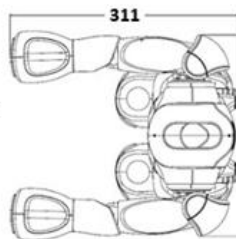


Fig. 5. A sample choregraphe program

- Two video cameras on the forehead
- Two infra-reds in it's eyes
- LEDs in it's eyes, ears, head and feet

The software that runs on the robot and controls it is named Naoqi³. Using the Naoqi framework, NAO can be programmed and all the sensors and interaction devices of NAO can be controlled. The Naoqi framework/API can be used by the major programming languages like C++, Java, Python, .NET, etc. to develop programs with NAO from scratch. NAO can also be programmed with a software named Choregraphe.

³<http://doc.aldebaran.com/1-14/dev/naoqi/index.html#naoqi-framework-overview>

In this project, we used Choregraphe to develop the interaction modules. It is a cross-platform graphical programming interface that allows a programmer to build NAO programs with Python programming language using graphical flow control. In this software, a Python box or simply a box is the unit of programming. Each box is a python module that does a specific task. These tasks are also referred to as behaviors. So, boxes are basically Python implementations of different behaviors. In Choregraphe, these boxes can be dragged and dropped in a graphical interface. Then these behaviors can be stacked and connected with each other to develop a complex behavior. Choregraphe gives a

structural and graphical way of representing the flow control of a complex program. There are plenty of default behaviors already built into the software. At the same time custom behaviors can also be created as new boxes for reusing. Combining the default and the custom behaviors/boxes, complex programs can be built rapidly using Choregraphe. All the interaction modules programmed with NAO in this project are developed using Choregraphe. Custom behaviors are built within Choregraphe using Python and Naoqi API.

We have developed a logging module and installed it into the Naoqi library of NAO that is compatible with Choregraphe to log the starting timestamp of each behavior of each interaction module. The sensors we have used in this project store EDA, HRV, and eye gaze data with their corresponding timestamp. The purpose of this module is to map sensor data with the interaction modules developed with NAO. After an experiment is performed with a participant, a student investigator will immediately transfer the log data to a secure password protected computer along with the sensor data. A separate Python module is built that maps the activities of the interaction modules with sensor data using the log data for analysis in the future.

All the programs developed in this project are hosted in private github repositories for version controlling and future collaboration.

E. Study Population

The population will be comprised of individuals who are aged 5-18 with communication disorders. Communication conditions that would be included are: articulation disorders, phonological disorders, mild receptive and/or expressive disorders, fluency disorders, and mild social communication disorders. The condition can not impair the participants ability to understand and speak English and to follow basic directions. All conditions that are targeted should be able to follow basic three step commands that are mastered by age four. Therefore, to minimize the risk of participation variance across all conditions, the activities presented do not exceed this threshold. According to statistical power analysis, at least 33 participants will be needed for the experiment. Demographics is not a parameter being examined for research, so no efforts will be made to equalize gender of participants.

F. Data Banking

Potential data containing private health information will be from video recordings via VALT, a secure video recording and reviewing system in the designated site. VALT is only accessible on the secure computers on the university with a secure login only available to individuals involved on the research team. Videos will be permanently removed at the end of the study.

Paper products, such as consent forms, will be kept in a locked drawer at a secure site. This will only be accessible by the student and principal investigators. Any paper forms will be shredded at the end of the study.

Language, eye gaze, EDA, and HRV data will not contain any private health information that can be linked to the identity of the individual. EDA and HRV data will be uploaded to the

secure cloud platform of Empatica (The manufacturer of Empatica E4 sensor) server during the sessions with no PHI. Post experiment, the data will be downloaded and removed from the server immediately. The eye gaze data will initially be stored on a SD memory card within the eye glasses recording unit. The data from both sensors will be immediately downloaded and stored on secure password-protected computers only accessible by the primary and student investigators. Once ensured the data has been properly stored, the information from the recording devices will be removed. Language data will also be stored with the sensor data on the secure computers. Once the data has been analyzed to its fullest extent, the student investigator(s) will ensure it has been deleted from the secure computers.

All members of the research team (principal and student investigators) at the University of Minnesota (UM) are required to have the below certifications in order to be a part of the study and access the information:

- A state of Minnesota General Information Services (GIS) background check to comply with the Board of Regents policy, Safety of Minors
- Computer based training (CBT) to comply with UM requirements for: HIPAA Training (HIP019)
- Information Security Awareness Training (Required) (ISA101)
- Safety of Minors (OHR138)
- UM Emergency Preparedness (UHS700)

IV. POTENTIAL STUDY ENDPOINTS/OUTCOMES

Following is a list of identified potential research questions and their potential outcomes for the proposed study:

- 1) How do arousal rates, such as heart rate variability and electrodermal activity, correlate to nonverbal social communication behaviors (including facial expressions, gestures, tone of voice, space, etc.) in a child who has a communication disorder during different interactions with a robot partner as compared with a human therapist partner?
 - a) Is there a difference between the physiological measures, such as heart rate variability and electrodermal activity, and/or external nonverbal communication behaviors, such as facial expressions, gestures, tone of voice, and proxemics, when the child with a communication disorder is participating in social play games with a human therapist versus while participating in social play games the humanoid robot, NAO?

Potential Outcome: Outcomes for this question will help therapists understand if robots such as NAO, will be a beneficial facilitator for groups of children who have social communication disorders. This can be seen through comparing the internal physiological responses, electrodermal activity and heart rate variability, to the external actions of the child, or the nonverbal social communication behaviors, such as facial expressions, gestures, tone of voice, and proxemics. The goal is to find the partner who elicits the most communication attempts in this population to

increase therapeutic outcomes across a variety of communication disorders.

2) For a child with a communication disorder, how does their joint attention, turn taking, and eye gaze compare between interaction with a humanoid robot and a human therapist?

a) With which interaction partner are there more fixations (gaze behavior measurement of attention), with the human therapist or with the humanoid robot?

Potential Outcome: An outcome from this question will be to determine the difference in fixations (if any) between the two partners. This could help determine if a robot could be used as a facilitative tool in therapy sessions where increasing fixation is a goal.

3) Does the amount of verbal expressive language, as measured by MLU, TTR, and pragmatic language use, change in children with communication disorders when interacting with an adult human therapist compared to interacting with a humanoid robot?

a) Is there a relationship between the mean length of utterance (MLU) when a child interacts with a human therapist versus with the humanoid robot?

Potential Outcome: An outcome from looking at the mean length of utterance (MLU) will be to determine the difference in average length of utterance (if any) between the two partners. This could help determine if a robot could be used as a facilitative tool in therapy sessions where increasing overall utterance length is a goal.

b) Is there a relationship between the type of language used, measured by type-token ratio (TTR), when a child interacts with a human therapist versus with the humanoid robot?

Potential Outcome: An outcome from looking at type-token ratio (TTR) will be to determine the difference in type of language used (if any) between the two partners. This could help determine if a robot could be used as a facilitative tool in therapy sessions where increasing language complexity is a goal.

c) Is there a difference in pragmatic language (greeting behaviors, proximity, parting behaviors) when a child interacts with a human therapist versus with the humanoid robot?

Potential Outcome: An outcome from looking at pragmatic language will be to determine the difference in pragmatic language used (if any) between the two partners. This could help determine if a robot or human therapist could be a better model for teaching pragmatic language targets to children.

V. PROTOCOL MODIFICATION FOR REMOTE STUDY

In response to the Covid-19 pandemic situation, when an in-person human study poses a health risk to the participants and the research investigators, a modified version of the protocol was also designed so that the study can be performed online through video conferencing. This version is designed in a way so that it imitates the originally proposed protocol as close as possible. The key ideas behind this version are mailing out the sensors to the participants, conducting the

study intervention through a video conferencing tool like zoom, and getting back the sensors through the mail. Though this process will potentially slow down the data collection process, it can be an alternative when in person study is not an option. Just like the original one, the course of this version is designed to take place in two major parts.

In part 1, the participants and legal guardians will attend the orientation through a zoom meeting. The timeline and agenda of the orientation zoom meeting will be the same as the original study. The timeline is described in Table I.

After the orientation, the necessary items will be mailed to the participant. Mailing out the sensors to the participants has a potential risk of loss and damage to the sensors. Considering this, Tobii Glasses eye gaze tracker, which costs more than \$10000, won't be used in this design. Instead, procedures will include taking subjective data, the approximate positioning of eye direction, such as towards the video conferencing camera or away from the camera. Also, the Empatica E4 EDA and HRV sensor won't be used in this design since it is also an expensive one. For this design, we have bought and tested the Mindfield eSense Skin Response EDA sensor which costs around \$100 but still gives reliable data. It is developed by Mindfield Biosystems Ltd. It's not a wristband-like sensor like the previous Empatica E4 one. Instead, one needs to wear two electrodes around the index and middle fingers. See figure 6



Fig. 6. Mindfield eSense Skin Response EDA sensor

It can easily be connected to an android/iphone and it has a free android/iphone app. EDA data recording session can be started and stopped easily with the app. Once a session is finished, users can share the data through email within the app. After careful consideration, this EDA sensor is decided as the only sensor to be used in this design. Since both EDA and HRV measure one's arousal or stress level, HRV won't be measured in this design as well. So, the participants need to deal with only one sensor. This EDA sensor along with a simplified instruction set to use the sensor will be mailed to the participants.

In part 2, on the experiment day, the participant along with the legal guardian will join the experiment through a zoom meeting. A student investigator will join the meeting along with NAO and control NAO throughout the meeting. A laptop will be placed in front of NAO in a way so that the participant

can have a full view of NAO through the webcam. The laptop will be connected to a noise cancelling and sound amplifying microphone placed in front of NAO so that the participant can listen to NAO properly over zoom. A speaker will also be connected with the laptop to amplify the sound so that NAO can properly listen to the participant’s response. Another student investigator will join the meeting who will act as the human therapist. The primary investigator/another student investigator will preside over the meeting and guide everyone throughout the process. The zoom meeting will be recorded directly onto Box, which is a very secure cloud storage service used by the University of Minnesota system. The recorded video will be securely accessed through Box only by the research team later for non-verbal social communication behavior and language characteristics analysis. At the end of the study the recordings will be permanently removed. The agenda and timeline for this part is described in table III.

After the study intervention, steps will be taken to get back the sensor through mail. All the other parts of the study will be exactly the same as the originally proposed one.

TABLE III. Online Protocol Part 2 Timeline

Activity	Time (Minutes)	Elapsed Time (Minutes)
Place, connect, and calibrate EDA Sensor	2-6	2-6
Rest to obtain baseline data	3	5-9
Playing Game: Activity One (Simon Says) with Human Therapist	2-4	7-13
Playing Game: Activity Two (Guess the Animal Game) with Human Therapist	2-4	9-17
Playing Game: Activity Three (Storytelling Activity) with Human Therapist	2-4	11-21
Rest: Reobtain baseline	3	14-24
Playing Game: Activity One (Simon Says) with NAO	2-4	16-28
Playing Game: Activity Two (Guess the Animal Game) with NAO	2-4	18-32
Playing Game: Activity Three (Storytelling Activity) with NAO	2-4	20-36
Instruct the legal guardian to share EDA data with the research team	3	23-39
End of activities: Brief survey and answer participants’ questions	4	27-43
Total Time	27-43	27-43

VI. POTENTIAL STUDY EXTENSIONS

A potential extension of the proposed study could be involving typically developed children along with children with communication disorders to engage in child-robot interaction. The extension of the study will focus on a comparative analysis of non-verbal social communication behaviors (facial expressions, tone of voice, gesture, proxemics, etc.), language characteristics (mean length utterance, type-token ratio, pragmatic language usage), and physiological patterns (EDA, HRV, and eye-gaze) between typically developed ones and children with communication disorders during different stages of child- robot interaction. While the originally proposed study would concentrate on how effective humanoid robots are as a therapeutic tool for children with communication disorders, this extension would concentrate on the similarity or dissimilarity in the correlation

of physiological patterns with verbal or non- verbal communication characteristics during different stages of child-robot interaction between these two groups. This will allow researcher to gain a better understanding of the internal behavior of the children having communication disorders. As we would solely focus on child-robot interaction in this version of the study, two more interaction modules have been developed with NAO. While the initial interaction modules focused on turn taking, these two interaction modules will focus on giving and following commands.

1) *Eye Color Changing Game*: In this activity, Nao will change its eye color according to the child’s command. NAO has LEDs in its eyes and it can be programmed to light its eyes with any color. At the beginning of the activity, NAO will show the participant a demo of how it can light its eyes with different colors. Then NAO will ask the child to say out loud any color of choice and wait for the command of the child. NAO will change its eye color according to the command of the child. This will continue for another 4 cycles.

2) *Touch Challenge*: In this activity, NAO will command the child to touch different parts of its body within a set period of time. If the child succeeds to touch the correct part within the time-frame, the child will score a point. At the beginning of the activity, NAO will sit down so that the child can easily touch different parts of its body. Then it will show the participant the exact parts of its body where the participant would need to touch when the challenge starts. The parts are:

- Left foot bumper
- Right foot bumper
- The front button on top of its head
- The middle button on top of its head
- The rear button on top of its head
- The back of its left hand
- The back of its right hand

Then NAO will ask the participant if they want to play the game in easy or hard mode. In easy mode, the child has to touch the correct spot in 4 seconds. In hard mode, the child has to touch the correct spot within two seconds. Once the difficulty is set, NAO will randomly choose any of the above mentioned parts and ask the child to touch it within the time-frame. If the participant succeeds to touch the correct spot within the time-frame, the participant will score a point. NAO will keep track of the score and announce after each round. This will continue for another 4 cycles.

Another potential extension of the protocol could be performing the original study on children with Autism Spectrum Disorder (ASD). Autism Spectrum Disorder is a neurological developmental disorder that generally affects a person in their early childhood and lasts for a lifetime. It generally affects a person’s ability to communicate with others and engage in social interaction. Like communication disorders, early diagnosis and therapeutic interventions are often effective measures for treating this disorder. The application of humanoid robots through child-robot interaction has become widely popular recently as a part of these therapeutic interventions. The proposed protocol can be used to study if

robots such as NAO is a beneficial facilitator for this group of children for therapeutic interventions.

VII. CONCLUSION AND FUTURE WORKS

The designed study protocols have been successfully tested through extensive dry-runs by the primary and student investigators. This comprehensive experimental setup can readily be used to perform the study on children with communication disorders. The results can be used to better understand the social interactions between humanoid robots and children with communication disorders for therapeutic intervention. This will allow researchers to gain a better understanding of the expressive language characteristics of children with communication disorders using non-verbal and verbal measures in correlation to physiological response data. Answering the identified potential research questions, future researchers will have a better comprehension if robots such as NAO is a beneficial facilitator for therapeutic intervention for groups of children who have communication disorder.

ACKNOWLEDGMENT

We acknowledge the collaborative support we got from Mark Mizuko, David Bohil, Brianna Stevensen and Claire Bieganek from the Dept. of Communication Sciences and Disorders, University of Minnesota-Duluth and Viann Nguyen-Feng and Lexi Asplund from the Dept. of Psychology, University of Minnesota-Duluth.

REFERENCES

- [1] Leo Kanner (1943), "Autistic Disturbances of Affective Contact", <http://simonsfoundation.s3.amazonaws.com/share/071207-leo-kanner-autistic-affective-contact.pdf>
- [2] Critchley, Hugo, Yoko Nagai, "Electrodermal Activity (EDA)", *In: Encyclopedia of Behavioral Medicine 2013*. Ed. by Marc D. Gellman and J. Rick Turner. New York, NY: Springer New York, pp. 666–669. https://doi.org/10.1007/978-1-4419-1005-9_13
- [3] D. Feil-Seifer and M. J. Mataric, "Defining socially assistive robotics," 9th International Conference on Rehabilitation Robotics, 2005. ICORR 2005., 2005, pp. 465-468, doi: 10.1109/ICORR.2005.1501143.
- [4] Dumouchel Damiano. (2017). Living with Robots. <https://doi.org/10.4159/9780674982840>
- [5] Kumazaki, Hirokazu et al. (2018). "The impact of robotic intervention on joint attention in children with autism spectrum disorders". *In: Molecular Autism* 9.1, p. 46. doi: 10.1186/s13229-018-0230-8.
- [6] S. Shamsuddin, H. Yussof, L. Ismail, F. A. Hanapiah, S. Mohamed, H. Piah, "Initial response of autistic children in humanrobot interaction therapy with humanoid robot NAO", *In: 2012 IEEE 8th Inter11 national Colloquium on Signal Processing and its Applications*, pp. 188–193.
- [7] Warren, Z. E., Zheng, Z., Swanson, A. R., Bekele, E., Zhang, L., Crittendon, J. A., Weitlauf, A. F., Sarkar, N. (2015). Can Robotic Interaction Improve Joint Attention Skills? *Journal of Autism and Developmental Disorders*, 45(11), 3726–3734. doi: 10.1007/s10803-013-1918-4
- [8] Z. Zheng, L. Zhang, E. Bekele, A. Swanson, J. A. Crittendon, Z. Warren and N. Sarkar. "Impact of robot-mediated interaction system on joint attention skills for children with autism" *In: 2013 IEEE 13th International Conference on Rehabilitation Robotics (ICORR)*, pp. 1–8.
- [9] Kennedy, J., Baxter, P. Belpaeme, T. Nonverbal Immediacy as a Characterisation of Social Behaviour for Human–Robot Interaction. *Int J of Soc Robotics* 9, 109–128 (2017). <https://doi.org/10.1007/s12369-016-0378-3>
- [10] Druga, Stefania Williams, Randi Breazeal, Cynthia Resnick, Mitchel. (2017). "Hey Google is it OK if I eat you?": Initial Explorations in Child-Agent Interaction. 595-600. 10.1145/3078072.3084330.
- [11] Miskam, Mohd Hamid, Mohd Yussof, Hanafiah Shamsud- din, Syamimi Malik, Norjasween Basir, Siti. (2013). Study on Social Interaction between Children with Autism and Humanoid Robot NAO. *Applied Mechanics and Materials*. 393. 573-578. 10.4028/www.scientific.net/AMM.393.573.
- [12] Conti, Daniela Trubia, Grazia Buono, Serafino Di Nuovo, Santo Di Nuovo, Alessandro. (2019). Affect Recognition in Autism: A single case study on integrating a humanoid robot in a standard therapy. 14. 66-87. 10.30557/QW000018.
- [13] M. A. Miskam, S. Shamsuddin, M. R. A. Samat, H. Yussof, H. A. Ainudin, A. R. Omar. 2014. "Humanoid robot NAO as a teaching tool of emotion recognition for children with autism using the Android app", *In: 2014 International Symposium on Micro-Nano Mechatronics and Human Science (MHS)* pp. 1–5.
- [14] Feil-Seifer D., Mataric c' M.J. (2009). Toward Socially Assistive Robotics for Augmenting Interventions for Children with Autism Spectrum Disorders. *In: Khatib O., Kumar V., Pappas G.J. (eds) Experimental Robotics. Springer Tracts in Advanced Robotics*, vol 54. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-00196-3_24 [15]Breazeal, C., Harris, P. L., DeSteno, D., Kory Westlund, J. M., Dickens,
- [15] L., and Jeong, S. (2016b). Young children treat robots as informants. *Top. Cogn. Sci.* 8, 481–491. doi: 10.1111/tops.12192
- [16] Kory, J. (2014). Storytelling with robots: effects of robot language level on children's language learning (Master's Thesis). Retrieved from ACM Digital Library Complete.
- [17] Kory Westlund, J. M., Jeong, S., Park, H. W., Ronfard, S., Adhikari, A., Harris, P. L., Breazeal, C. L. (2017). Flat vs. expressive storytelling: young children's learning and retention of a social robot's narrative. *Frontiers in Human Neuroscience*, 11. doi:10.3389/fnhum.2017.00295
- [18] Kory Westlund, J. M., Breazeal, C. (2019). A long-term study of young children's rapport, social emulation, and language learning with a peer-like robot playmate in preschool. *Frontiers in Robotics and AI*, 6. doi: 10.3389/frobt.2019.0008
- [19] Krupa, N., Anantharam, K., Sanker, M., Datta, S., Sagar, J. V. (2016). Recognition of emotions in autistic children using physiological signals. *Health and Technology*, 6(2), 137–147. <https://doi.org/10.1007/s12553-016-0129-3>
- [20] Bal, E., Harden, E., Lamb, D., Van Hecke, A. V., Denver, J. W., Porges, S. W. (2010). Emotion recognition in children with autism spectrum disorders: relations to eye gaze and autonomic state. *Journal of Autism and Developmental Disorders*, 40(3), 358–370. <https://doi.org/10.1007/s10803-009-0884-3>
- [21] Stagg, S., Davis, R., Heaton, P. (2013). Associations between language development and skin conductance responses to faces and eye gaze in children with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 43(10), 2303–2311. <https://doi.org/10.1007/s10803-013-1780-4>
- [22] Leite, I., Henriques, R., Martinho, C., Paiva, A. (2013). Sensors in the wild: Exploring electrodermal activity in child-robot interaction. 2013 8th ACM/IEEE International Conference on Human-Robot Interaction (HRI), 41–48. <https://doi.org/10.1109/HRI.2013.6483500>