2013

Charlie: A New Robot Prototype for Improving Communication and social Skills in Children with Autism and a New Single-point Infrared Sensor Technique for Detecting bBeathing and Heart Rate Remotely

Laura Boccanfuso
University of South Carolina - Columbia

CHARLIE: A New Robot Prototype for Improving Communication and Social Skills in Children with Autism
and
A New Single-Point Infrared Sensor Technique for Detecting Breathing and Heart Rate Remotely

By

Laura Boccanfuso

Bachelor of Arts
George Mason University, 1990

Master of Science
Bowling Green State University, 2001

Submitted in Partial Fulfillment of the Requirements

for the Degree of Doctor of Philosophy in

Computer Science

College of Engineering and Computing

University of South Carolina

2014

Accepted by:

Jason M. O'Kane, Major Professor

Manton Matthews, Committee Member

John Rose, Committee Member

Yan Tong, Committee Member

Ruth K. Abramson, Committee Member

Lacy Ford, Vice Provost and Dean of Graduate Studies

Acknowledgments

Thanks to my family. My amazingly supportive and bright husband Tony who never blinked when I told him that I wanted to go back to graduate school to complete a Ph.D. Thank you for being an incredible source of support, springboard and resource for me. I could not have completed this work without you. Carolina, Michael, Ana: thank you for understanding when I had to carve out a “little" time here and there to work during family time. You are each so remarkable in your own way and I am exceedingly proud to be your mom. You inspire me!

Many thanks to Jason M. O’Kane. Thank you for your valuable input and insightful guidance along the way. The countless electrical, mechanical and strategic implementation efforts you provided were key to building and making CHARLIE and the remote stress detection system work effectively. I’ve learned so much from you and I am grateful for the opportunity to step into the HRI research domain. It has made all the difference.

I also want to extend my heartfelt gratitude to Ruth K. Abramson, Harry H. Wright and Alicia V. Hall. Thank you for your unwaivering support in refining and field testing CHARLIE and for generously sharing your valuable time to make the CHARLIE study successful. Your insight and expertise was crucial to completing this research.

Thanks to Manton Matthews, John Rose and Yan Tong. I am honored to have each of you on my committee. Thank you for the many ways in which you helped me get through the program to the finish!

Finally, many, many thanks to Sarah Scarborough. I’ve been so very fortunate to have had the opportunity to work with you this past year. You have selflessly devoted well over 100 hours to field testing CHARLIE in addition to the many lives you regularly change through your important work. I am grateful for your generosity, insight and expertise.

Abstract

This research delivers a new, interactive game-playing robot named CHARLIE and a novel technique for remotely detecting breathing and heart rate using a single-point, thermal infrared sensor (IR). The robot is equipped with a head and two arms, each with two degrees of freedom, and a camera. We trained a human hands classifier and used this classifier along with a standard face classifier to create two autonomous interactive games: single-player ("Imitate Me, Imitate You") and two-player ("Pass the Pose"). Further, we developed and implemented a suite of new interactive games in which the robot is teleoperated by remote control. Each of these features has been tested and validated through a field study including eight children diagnosed with autism and speech delays. Results from that study show that significant improvements in speech and social skills can be obtained when using CHARLIE with the methodology described herein. Moreover, gains in communication and social interaction are observed to generalize from child-to-robot to co-present others through the scaffolding of communication skills with the systematic approach developed for the study. Additionally, we present a new IR system that continuously targets the sub-nasal region of the face and measures subtle temperature changes corresponding to breathing and cardiac pulse. This research makes four novel contributions: (1) A low-cost, field-tested robot for use in autism therapy, (2) a suite of interactive robot games, (3) a hand classifier created for performing hand detection during the interactive games, and (4) an IR sensor system which remotely collects temperatures and computes breathing and heart rate.

Interactive robot CHARLIE is physically designed to be aesthetically appealing to young children between three and six years of age. The hard, wood and metal robot body is cov-

ered with a bright green, fuzzy material and additional padding so that it appears toylike and soft. Additionally, several structural features were included to ensure safety during interactive play and to enhance the robustness of the robot. Because children with autism spectrum disorder (ASD) often enjoy exploring new or interesting objects with their hands, the robot must be able to withstand a moderate amount of physical manipulation without causing injury to the child or damaging the robot or its components. CHARLIE plays five distinct interactive games that are designed to be entertaining to young children, appeal to children of varying developmental ability and promote increased speech and social skill through imitation and turn-taking.

Remote breathing and heart rate detection Stress is a compounding factor in autism therapy which can inhibit progress toward specific therapeutic goals. The ability to non-invasively detect physical indicators of increasing stress, especially when they can be correlated to specific activities and measured in terms of length and frequency, can relay important metrics about the antecedents that cause stress for a particular child and can be used to help automate the evaluation of a child's progress between sessions. Further, collecting and measuring critical physiological indicators such as breathing and heart rate can enable robots to adjust their behavior based on the perceived emotional, psychological or physical state of their user. The utility and acceptance of robots can be further increased when they are able to learn typical physiological patterns and use these patterns as a baseline for identifying anomalies or possible warning signs of various problems in their human users.

We present a new technique for remotely collecting and analyzing breathing and heart rates in real time using an autonomous, low cost infrared (IR) sensor system. This is accomplished by continuously targeting a high precision IR sensor, tracking changes in the sub-nasal skin surface temperature and employing a sinusoidal curve-fitting function, Fast Fourier Transform (FFT), and Discrete Wavelet Transform (DWT) to extract the breathing and heart rate from recorded temperatures.

Table of Contents

LIST OF TABLES

List of Figures

CHAPTER 1

INTRODUCTION

The use of robots for cooperative work with humans is becoming increasingly pervasive across an ever widening range of disciplines. Medical procedures using robots are reported to be less invasive, result in faster recovery times and are estimated to have nearly tripled from 2007 to 2010 [1]. Robots have been employed for use in post-stroke rehabilitation [2], as assistive feeding systems for the physically handicapped [3] and in therapeutic roles such as robotic pets for the elderly in nursing homes [4]. Engineers across multiple disciplines have capitalized on the unique qualities of robots to perform autonomously and predictably and repeat mechanical tasks consistently. These characteristics also make robots well suited as part of an early intervention strategy for many autistic children who tend to perceive them as nonthreatening and intrinsically interesting.

Robots have been used to effectively engage autistic children in interactive game playing and research has demonstrated that robot-assisted autism therapy promotes increased speech and increased child-initiated interactions in children with Autism Spectrum Disorder (ASD) [5, 6]. However, more research is needed to develop definitive paradigms that describe which types of autonomous robot designs and interactive modalities will most benefit children with autism.

According to a 10-state study conducted by the Center for Disease Control (CDC) [7] the number of diagnosed autism cases increased an average of 57% from 2002 to 2006. In early 2013 the CDC reported that the occurrence of autism spectrum disorders (ASD) in the United States was 1 in 88 births. Not only does this translate to a growing population of autistic children but it also means that existing resources used to treat and care for children

with autism are under greater strain. Further, because of the added expense of therapy and specialized medical care, the cost of raising an autistic child in the United States is estimated to be between 8.5 to 9.5 times greater than raising a typically developing child. This additional financial burden may mean that some families have to choose whether to incur significant debt to get the proper care for their child or limit the amount of therapy their child receives. Even though robots have been proven to be effective for promoting communication with some autistic children, there are few existing robots currently in use for autism therapy and those that do exist are cost prohibitive for widespread use.

Our research focuses on achieving two primary objectives. The first objective is to design and develop an interactive robot that can be used to promote speech and social interaction among children with autism, is suitable and sufficiently robust to be handled by children and is financially accessible to those who would most benefit from its use. Our second research focus is inspired by the desire to make human-robot interactions more natural and productive by providing a technique for remotely detecting subtle physiological changes that correspond to stress. The two most important questions this research seeks to answer are: (1) Can a simple, low-cost robot design be effective for promoting human-to-human interaction with autistic children? and (2) Can minimal temperature data collected with a single-point, non-contact infrared sensor be sufficient to accurately calculate breathing and heart rate?

A simple, low-cost robot design for promoting imitation and turn-taking skills

Basic turn-taking and imitation skills are imperative for effective communication and social interaction [8]. Research has shown that interactive games using turn-taking and imitation have yielded positive results with autistic children who have impaired speech or social skills [9]. In [10] we present research in which we designed and built a toy-like robot with face and hand detection capabilities to autonomously engage autistic children in interactive games using imitation and turn-taking skills. The robot is equipped with a head and two

Figure 1.1: Complete robot (top left). Snap-off arm (top right). Snap-off head (bottom)

arms, each with two degrees of freedom, and a camera. For robustness and safety during play, the robot's arms and head are fully detachable (Figure 1.1). Additionally, a human hands detector was trained and subsequently, this detector was used along with a standard OpenCV face detector [11] to create two autonomous interactive games: single-player ("Imitate Me, Imitate You") and two-player ("Pass the Pose".)

In "Imitate Me, Imitate You", the robot has both passive and active game modes. In the passive mode, the robot waits for the child to initiate an interaction by raising one or both hands. In the active game mode, the robot initiates interactions by assuming a pose

and detecting when the child imitates. The "Pass the Pose" game engages two children in cooperative play by enlisting the robot as a mediator between two children who alternately initiate and imitate poses. These games were expressly designed to increase joint attention and encourage child-led interactions through games that are based on turn-taking and imitation. Because the frequency and duration of each child's participation is continuously measured, the robot is able to adapt its game mode(s) based on the perceived interest of the child. Three additional teleoperated, human-in-the-loop games were also developed where CHARLIE is employed as a catalyst for improved social interactions. During gameplay in one of the teleoperated modes, children learn and practice verbal requests with the robot (and receive positive reinforcement for each attempt) before generalizing the game to co-present others. The three new contributions presented in this part of our research are: (1) a new low-cost robot design which measures and adapts its behavior according to a child's actions, (2) five new interactive games (two are autonomous and three are teleoperated) and, (3) a new hand classifier used for hand detection, which is now freely available for use in various kinds of human-robot interactions.

Remote collection of breathing and heart rate

Remote breathing and heart rate detection is valuable for a multitude of applications including rehabilitative robotic applications such as post-stroke and post-operative cardiac therapies, socially assistive robots used to help developmentally disabled children and cognitively impaired adults, search and rescue robots which may evaluate the physical condition of victims found at a disaster site and personal or home robots which work in close proximity to humans. Our approach is an important potential improvement in scenarios where user mobility is an inherent part of the therapy, when users have a general aversion to being fitted with sensors or when the use of biofeedback sensors is otherwise impractical. Further, due to its relatively small size and modular design, existing robot systems can be retrofitted with the proposed detection system to enhance and extend their functionality.

Detecting and tracking the physiological state of humans is an important focus for research in human-robot interaction (HRI) because it promises to make robots better-suited to work in close proximity and more cooperatively with humans. Collecting and using physiological indicators can enable robots to adjust their behavior based on the emotional, psychological or physical state of their user. In addition, the overall utility of robots can be further increased when they are able to learn typical physiological patterns and use these patterns as a baseline for identifying anomalies or possible warning signs of various problems in their human users. For example, if an autistic child becomes distraught during the course of therapy, he or she may not be able to appropriately communicate this fact to the therapist or teacher. A robot that can detect and monitor a child's breathing and heart rate may track subtle shifts in his or her emotional state and change its behavior before the child's frustration escalates.

Additionally, a robot that continuously collects heart rate and breathing data from a patient undergoing post-stroke therapy can adjust the amount of exertion in a given exercise or the duration and number of repetitions so as to challenge the patient without pushing them beyond their physical limits. Finding an efficient way to accurately detect stress remotely for real-time applications is the necessary next step towards fully realizing this potential.

Contact modalities exist for obtaining physiological information from a user but they require that the user wear specialized sensors or that the user make repeated or continuous contact with the part of the robot fitted with a specialized sensor. These techniques have relied on wearable sensors such as thermistors, respiratory gauge transducers, pulse oximeters and acoustic sensors. While contact devices typically deliver accurate physiological data, they are not suitable for many mobile applications or for people who are generally averse to wearing sensors. Further, although solutions exist using non-contact methods such as infrared video cameras, radar and doppler techniques, these approaches rely on high-cost equipment and collecting and analyzing very large amounts of data at a

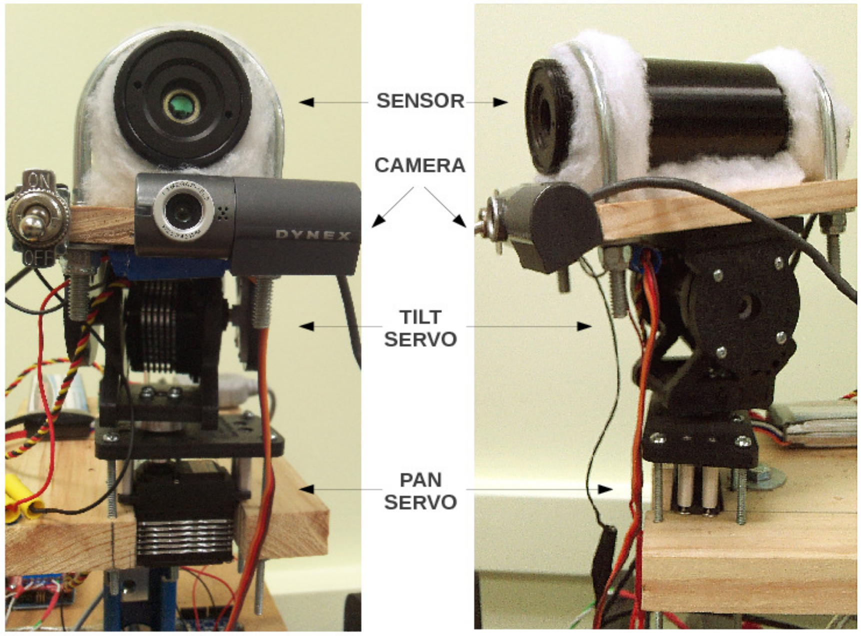

Figure 1.2: Single-point FAR infrared sensor.

high processing cost.

In a submission to the International Conference on Biomedical Robotics and Biomechatronics (BioRob) [12], we present a new technique for capturing changes in the sub-nasal skin surface temperature to monitor breathing events remotely. Temperatures are recorded in real-time using a high precision, single-point infrared (IR) sensor and the breathing rate is automatically extracted using a sinusoidal curve-fitting function which provides an estimated rate in breaths per minute. Results from preliminary tests show this system effectively captures breathing rates within an error rate of under 2 breaths per minute in approximately 70% of typical test cases.

In subsequent research presented at the International Conference on Social Robotics (ICSR) [13], we extend the IR technique to perform real-time collection and analysis of heart rate using the same IR sensor system we used for breathing detection. Due to the sometimes poor curve-fitting performance that resulted from occasional irregularities in

breathing and given that heart rate data is known to be nonstationary, we implemented a Discrete Wavelet Transform (DWT) to process collected temperatures. Accuracy was improved over initial results obtained from applying the Fast Fourier Transform (FFT) technique and experiments showed that in 72.7% of typical cases heart rate was successfully detected within 0-9 beats per minute over a ten-minute session as measured by root-mean-square error (RMSE).

Chapter 2

Related Research

2.1 Autism, Impact and Computer-Assisted Therapies

Social psychology literature emphasizes three basic developmental milestones in human social interaction during the first year of life between a child and the caregiver [14,15]:

> **One to two months of age** - Initiates and reciprocates eye contact with a caregiver; exchanges vocal and facial expressions, establishing a pattern of interaction based on prompt and response between child and caregiver.
>
> **Three to nine months of age** - The child expresses his/her desires, displeasure and pleasure and the caregiver interprets and responds. Gradually the child begins to develop the ability to predict the caregiver's response, making their interactions more symmetric.
>
> **Ten months and up** - Child-caregiver interactions continue to develop, resulting in the emergence of joint attention (where two people attentively look at the same object, as a result of either by pointing or directing their gaze.) Moreover, their attentiveness is sometimes accompanied by an awareness of the emotional assessment associated with the object. This important step includes vocalizations and facial expressions which relay an interpretation or meaning of the target of joint attention.

These critical social milestones provide the framework for identifying cognitive impairments relating to communication disorders. Since it was first described in 1943 by Leo Kanner, autism has been classified as a pervasive development disorder or cognitive impairment characterized by deficiencies in communication, social interaction, and creative or imaginative play [16]. The American Psychiatric Association's Diagnostic and Statistical Manual-IV (DSM-IV) provides standardized criteria describing the major impairments characterisitic of children diagnosed with autism which include :

> **Social (non-verbal) impairments** - Marked impairment in the use of non-verbal behaviors such as eye-to-eye gaze, facial expressions, body postures, and gestures used for social interactions. Also characterized by a failure to develop peer relationships appropriate to developmental level, a lack of spontaneous seeking to share enjoyment, interests, or achievements with other people and/or a lack of social or emotional reciprocity.
>
> **Linguistic (verbal) impairments** - Delay in, or total lack of, the development of spoken language (not accompanied by an attempt to compensate through gesture or mime.) For individuals with adequate speech, a marked impairment in the ability to initiate or sustain a conversation with others. Additionally, stereotyped and repetitive use of language or idiosyncratic language and a lack of varied, spontaneous make-believe play or social imitative play appropriate to developmental level.
>
> **Imaginative impairments** - Encompassing preoccupation with one or more stereotyped and restricted patterns of interest that is abnormal either in intensity or focus. Apparently inflexible adherence to specific, nonfunctional routines or rituals stereotyped and repetitive motor manners (e.g., hand or finger flapping or twisting, or complex whole-body movements) persistent

preoccupation with parts of objects.

The Autism Society of America (ASA) [17] estimates that the annual cost for services for autistic persons is $90 billion, with 90% of these costs dedicated to adult services. In addition, the ASA projects 10-17% annual growth in the occurrence of autistic diagnoses and the annual cost to be approximately $200-$400 billion in 10 years. To add to the sizeable financial strain and the emotional toll it takes, raising an autistic child can be taxing on relationships between family members.

The substantial impact of autism on familial relationships, steep medical costs and the debilitating challenges faced by those diagnosed with the disorder, has driven research over the last few decades to study the possible causes and to design reliable diagnostics and effective therapies for autism. It has been estimated that lifelong costs of care can be reduced by as much as 2/3 when autism is diagnosed and treated at an early age [17].

Since autistic children tend to show a partiality for interacting with computers, programs for the treatment of young autistic children have included various computer-aided therapies aimed at improving vocabulary and grammar acquisition [18]. While the immediate goal of teaching autistic children some of the basic tools use for language and expression using computer-assisted methods is often achieved, the ultimate measure of success in autism therapy encompasses more than acquiring the semantics of language. Successful therapies should promote long-term social integration and effective human-to-human communication. An example of computer-assisted autism therapy designed to mediate and promote social communication is a program designed at Stanford called SIDES.

In 2006, a group of researchers from Stanford University [19] designed a case study featuring a cooperative tabletop computer game aimed specifically at developing the social skills of children with Asperger's Syndrome (considered "high functioning" on the spectrum of autism disorders.) The Stanford team's experiment included 12 middle school students, most of whom had been diagnosed with Asperger's Syndrome and exhibited some

challenge in social interaction. Over a six-month period they conducted interviews with school mental health therapists and attended discussions to learn how to identify potential solutions for teaching group work skills. They learned that games had been frequently used in therapy sessions and that highly visual games were most effective for teaching children with Asperger's. The resulting game design was SIDES (Shared Interfaces to Develop Effective Social Skills) which encourages face-to-face interaction and promotes "listening, negotiation, and group work skills."

The tabletop computer game showed promise with the group of high-functioning students with Asperger's while raising additional questions about the best way to teach effective communication among the participants. The premise of the game, which features frogs and insects and an electronic game panel for each participant, challenges the players to play cooperatively to find the best path for the frog to eat the insect. One very positive outcome from the experiment was that the students were so engrossed in playing the game that they did not realize they were actively working in a group and building confidence in their own social abilities. The SIDES game proved to be engaging for the students and promoted positive social interaction among the participants.

2.2 Robot-Assisted Therapies for Autistic Children

Autism therapy should ultimately seek to promote human-to-human interaction. The aim of the Stanford study reflects the direction of much of the autism therapy research being carried out today. The intent to ultimately promote human-to-human interaction is realized with the SIDES game, where players must collaboratively work together and communicate effectively in order to succeed. The proliferation of robots in our culture today has presented similar challenges for developers of robust, modern robots. The recent focus on Human-Robot Interaction (HRI) and Socially Assistive Robots (SARs), has accelerated the efforts of robotics researchers toward developing approaches to promote effective and

natural communication between humans and robots. As this type of research progresses, the utility of robots and the benefits we receive from them increases. Autism treatment is a prime example of the direct benefits possible with natural and effective human-robot interaction.

It is widely accepted that autistic children tend to prefer computers and mimic robots. Recent research conducted by a team of psychologists from the University of Padua, University of Melbourne and Royal-Hollaway University of London, provides empirical evidence that "interaction with robots can trigger imitative behavior in children with autism" [20]. The study featured a group of twelve high-functioning autistic children, twelve normally developing children, a human model and a robot with a remotely controlled robotic arm. The experiment consisted of two basic conditions and one basic action. In the first condition, a participant was seated at a small table across from a human model. The human model reached for a small ball in the center of the table. Once this action was completed, a sound would signal for the study participant to perform the same reach-to-grasp object. The second condition, featured the same action, this time with a robot arm performing the reach-to-grasp action.

Results showed that autistic participants had faster response times and displayed more consistent responses with the robotic arm whereas control participants had faster response times with the human model. Researchers concluded that children with autism are visually primed predominantly by robots over humans, thus confirming what had been observed by other therapists. Because human action is characterized by great variability, the same human action repeated 20 times produces 20 unique trajectories and kinematics. It is likely that minute variances in human actions are overlooked by neurologically healthy children while children with autism may not know how to deal with such differences and therefore, respond better to action which is more predictable and repetitive. This study provided insight into the reasons some autistic children tend to prefer interacting with computers or robots over people.

The observations documented in the robotic arm study show great promise for successful robot-assisted therapy. Still, to be completely effective in this capacity, robots must be used for more than just visual priming and prompting action in the autistic child. Robots must be implemented as mediators whose role is more dominant at the start of therapy and becomes less so as the autistic child progresses. While empirical validation of visuomotor priming with robots is a fairly new course of research, the introduction of robots to autism therapy is not new.

The first experiment using a robot as a therapeutic tool for autistic children was conducted in 1976 by Weir and Emanuel [21]. Using the LOGO programming language (developed by MIT in the 1960s and widely used for teaching children), Weir and Emanuel designed a remote-controlled turtle robot to act as a catalyst for communication with an autistic child. Although the robot did not act autonomously and did not interact physically with the children in the study, the results from their research demonstrated the therapeutic potential that exists when pairing robots with autistic children and provided the foundation for the research that has since followed.

Another more recent example of robot-assisted autism therapy was conducted by artificial intelligence researchers at the University of Hertfordshire who initiated the AURORA (AUtonomous RObotic platform as a Remedial tool for children with Autism) Project in 1998. AURORA consists of a multidisciplinary research team dedicated to the exploration of the role of robots in autism therapy. According to their website (www.aurora-project.com), AURORA researchers are using the "robotic platform to attempt to bridge the gulf between the stable, predictable and safe environment of a simple toy (robot), and the potentially unpredictable world of human contact and learning." By introducing a simplified, predictable, interactive world with the robot, and gradually integrating more complicated interactions, it is believed that an autistic child can acclimate to the world at their own pace.

The initial approach of the AURORA project was to introduce robotic agents as social

mediators to provide a stable environment which would disarm the autistic child's fears and reduce the stress and pressure inherent to social interaction. During a symposium on Intelligent Robotics Systems in 1999 [22], Werry and Dautenhahn proposed that the robot's ability to provide structure and repetition as a friendly agent, makes it an ideal candidate for therapeutic use. Although robots have traditionally been designed to perform a task or to carry out a specific set of actions, researchers involved in the AURORA project make the distinction that the emphasis for robots in autism therapy will be on the interactions and expression of actions, not necessarily the completion of a given task. This is a fundamental departure from the standard design of robots and means that even simple tasks or actions can have significant impact if they are successful in producing the intended response. For example, if the sole objective of a robot is to establish and build trust in a user, the set of actions implemented can be effectively small so long as the actions are predictable, repetitive and non-threatening.

Of course, building trust is just one of the objectives for robots used in autism therapy. The robots must also be fun, intriguing and engaging in a non-threatening way for the children to remain actively interested. The original robot used by the AURORA team was a relatively small, flat-topped mobile robot with eight infrared sensors used for avoiding obstacles and a single positional heat sensor. A behavior-based architecture was implemented featuring a central decision module which would select appropriate actions from a menu of available choices. The architecture was further broken down into two levels. The first (lower) level actively monitored data read from the sensors, maintained the timer and was in charge of avoiding obstacles. The second level managed the selection of behaviors, deciding when to activate each new behavior and for how long.

The initial study included a group of five children, each of whom was given the opportunity to interact individually with the robot designed for the experiment. When the child became disinterested or showed signs of boredom, the supervising teacher would end the session. The team reported a number of very positive and promising results from the initial

interaction sessions observed. First, the study participants showed no fear of the robot and engaged in a positive manner for a relatively extended period of play. Second, the children enjoyed the robot interaction so much that they responded with laughter, vocalization and substantial eye focusing and attention focusing which is typically uncharacteristic of children with autism. The children were so engaged by the robot that stereotypical autistic "empty gazing" and repetitive behaviors were also reduced.

At the conclusion of this study, Werry and Dautenhahn reported that robots have an important role in rehabilitation partly because they are able to produce consistent, repeatable and reliable behaviors. Consistency is the key in establishing a level of trust between the robot and each participant, and serves as the basis for introducing gradual changes without provoking fear. Finally, but equally importantly, robots have proven that they can capture a child's attention and engage a child in activities that promote long-term learning.

2.3 Designs for Robot-Assisted Autism Therapy

Although the last decade has seen a significant rise in the development of humanoid robots, not all researchers agree that increasingly lifelike robots are more effective for all applications [5]. Dautenhahn argues that using human models for creating life-like robots cannot be applied universally for designing robotic social actors. She explains that creating robots to be as life-like as possible anthropomorphizes the robot and sets up unrealistic expectations which are likely to go unmet. Once behavioral expectations are not met, the believability of the agent is severely reduced. Furthermore, giving a robot life-like qualities drastically limits its usefulness. The example she uses to support this concept is a child and a simple wooden stick. Since the stick is not well-defined, it allows for creative play to take place and, with imagination, the stick can become a sword, an Indian's arrow, or a simple tool used for reaching a kite in a tree. Instead of creating a robot which closely imitates the unpredictability of real-world human behavior, she proposes introducing a fairly ritualized predictable robot which initially mediates between the autistic child and the unpredictable

world and gradually incorporates increasingly complex behaviors as the child develops.

Another interesting viewpoint of hardware design is based on the exploration of *how* autistic children interact with specific kinds of robots. In 2007, a study was conducted to examine the extent to which proprioceptive perception can be used to identify various positions and movements which indicate whether a child is carrying, rolling or throwing a robot named Roball [23]. The study featured a spherical robot (Roball) equipped with three accelerometers and three tilt sensors which were used to collect information about the robot's position and orientation. The data collected was then used to make inferences about its own state. Sensory information from the accelerometers was analyzed to determine three physical states including: (1) Alone, (2) Interaction and (3) Carrying (being carried.) Tilt sensor information was used to determine two additional states: (4) Spinning and (5) No Condition. The results of the study were mixed. The robot was highly likely to determine its correct state (greater than 90% accurate) when its sensors were consistent with Alone and Carrying. It was fairly likely to determine its state when the sensory data mapped to Spinning (77%) but highly unlikely to determine Interaction (10%).

Ultimately, this experiment confirmed that general environmental conditions can be detected through the collection of relatively crude sensory information. Information about how the user interacts with the robot is important information that may be insightful for both inferring the emotional state of the user and as a determinant for robot behavior. The authors expect that this type of proprioceptive information be used to adapt the robot's behavior to "create and sustain more meaningful and a broader range of interactions." Although there is still much research to be done in the area, this implementation of proprioceptive perception is one approach that seeks to integrate emotional or stress detection into the overall architecture of the robot.

Another study explores the use of robot sensors to collect and analyze tactile information during interactive play with autistic children by employing force sensors located at strategic points on the robot's body [24]. Touch is a necessary and important part of so-

cial development and is one of the most basic forms of communication. For many autistic children, it is the primary vehicle for exploring the world around them. The authors of this paper incorporate force sensors on the KASPAR robot in order to capture and classify characteristic touch patterns by autistic children engaged in interactive play with the robot. Three sensors are placed on each of KASPAR's hands, three on each arm, two on each shoulder and two on the head.

While preliminary, the study concluded that there is strong indication for high frequency of touch occurring in the hands, arms and head, autistic children tend to focus on one part of the robot's body during interaction (with the exception of the hands) and that it is possible to detect the length, location and extent of touch using the simple sensors included in this study. The total cost of the system is approximately 2500 USD.

Over the last 2 years, there has been a rapid acceleration of research which uses the humanoid NAO robot in intervention studies for children with ASD [25], [26], [27], [28], [29]. The NAO is a sohpisticated and versatile robotic research platform with 25 degrees of freedom, two cameras, four microphones, a sonar rangefinder, two IR emitters and receivers, one inertial board, nine tactile sensors, eight pressure sensors, a voice synthesizer, LED lights, and two high-fidelity speakers. For research purposes, the NAO provides a robust programmable platform which affords scientists and skilled educators the opportunity to employ the robot for many diverse investigative purposes. However, there are several functional factors that greatly limit the NAO's translational value outside of the laboratory or clinic as a socially assistive device for children with autism. First, the NAO is an expensive robot that is prone to damage from falling or overheating. Further, because each joint requires precise control for the robot to actuate smoothly and receive input from its sensors, even subtle damage can significantly affect the robot. Hardware such as force sensors, accelerometers, tilt sensors, servos and actuators give the robot more sophisticated capabilities and potentially deliver a richer set of data for the robot to base its decision-making. However, the incremental trade-off for incorporating each additional hardware component

is a corresponding increase in overall cost, complexity and, especially if deliberate attention is not given to protecting the structural integrity of the robot, a decrease in the robot's utility. In summary, complex and expensive robots are inaccessible to the majority of the large population of users for which they are intended and have limited translational and research value. As described earlier, physical manipulation and tactile exploration is an essential part of learning for children with and without ASD. Therefore, robots that are designed without careful regard for the human-robot interactions that will necessarily be part of their use, further limits the robot's overall practical usefulness.

Minimalistic Robot Designs

Over the past decade, the use of robots as social mediators has been explored as a tool for supplementing traditional autism therapies in order to teach and improve social skills. Robots are uniquely suited for engaging children with ASD since they tend to be perceived as predictable, non-threatening, and are able to perform repetitive tasks consistently and reliably [20, 30]. Most importantly, an increase in basic social and interaction skills has been observed when using robots for turn-taking and imitation games [31]. Some of the most promising results from robot-assisted autism therapy include an increased attention span, eye contact, child-led speech, improved turn-taking and imitative game playing skills and overall use of language [32]. Since social behavior is known to be very complex and subtle in nature, social interaction can appear to be unpredictable and extremely difficult to comprehend for a child with ASD and impaired social skills.

In one study, the use of a minimally expressive humanoid robot named KASPAR is tested as a communication facilitator for children with ASD [33]. KASPAR is a child-sized robot which uses an 8-degree of freedom (DOF) head, two 3-DOF arm movements and minimal facial expressions to interact with a human. Research with KASPAR assesses improvements in and the acquisition of interaction competencies of children with ASD while interacting with the robot by measuring body movement, eye gaze and non-verbal

communication with co-present others.

Experiments included a child with ASD, the investigator and the child's caregiver (i.e., teacher or parent), and were videotaped for subsequent analysis by a social psychologist. Trials were designed to allow the child(ren) to interact freely with the robot, in order to allow the child to explore the robot under their own terms and establish a level of trust. If the child indicated interest in interacting physically with the robot, they were allowed to touch, handle or teleoperate the robot using a remote control. In some scenarios, the investigator or caregiver would manipulate the robot. Preliminary observations of the interactions between KASPAR and children with ASD indicated that the combination of subtle changes in facial expression along with simple gestures was sufficient to convey various emotions to the child.

Field studies with KASPAR revealed that relatively low functioning children with ASD, who would not normally seek physical or eye contact, directly engaged with the robot and, in some cases, proactively touched and gazed at co-present others during sessions with KASPAR. Three successful cases were highlighted. In the first case, a six year-old girl with severe autism was introduced to KASPAR. At the time of the study, the child did not talk and refused all eye contact, even with her own family members. The girl was brought into the room by her mother and after a short acclimation period, she indicated her interest in KASPAR by reaching out to the robot. After she was moved closer to the robot, the girl used her hands to explore the robot's face, paying particular attention to its eyes. When KASPAR played the tamborine, the child attempted to imitate the motion. At one point, the girl even focused her attention on the investigator and reached out her hand to him, demonstrating the same kind of interactional practice she had shown with the robot.

The second case included a boy with severe autism who would interact regularly with family members at home, but would not proactively seek interaction with others at school. The child immediately showed interest in the robot, especially focusing on its face and eyes. The child was able to touch and explore KASPAR's face and eyes and he later turned

to his teacher to touch her eyes and then his own. After several sessions with KASPAR, the child began to share his excitement with his teacher by turning to her, reaching out to her and non-verbally encouraging her to engage her in the game with KASPAR.

Research with a robot named Keepon [34] has shown that it is also possible to promote social and communication skills in children with ASD using a non-humanoid robot. Keepon is a small, toylike robot with four degrees of freedom, with a simple physical design and is used for nonverbal interaction with children. Keepon's predecessor, Infanoid [35], is an upper-torso robot with 29 actuators, capable of expressing its attention and emotions (using eyebrows and lips.) Due to its many moving parts and the large amount of information that is conveyed when gestures and facial expressions are used, researchers found that the robot was perceived by some children with ASD to be overwhelming.

As a result, the research team developed Keepon which is capable of conveying emotion in a simplistic manner. The robot can convey excitement by bobbing up and down, pleasure by rocking from side to side, and fear by vibrating. The cameras in each of the robot's eyes and the robot's ability to orient its head provide the capability for establishing and measuring eye contact, directing gaze and identifying objects of joint or shared attention. By combining these simple actions, the robot can convey not just *what* the object of interest is, but also *how* it perceives the target.

Three sets of experiments were conducted. In the first tests, three age groups of typically developing children 9 months to 3 years old were allowed to interact with Keepon in a controlled setting. These experiments showed that children in each age group interacted with the robot in fundamentally different ways. The youngest test group (0-1 year olds) primarily explored the robot using their hands or mouth and although they did not seem to respond to Keepon's directed attention, they responded positively to the robot's emotive actions by laughing or bobbing their bodies when the robot bobbed its own body. The middle group (1-2 year olds) also examined the robot through tactile exploration, but showed an awareness of the robot's attention focus, sometimes even following its gaze. Addition-

ally, several children in this age group mimicked the robot's positive emotive expressions by bobbing and rocking their own bodies. The oldest age group (2+ year olds) showed a progressive understanding of the robot. Upon first being introduced to it they would watch it carefully and watch how the caregivers interacted with it. Then, upon recognizing that its actions were predicated not only on an object of interest but also by an appraisal of the object of its attention, would begin to treat it as a social agent - showing it toys, stroking its head and verbally interacting with it.

The second tests were conducted in a preschool playroom with approximately 30 typically developing children between 3 and 4 years of age, where the robot was present but no instructions for when or how to interact with Keepon were given to the children. Observations regarding how the robot's various actions were interpreted, expressed and shared among the children were documented. Four basic styles of play among the children were observed: (1) Violent/Protective, (2) Caregiving, (3) Demonstrative and (4) Self-conscious. Each of these expressions during play are indicative of the children's perception of the robot as more than a mobile "thing." Instead, the children attributed communicative meaning to the robot's simple gestures and sounds. Further, during free-play time dyadic interactions between one child and the robot were observed in addition to several cases of n-adic interactions. In the n-adic interactions, the robot became an object of shared interest which spawned interpersonal play with other children and the school teacher.

The last set of experiments were conducted in a day-care center for children with some form of autism between 2 and 4 years of age, over a period of 3 years. Of the approximtely 30 children observed during the course of this study, three representative cases were detailed in the research.

The first child was a non-verbal 3-year-old girl diagnosed with autism and moderate mental retardation. In her first 4 sessions the girl would avoid the robot, keeping her distance and averting her eyes from its gaze. After observing a boy interacting with the robot (during the 5th session), the girl went to her therapist, pulled her by the arm and indicated

that she wanted the therapist to imitate the boy's actions. During her last few observed sessions with the robot (sessions 11-15), the girl would clothe the robot, look into its eyes, kiss it and vocalize non-words to it. Of specific value are the dyadic interactions between the girl and the robot that emerged from the first and the last sessions.

The second case featured another non-verbal 3-year-old girl with autism and moderate mental retardation. Her participation in the study lasted approximately 17 months, and about 40 sessions. For the first 9 sessions the girl would not interact or pay attention to the robot at all, even though she would glance at it occasionally when it made a noise. After her 10th session, the girl began to touch and interact with the robot. During the 16th session she poked the robot in the nose, causing the robot to bob up and down. The girl showed her surprise and smiled while the others in the playroom burst into laughter. In subsequent sessions, the girl would interact with the robot, smile and look referentially at her therapist and mother. For the last 10 sessions or so, the girl would regularly participate in a game of imitation and turn-taking with the robot, while repeatedly and referentially looking at her therapist and mother. The marked change from non-interaction to triadic interaction, especially given her reluctance to engage in this kind of play prior to the test, is an important observation.

The last case describes a boy (of an undisclosed age) diagnosed with Asperger's syndrome and mild mental retardation. In his first encounter with the robot, the boy acted aggressively towards it and knocked it over but in later sessions began to act protectively and interacted with the robot as if it were capable of perceiving the emotional valence of its environment and understood spoken language.

Through their experiments with the Keepon, Kozima et. al. learned that an appropriately designed robot can facilitate not only dyadic interaction between a child with ASD and the robot but also triadic interaction and empathetic interactions. Further, it was shown that a very simple robot interface could be used to attract and maintain the attention of children with ASD and facilitate social interaction.

Adaptive Robot Design

Research conducted at the University of Southern California explored the use of a robot whose actions are contingent on user actions to determine the effect on social interaction in children with ASD [36]. Results obtained from the Bubblebot research show that human-robot and human-human social interaction is increased with a robot that responds in a predictable way to user commands. Bubble-blowing games are a common technique for diagnosing children with autism since they tend to provoke social behavior including joint attention and pointing. The interesting aspect of the Bubblebot research is that two distinct robot modes are tested to evaluate whether the actions of the robot effect the behavior of the children participating in the study. In the first mode, the robot would blow bubbles randomly and in the second mode, the robot would only blow bubbles when a large button on the robot's body is pushed.

Five participants (4 with ASD and 1 typically developing) were included in the preliminary pilot study and ranged from approximately 1.5-12 years of age. Quantitative measurements such as the number of social behaviors exhibited by the children in the study and qualitative observations such as the type of behavior (human-human, human-robot) were collected during each of the play modes. During the trial, video recordings were annotated to identify speech, gestures, movement and physical contact in addition to the target(s) of the behavior (the robot or a co-present parent) and whether the event was proactive or in response to the parent or robot.

Results show that all the selected social behaviors measured by this study increased when the robot operated in contingent mode. Total speech increased from 39.4 to 48.4 utterances. Total robot interactions increased from about 43.4 to 55.3 and total directed interactions (those which were clearly directed toward the robot or parent) increased from approximately 62.7 to 89.5. The definitive increases in social interactions presented by this study provide a compelling case for the design and use of adaptive robots which are in some way responsive to the user.

Interactive Game Design

More recently, research in the area of robotics for children with special needs has yielded a comprehensive study by the IROMEC project [37] which describes the types of robot technologies and play scenarios most effective for children with various disabilities, how robots can be best used in therapeutic or educational settings, as well as detailed accounts involving the use of robots used for play activities and possible play-based methodologies. The IROMEC project [38] identifies three play scenarios and five distinct developmental areas most beneficial for collaborative, interactive play with children with ASD [39]. The testing protocol developed for the introduction and use of CHARLIE as a play tool for children with autism is based on the guidelines detailed in the IROMEC study.

CHARLIE incorporates key characteristics from each of the above studies. The toylike, non-humanoid appearance of the Keepon and the user-directed modality of the Bubblebot were used as the basis for the development of the robot architecture and the three types of play scenarios identified in the IROMEC study, (1) turn-taking, (2) sensory reward and (3) imitation were used to design the games detailed in this paper. One of the unique contributions made by this research is the low-cost robot design and additional functionality provided by the hand classifier. With hand detection, the robot is not only able to participate in qualitatively different interactive games but it also allows the robot to collect pertinent information regarding a child's specific progress that may be difficult or impossible to obtain otherwise.

2.4 Remote Stress Detection

Studies related to the remote collection and use of physiological information have been published across multiple disciplines including computer vision [40], image and signal processing [41], human-computer interaction [42], biomedical engineering [43], plant sci-

ence [44] and robotics [45].

Traditional approaches rely on devices which capture changes in air temperature, the circumference of the chest or abdomen, or the sound created by breathing events. Thermistors measure the air temperature near the nasal region during inhalation and exhalation to detect breathing events [46] while respiratory belt transducers rely on changes in the circumference of the chest or abdomen to capture the breathing cycle [47]. A third approach uses battery-powered wearable sensors to detect the sound created by turbulence occurring in the human respiratory system [48]. In addition to being impractical for use in many real-world scenarios, these devices are generally uncomfortable or impractical to wear. Respiratory belt transducers can sometimes even interfere with the breathing process. None of these options are suitable for mobile applications or for people who are sensitive or disinclined to wearing sensors of any kind.

The detection and tracking of stress in humans has become a major focus of current research in robotics and promises to lead to many exciting breakthroughs. The ability to detect shifts in human physiological and/or physical patterns has numerous applications including airport security, military reconnaissance, autonomous home robots, entertainment, and a wide array of medical diagnoses and therapies. Two branches of robotics currently exploring this field of study are Human-Robot Interaction and Socially Assistive Robotics.

Socially Assistive Robotics (SAR) is a relatively new field in robotics which focuses on designing robots to assist people through social interaction while Human-Robot Interaction (HRI) focuses on algorithms which promote more natural and effective communication between people and robots. Both fields require a multidisciplinary approach drawing on the expertise of computer scientists, psychologists and cognitive scientists. In addition, both fields share research with robots whose actions are determined by the emotional state or mindset of the user with whom they are interacting. Due to recent advancements of technology in this area and the collaboration of experts across multiple disciplines, both fields are poised to make substantial contributions in the very near future. It is quite feasible

that stress-sensing robots will be a pervasive, cross-culture technology useful in many areas of life.

Stress Detection Techniques

One of the earliest examples of stress detection is the polygraph. In 1908, an English doctor named James McKenzie designed a device called the “ink polygraph” that measured fluctuations in pulse and blood pressure. It was not until several years later that this particular technology was used to determine whether a subject was being deceptive. Over the years, the device has remained essentially the same, adding the measurement of breathing and perspiration to pulse and blood pressure. In terms of detecting deceptiveness, the reliability of the polygraph has long been a source of controversy, however, its ability to detect physiological changes that are known indicators of stress is undisputed.

One major limitation of the polygraph is the requirement that a subject be “hooked up” to a number of measurement devices. In controlled environments, collecting physiological information from a stationary subject may be effective and useful, but in the majority of real-world applications, subjects are mobile and this mode of data collection is not practical. Since the advent of the polygraph, a variety of sensor/receptor designs have been developed which employ the use of specialized sensors placed on a subject and physiological information is wirelessly transmitted back to a central computer. The application of wireless technology has yielded many new possibilities in general, and specifically, in the field of robotics.

A robotics team at the University of Calgary presented a poster at the Human-Robot Interaction 2009 Conference in LaJolla, California [49], which uses a reasonably-priced commercially-available device to control an iRobot Roomba. In the first phase of their experiment, they tried to issue direct motion commands to the iRoomba, which yielded unreliable results. During the second experiment, they used the biofeedback signals from an OCZ NIA neural impulse actuator (typically used for video games) to influence the

iRobot's movements.

The OCZ NIA is basically a headband that "reads bioelectric signals that are amplified, digitized and further de-convoluted into computer commands." The "de-convolution" is used in video game applications where sensor readings are mapped to specific keystrokes used in various computer games. For the purposes of their experiment, the Calgary team customized this convolution to control the iRobot Roomba through its API (application programming interface). Since the most reliable information collected by the OCZ NIA was muscle tension, the team focused solely on interpreting muscle tension readings which it used to infer a person's emotional state. Instead of directly controlling the robot's actions using the level of muscle tension, the emotional state is estimated from muscle tension readings and that state is used to influence the robot's behavior.

Every five seconds, muscle tension readings were averaged and mapped onto one of four stress levels. The higher the muscle tension, the higher the stress level inferred. When the user is experiencing high levels of stress (levels 3-4), the robot's corresponding action is to enter cleaning mode and avoid the user. When the stress level is low (level 1), the robot will approach the user and stop, behaving as a pet would. Saulnier and his colleagues concluded that crude stress-level readings can be a useful tool for influencing a robot's behavior. In addition, although trying to directly control robot actions produces unreliable results, using inferred emotional states based on stress levels to influence corresponding robotic behavior can be implemented in a fairly simple, straightforward way.

In another experiment, with the support of a grant from the National Science Foundation and the NASA Institute for Advanced Concepts, a team of researchers at Vanderbilt University applied the fundamental concept of human-stress detection to the study of autism therapy [50]. The basic objective of the Vanderbilt team was to "teach" a robot to recognize physiological indicators of stress and determine when to respond with help.

The principal investigator, Nilanjan Sarkar, enlisted the help of a psychologist in order to design stress tests that would be used to produce physiological symptoms of stress.

The tests included playing video games, anagram word puzzles and solving mathematical equations. The team fitted subjects with several (wearable), biofeedback sensors which measured heart rate variability, skin conductivity, eyebrow movement, jaw clenching, and body temperature. The sensors then relayed the information from the subject to a computer through a cell phone-sized data acquisition box which is also worn by the subject. To complete the circuit, the computer was used to communicate wirelessly with the robot.

While biofeedback sensors continuously collected physiological data, the robot monitored the anxiety level of the subject using wavelet signal processing to analyze the sensory data. Results from the data analysis were then used to develop indices correlating to a person's anxiety level. These indices form the basis for building the robot's control architecture, defining for the robot some threshold anxiety level which initiates a state change (i.e, respond and help.)

The key to this approach includes designing an affective control architecture and creating rules by which the robot decides how to respond when the threshold anxiety level is reached. In preliminary studies, Sarkar and his colleagues created rules that directed the robot to simply offer assistance when a specific level of stress registered. Later, however, the robot learned to make choices between response options such as protecting itself, moving toward a human to offer assistance, or sounding an alarm. According to Sarkar, the action plan executed by the robot may require the robot to change its level of autonomy or simply adjust the priority of tasks within the same autonomy level.

The modes of collecting physiological data described by Sarkar et. al. have been effective but have limitations similar to those presented by the polygraph. Both methods still require that the subject be fitted with the proper biofeedback sensors. In certain controlled settings and with certain subjects, this may not be an issue. However, its efficacy in many real-world settings is still somewhat limited. For example, broad applications such as stress-sensing technology in airports, military reconnaissance and other dynamic environments where subjects to be studied cannot be fitted with biofeedback sensors, cannot be

effective without some way to conduct stress-sensing remotely. Another major limitation is with persons (namely children and children with autism, in particular) who are averse to wearing sensors of any kind.

According to a study conducted in 1994 persons with autism may be particularly prone to stress [51]. From an insightful book by Olga Bogdashina [52] “Many a time autistic individuals have been ’pushed’ beyond their limits of sensory endurance. Often this is due to those relating to them not having understood how ’painful’ it is to be overloaded by too much sound; visual stimulation; emotional and/or physical demand and environmental expectation.” For these individuals, effective treatment requires the reliable detection of stress and the minimization (or elimination) of stressors like potentially invasive therapies that require wearing sensors of any kind.

Another study conducted at the Washington University in St. Louis set out to explore the remote detection of stress through the use of Laser Doppler Vibrometry (LDV) [53]. The research team included a professor of psychiatry, university experts in the fields of computer vision and psychology, and researchers in the photonics group at the Boeing Company. The project goal seeks to use LDV to obtain useful physiological information by directing a laser beam at exposed skin.

LDV devices have traditionally been used for the inspection of mechanical components, structural dynamics and even for eardrum diagnostics and detecting insect communication. Where LDVs have shown great promise is in the detection of landmines. Sound is introduced to the area to be inspected (through the use of a loudspeaker, for example), and ground vibrations are measured. If there is a landmine, the area above it will produce enhanced ground velocity at the resonance frequency of the mine-soil system. LDV technology is widely used in industrial and military applications, but their use in studying human physiology has not been widely explored. Technology like the LDV device, when customized, could feasibly close the gap between machines and humans, where affective computing meets human-robot interaction in the most natural way possible.

Experiments were carried out at Washington University, where an LDV device was aimed at the general area on the neck which overlies the carotid artery. Skin vibrations resulting from the pulsating artery were measured for approximately five-minute intervals, three times over a period of several months. Each time, the LDV signal was downsampled to 1kHz and the raw data was extracted. The raw signal easily identified the cardiovascular spike which corresponds to the same R wave in an electrocardiogram (ECG.)

Some of the drawbacks posed by using LDV for physiological data collection are variances in physiology and tracking. It is still unclear how much natural human movement can be tolerated before adversely affecting the accuracy of measurements taken by the LDV. Additionally, it is not known if these variances in tracking can be overcome and whether the differences in physiology can be normalized to a specific population. More research is required to determine the limitations of this type of technology in order to ascertain its potential role in remote human stress-sensing in general, and more specifically, its usefulness in autism therapy.

Stress-Detecting Robots

While the collection of physiological data for diagnosing disorders and stress in humans is not new, remotely recovering this information for use in robotics is an emerging field. Recently, a remote-controlled robot was developed that is capable of detecting motion and breathing though building walls using millimeter-wave miniaturized radars [54]. Although such systems have high utility for search and rescue, military and law enforcement applications, they are not suitable for most HRI scenarios because of their size, cost and the fact that they are non-autonomous. Other important research contributions in robotics using contact sensors have shown that physiological responses alone can be used to successfully recognize anxiety in humans [55].

A study using a reasonably-priced, commercially-available device to control an iRobot Roomba [49] obtains biofeedback signals from an OCZ NIA neural impulse actuator (typ-

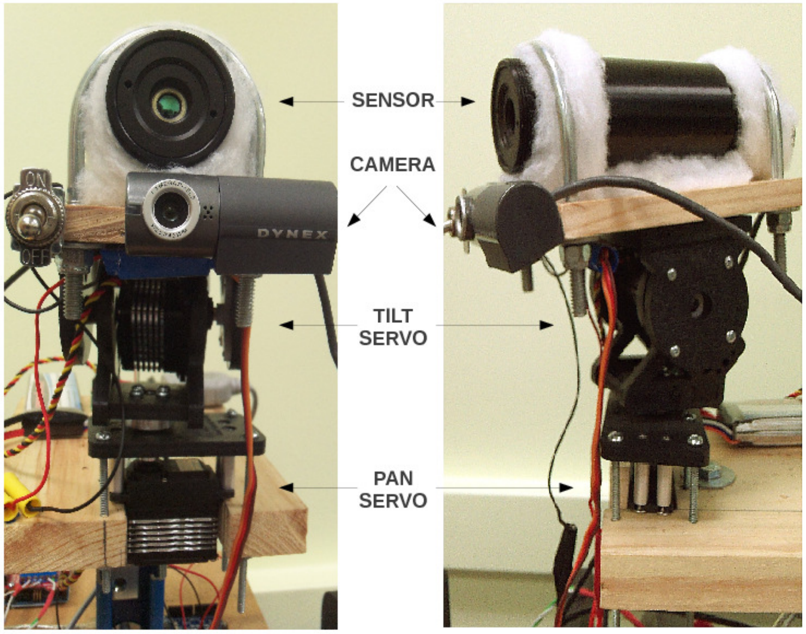

Figure 2.1: Remote breathing monitoring system. Front view (left) and profile view (right).

ically used for video games) and deduces a user stress state to adapt the Roomba's movements accordingly. The OCZ NIA is basically a headband that collects bioelectric signals, amplifies, digitizes and further de-convolutes the signals to convert into computer commands. For the purposes of their experiment, researchers customized this convolution to control the iRobot Roomba. Since the most reliable information collected by the OCZ NIA was muscle tension, the team focused solely on interpreting muscle tension readings which it used to infer a person's emotional state. Instead of directly controlling the robot's actions using the level of muscle tension, one of four stress states is estimated from muscle tension readings and that state is used to influence the robot's behavior. The higher the muscle tension, the higher the stress level inferred. When the user's stress state was perceived to be high, the robot would enter cleaning mode and avoid the user. When their stress level was perceived to be low the robot will approach the user and stop, behaving as a pet would. An important contribution of this paper was the finding that crude stress-level readings can be a useful tool for influencing a robot's behavior.

Another study applies the fundamental concept of human stress detection to the study of autism therapy [50]. The basic objective was to "teach" a robot to recognize physiological indicators of stress and determine when to respond with help. Stress tests were designed to produce physiological symptoms of stress and included activities such as playing video games, solving anagram word puzzles and mathematical equations. Participants of the study were fitted with biofeedback sensors which measured heart rate variability, skin conductivity, eyebrow movement, jaw clenching, and body temperature. The sensors then relayed the information from the subject to a computer which communicated wirelessly with the robot.

While biofeedback sensors continuously collected physiological data, the robot monitored the anxiety level of the subject using wavelet signal processing to analyze the sensory data and indices correlating to a subject's anxiety level were developed. These indices form the basis for building the robot's control architecture, defining for the robot some threshold anxiety level which initiates a state change. The key to this approach includes designing an affective control architecture and creating rules by which the robot decides how to respond when the threshold anxiety level is reached. Ultimately, the robot learned to make choices between response options such as protecting itself, moving toward a human to offer assistance, or sounding an alarm.

The modes of collecting physiological data described in these studies have been effective but each method still requires that the subject be fitted with the proper biofeedback sensors. In certain controlled settings and with certain subjects, this may not be an issue. However, their efficacy in dynamic environments where people cannot be fitted with biofeedback sensors or in medical or therapeutic settings where persons are averse to wearing sensors is still somewhat limited.

Remote Stress Detection

Remote breathing detection can be used in many applications when it is important to measure changes in breathing rate but it is not practical to attach sensors or receive frequent feedback from the user. One of the first published works which measures breathing rate remotely uses an active radar detector to measure movements of the chest caused by cardiac and breathing events [56]. Since then, other non-contact modalities have been explored including laser doppler vibrometry (LDV) and mid-wave infrared video cameras. The LDV study remotely collected physiological information to deduce the stress state of an individual based on vibrations of the skin directly covering the carotid artery [53]. The objective of the study was to obtain useful physiological information such as that provided in an electrocardiogram (ECG) by directing a laser beam at exposed skin. However, two significant limitations posed by this approach include the challenge of tracking accurately caused by variances in physiology and the prohibitive cost of the equipment used.

The biomedical engineering field has published a great deal of research dedicated to the acquisition of a wide variety of physiological information. One recent study uses a mid-wave infrared camera to capture breathing rate based on air temperature changes near the nasal region [57]. This particular implementation was initially designed for polysomnography, or sleep studies, and relies initially on the manual identification of a primary region of interest, tracking the location of the nostrils and, more specifically, the outer extent of the nostril region. Because of the large amount of image and data processing required and the small size of the nostril location to be tracked, segmentation becomes challenging and computationally expensive.

Medical applications including polysomnography and the diagnosis and management of respiratory diseases may require a high level of precision that demands the collection and analysis of data relating to the entire breath waveform not just the number of breathing events. Our research uses a simpler, less expensive method to monitor *changes* in breathing and heart rate. It does not require capturing a large amount of precise data relating to the

full breath spectrum.

CHAPTER 3

INTERACTIVE ROBOT

It is widely known that the incidence of autism has been increasing over the last decade, with some reports citing a 57% increase in autism prevalence between 2002 and 2006 [7]. Two of the most significant problems stemming from the increased prevalence of autism are the additional strain placed on existing resources for treating children with ASD and the additional financial strain placed on families who care and seek treatment for their children. The costs associated with additional therapy, specialized and medical care for a child with ASD are estimated to be approximately 8.5 to 9.5 times more than raising a typically developing child [58]. For some families, this additional financial burden may mean having to choose between incurring significant debt in order to get the proper care for their child(ren) or limiting the amount of therapy their child receives. Although several existing robots have been used with children with ASD, they are still generally cost prohibitive for widespread use by special education instructors and therapists.

In response to these existing needs, the long term vision of our research is to produce a low-cost, adaptable robot which is widely accessible to a large population of autism therapists, teachers and parents for use as part of an overall early intervention strategy for children with ASD. Because the social and communication skills of children with autism vary as much as their individual preferences, one robot design will not be universally accepted or effective. However, while a one-size-fits-all approach is not appropriate, simple robot and game designs can be achieved which target a specific set of communication tasks or social skills. In addition to focusing on the careful design and development of hardware and software that is known to be effective for children with ASD, we paid special attention

Figure 3.1: CHARLIE. [left] Completed robot. [right] Internal structure.

to developing an appropriate testing protocol.

3.1 Methodology and Approach

The physical design of CHARLIE addresses three major objectives. First, the outward appearance of the robot was designed to be toylike and pleasant so as to invite the attention of young children with ASD and avoid being intimidating to the greatest extent possible. Second, we carefully designed the robot structure so as not to allow the robot, nor its constituent parts, to harm the child interacting with it. Third, we made the robot more robust by adding features to protect its mechanical components and allow children to explore and interact more freely with the robot without excessive concern for the physical integrity of the robot.

Our approach to the design of CHARLIE's interactive games is based on the integration of robot and game designs that are known to be effective with children with autism. Each game was designed to be entertaining to young children and to promote two fundamental requirements for communication: imitation and turn-taking. Further, we created three different types of games to appeal to children with varying levels of communication and

social skill. For children who are reluctant to play with a completely autonomous robot or for those who would benefit from a period of exploration before they begin playing the interactive games, the robot can be teleoperated using a simple remote control. For those who are ready to play directly with the robot, but who are not necessarily ready to play a co-operative game with another child, the single-player interactive game is available. Finally, a two-player interactive game was created to appeal to those children who have established some level of simple imitation and turn-taking but could use more practice with these skills using the robot as a social mediator.

3.2 PHYSICAL DESIGN

We deliberately designed the outward appearance of CHARLIE with the end-user in mind. Recent research has shown that robots with a simple interface are generally better received initially by children with autism, than robots with a more realistic, human-like appearance [59]. The implication is that low-tech robots, when designed appropriately for the particular needs of the child(ren) with ASD they will serve and the context in which they will be used, can be used effectively to teach and promote social skills. In addition to the low-cost design, CHARLIE'S physical appearance is intended to be toylike to create a friendly and approachable outward appearance and to more easily attract the attention of a child.

Basic hardware components

CHARLIE's hardware includes 6 servos, 3 pan-tilt platforms, an 8 channel servo controller, a consumer-grade web cam, and 2 D-cell battery packs. The robot's body is padded for safety, and its outer surfaces are covered with a bright green, fur-like material to achieve a non-threatening appearance. During active game play the child's attention is typically focused near CHARLIE's hands, so one LED is embedded in each of the hands to provide positive feedback during interactive games. A speaker is also included in CHARLIE's body

Figure 3.2: [top] Snap off arm. [bottom] Snap off head.

in order to provide optional auditory instructions for playing interactive games and positive feedback. Exclusive of the computing hardware, the retail cost of the robot's components is approximately $200 USD. In a production version of this robot, a computer could be integrated into the robot's body, or users could connect via USB to a standard laptop or desktop PC.

Features for robustness and safety

In general, children are curious about robots and many enjoy exploring the physical features of the robot as much as interacting with it. This can present hazards to both the child and to the robot's mechanical hardware. In order to minimize potential hazards and to improve the robustness of the robot, we included two characteristics in the robot's design. First, the body of the robot is secured to a platform that may be strapped to a desk or table. Immobilizing the robot in this way prevents the child from being able to pick up the robot and potentially harm him/herself, others in the room or the robot itself. Second, the arms and head of the robot are attached to the robot's body using snap fasteners so that excessive force will not cause damage to the servo motors, but will instead allow that piece to snap

off. Furthermore, allowing the arms and head of the robot to detach, affords the child more continuous free play since there will be less concern over the child's safety and the integrity of the robot's hardware. As described in the IROMEC study [37], while the adult must fulfill a more active role for promoting play skills with children with ASD, "much of the literature on childhood play emphasizes the importance of free play and the need to interfere as little as possible in the child's actions, thus underscoring the creative aspects that in essence cannot be controlled or oriented." We expect that longer, uninterrupted interactions will maximize the opportunity for each child to benefit from each session.

3.3 Interactive software design

We used the Open Source Computer Vision Library (OpenCV) [11], a cross-platform library for real-time computer vision applications, for training the hand classifier and for the implementation of hand and face detection. OpenCV provides a facility for object detection based on an extended set of Haar-like features [60]. Informally, this method works by screening small portions of an image for visual characteristics of the target object. To train a classifier to identify a specific class of objects, OpenCV uses Adaptive Boosting (AdaBoost) [61] to create a cascade of boosted classifiers defined over these features. We then included the resulting hand classifier along with a standard OpenCV face classifier to detect user hands, track the user's face and provide position information for managing three interactive games. In the first game the robot waits for the child to initiate an interaction by raising one or both hands. In the second game, the robot initiates interactions. The primary objective of our game designs is to increase attention, promote turn-taking skills and encourage child-led verbal and non-verbal communication through simple imitative play.

Figure 3.3: CHARLIE poses. From top to bottom, left to right : Left hand high. Both hands high. Right hand high. Peek-a-boo. Neutral.

Face detection and tracking

We relied upon the frontal face classifier provided by OpenCV (more specifically, a cascade of boosted classifiers working with Haar-like features) for face detection. Haar-like features are used as an abstraction of RGB pixel values for object detection since image intensities are computationally expensive to work with. Each feature type is used to screen a given portion of an image for different characteristics of the target object. The extended sets of rectangular Haar-like features used for the face and hand detectors described in this paper are applied to assess whether a particular rectangular portion of a video frame contains a

Figure 3.4: Face and hand detection.

face or hand by summing the pixels contained within the rectangle and determining whether it matches the characteristics of the target object as defined by the classifier.

To make the overall program as efficient as possible, we implemented a face tracking algorithm instead of repeating the computationally intensive detection process for each frame. Face tracking was accomplished using the Continuously Adaptive Mean Shift (CAMSHIFT) algorithm [62]. CAMSHIFT incorporates the MEANSHIFT algorithm which is based on a nonparametric technique for climbing density gradients to find the peak of the probability distribution of the position of a given target object. For face tracking, this translates to identifying the center of the target color distribution in a given video frame. In order to make face tracking fast and relatively robust (and appropriate for use in real-time tracking applications), we used the CAMSHIFT technique. This tracking method improves performance by eliminating the need to repeat the face detection for each frame of the video. To overcome errors resulting from drift in the CAMSHIFT algorithm, the robot periodically repeats the full face detection process. In the event that the robot cannot detect the face, the robot head is reset to a neutral position and searches outward in an increasingly larger area.

Figure 3.5: Sample images used to train the hand detector. [top] Positive examples. [bottom] Negative examples.

Whereas face detection is a well-studied problem [63, 64], and effective face classifiers are freely available through OpenCV, robust and real-time hand detection in diverse environments is a topic of continuing research.

Hand classifier and hand detection

Numerous approaches for developing robust hand detectors have been explored [65, 66], but the resulting classifiers have not been made available to the research community. Further, some hand classifiers that are freely available such as the gesture letter "A" detector by Juan Wachs from the Ben Gurion University of the Negev, Israel and Washington Hospital Center [67], are too narrow in scope for use in this context and others are not accurate or efficient enough for our application. In order to implement a hand detector suitable for our purposes, we trained a new hand classifier to detect hands in various lighting conditions, rotations, scales and finger positions. Approximately 750 positive hand images of various size, color and position and approximately 3300 negative images were collected and cropped to a uniform pixel size of 40x40. Representative examples are shown in Figure 3.5.

To create additional positive training samples representing variations in lighting, rotation and scale, ten distortions were applied to 100 of those samples, yielding a total of approximately 1750 positive hand samples. The resulting vector files were then merged and the AdaBoost training procedure was initiated using the combined vector file representing all positive hand samples and the complete set of negative samples.

Interactive game design

Research in robot-assisted autism therapy typically emphasizes specific objectives for ideal human-robot interaction including an increased attention span, eye contact, proactive interaction with the robot initiated by the child, verbal and non-verbal cues, turn-taking, imitative game playing and overall use of language.

First, we defined the play scenario in terms of: (1) a main target group, (2) a play type, (3) actors involved, (4) a setting, and (5) the duration of the play activity. The main target group consists of a small group of children ages 4 to 11 who have been diagnosed with autism and have documented communication deficiencies. The play type consists of a very simple game of imitation with a basic set of rules and is designed to engage one teacher or one child at a time. The tests take place in a closed classroom, where both the child and teacher are seated across from the robot and the robot will be seated atop and securely attached to a nearby desk so that the robot's head is at approximately the same height as the child's. The duration of the play activity is variable. The length of a typical session with the robot is based on the normal amount of session time allotted for that particular child, the perceived benefit of the robot to the child's development and the child's interest in the robot.

Second, we prepared a detailed description of how CHARLIE is introduced to each child and how play proceeds during the first and subsequent sessions. Prior to introduction, a baseline for communication skills and developmental ability is established for each child using assessment information provided by the child's teacher. At the first meeting, the

teacher introduces CHARLIE and explains and/or demonstrates how to play the imitation game. The teacher then invites the child to play with robot and provide guidance, when necessary. For children who prefer to examine the robot and learn about its capabilities independently, the teacher assumes a more passive role, as an observer and guide.

Third, we identified measures of success using the baseline communication skills identified prior to the child's first session. Initially, the child's level of interest in CHARLIE is noted in addition to any specific robot characteristics that are especially interesting to the child. During each session, communication between the child and robot, and the child and teacher is documented by the teacher or researcher (the author). Because the robot measures successful imitations between the robot and child it is not necessary to document these interactions, but other nonverbal and verbal communication occurring during the session is noted for subsequent analysis. Measures of success and user information collected during an interactive game can be used to assess the child's readiness for more advanced, child-initiated games such as collaborative group play and story-telling.

Ultimately, we designed and implemented two additional interactive games to appeal to children with ASD of a wider range of ability and skill. The original game developed is a single-player game which engages a child in a game called "Imitate Me, Imitate You". In this game, the child may either initiate a pose for the robot to imitate ("Imitate Me") or the child may follow the robot's pose ("Imitate You"). The single-player game is intended for the child with ASD who is comfortable interacting with an autonomous robot but who may not be ready for turn-taking with another child.

Single-player "Imitate Me, Imitate You"

The "Imitate Me, Imitate You" game is detailed in Figure 3.6 and consists of two primary modes: passive and active. Within each of the two modes, there are five poses: neutral (both hands down), left hand raised, right hand raised, both hands raised and peek-a-boo, as shown in Figure 3.3. In order to give the child initial control over the robot's actions, the

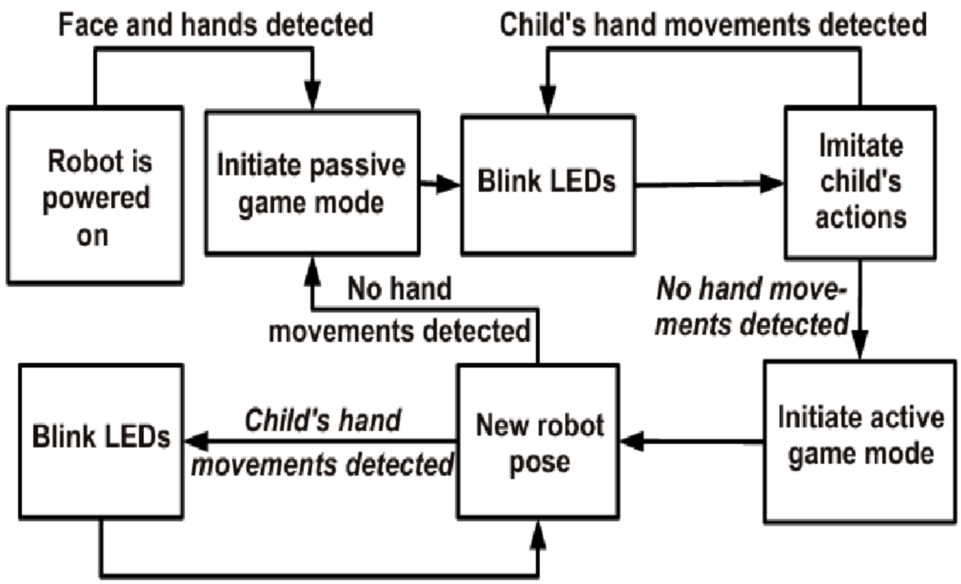

Figure 3.6: State diagram for CHARLIE's "Imitate Me, Imitate You" autonomous interactive game.

default robot state is the passive game mode. Once the robot detects and begins tracking the child's face and hands, the robot indicates that it is ready to interact by moving to the neutral pose and blinking the LEDs in its hands three times. The robot then immediately enters the passive game mode and waits for the child to initiate a game by raising one or both hands. As the child's hand movements are detected, the robot responds by imitating the child's hand positions and lighting the LED in the corresponding hand while simultaneously detecting any additional hand movements. If ten seconds elapse without any detected hand movement, the robot will transition to the active game mode.

During the active game mode, the robot initiates a new game and attempts to engage the child by raising or lowering one or both arms, or beginning a game of peek-a-boo. Each pose assumed by the robot in the active game state is selected randomly in order to avoid repetitive patterns of poses. When a positive outcome is detected (the child successfully imitates the robot's pose), positive sensory feedback is generated by the robot. A positive sensory response entails the robot lighting a small LED in the hand corresponding to the

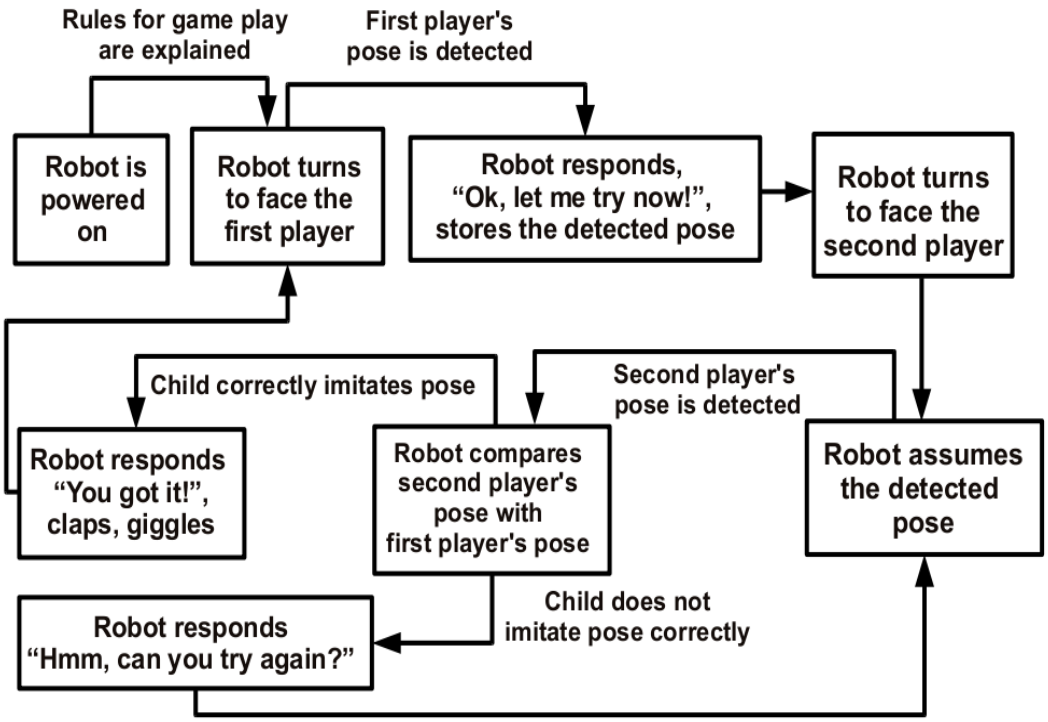

Figure 3.7: State diagram for CHARLIE's "Pass the Pose" autonomous interactive game.

raised hand or hands of the imitated pose. As with the passive game mode, the robot will wait ten seconds for the child's response. If ten seconds elapse and a positive response has not been detected, the robot will transition back to the passive game mode, waiting again for the child to initiate a new game.

Two-player "Pass the Pose"

The second interactive game is a two-player game described in Figure 3.7 called "Pass the Pose". In this game, two players interact directly with the robot and indirectly with one another. With the optional sound enabled, the "Pass the Pose" game works as follows: Game play begins with CHARLIE describing how to play "Pass the Pose" and asking the first player (seated to the right of the robot) to assume a pose. Once she has detected the pose, CHARLIE indicates that she has learned the pose by saying "Ok, I got it. Now let

me try”, turns to the second player (seated to the left of the robot), asks the child to follow her and then assumes the same pose learned from the first player. If the second player successfully imitates the pose assumed by CHARLIE, she responds by saying “You got it!”, claps her hands and giggles. If the player does not immediately imitate the correct pose, CHARLIE will ask the child to try again. If the child does not correctly assume the pose after three tries, the robot asks the current player to initiate a new pose and the game continues, this time with the second player initially “passing” the pose to the robot.

If the sound is disabled, we expect that the teacher, therapist or parent will describe how to play the “Pass the Pose” game. When the players are ready the teacher will start the game and CHARLIE will turn to the first player and wait for the child to assume a pose. Once CHARLIE has detected the pose, she turns to the second player and assumes the same pose. If the child correctly imitates the pose, CHARLIE claps her hands and waits for the second player to initiate the next pose. If the second player does not correctly imitate the pose, CHARLIE lowers her head and shakes it slowly from side to side. Should the child fail to imitate the pose correctly after three tries, CHARLIE resumes a neutral position and waits for the second player to start a new game. This two player game is ultimately designed to promote shared attention and cooperative play. We anticipate that the “Pass the Pose” game will be most useful for children who have already demonstrated some level of proficiency with turn-taking and imitation and who are able to play a game with a simple set of rules.

Teleoperation

In addition to the two autonomous games, we developed and implemented software that allows for the robot to be teleoperated so that when a button is pushed on the remote, the player is given complete control over CHARLIE’s limbs and head. While each of the four push buttons on the remote correspond to specific pre-programmed poses, the two joystick buttons provide continuous control for the movement of each arm and a single directional

button allows for continuous control of the head. We expect this game play to be useful for the child with ASD who may be initially wary or hesitant to interact with the robot. By temporarily disabling the robot's autonomous actions, the child is given the freedom to learn about CHARLIE's various capabilities at his or her own pace.

3.4 Data collection

There are two distinct kinds of user interaction information collected by the robot. Information pertaining to the user's overall progress such as (1) the total length of active engagement (time spent actively engaging in either passive or active mode), (2) number of child-led actions and (3) the number of successful interactions is continuously captured during each session. At the end of the session, this information is used to create a user progress report for analysis and for future sessions with the same child. The second type of user information, such as the length of the intervals between interactions, is used for controlling the robot state.

3.5 Initial robot experiments

Experimental setup

Face detection, tracking and hand detection results

As a proof-of-concept for CHARLIE's effectiveness, we conducted preliminary tests using the single-player game with a small group of typically developing children. See Figure 3.8. A relatively large age range (4-11 years) was selected to test the reaction times of the robot when used with children of varying levels of ability. Each child participated in an 8-10 minute session, in which both game modes (passive and active) were tested and the accuracy of the hand and face detectors was measured. The duration of each game mode was recorded to ensure that adequate time is given for the child to respond before a transition

Figure 3.8: Children Interacting with CHARLIE.

is made to the alternate game mode and the effectiveness of the positive sensory feedback (LEDs in hands indicating successful detection) was assessed.

Experiments were conducted to measure the speed and accuracy of the face and hand detector and to assess the appropriateness of CHARLIE's timed responses during game play (Table 3.1). The accuracy of the face detector and tracker was determined by calculating the ratio of successful face detection time to the total session time. The face detector averaged an accuracy of 86% across all sessions and users. In a typical session, users averaged 33 child-initiated hand movements and imitated 16 robot movements per minute. The hand detector accurately detected the child's hands an average of 92% of the total session time, with 244 hits out of 265 total hand events.

While face detector accuracy includes the time during which a false positive or detection failure occurs, it also counts as misses the aggregate time during which the child's face

Table 3.1: Data collected from an interactive session with CHARLIE.

Participant	Child1	Child2	Child3	Child4	**Average**
Age (years)	8	8	4	11	7.75
Interaction time	152s	198s	156s	144s	162s
Lost face time	22s	30s	36s	18s	26s
Face detection hit rate	**87.0%**	**87.0%**	**81.0%**	**89.0%**	**86.0%**
Passive time	30s (19.7%)	124s (62.6%)	89s (57.0%)	118s (82.0%)	90s
Active time	122s (80.3%)	74s (37.4%)	67s (43.0%)	26s (18.0%)	72s
Actual passive hand actions	29	48	37	84	50
Passive hand detections	24	41	35	81	45
Passive hand hits	**83.0%**	**85.0%**	**95.0%**	**96.0%**	**90.0%**
Actual active hand actions	19	39	4	5	17
Active hand detections	17	38	4	4	16
Active hand hits	**89.0%**	**97.0%**	**100.0%**	**80.0%**	**92.0%**

is actually not within the camera's field of view. Since an absent face is not possible to detect, the observed accuracy for face detection may be artificially low. Conversely, the accuracy of the hand detector may be artificially high since it is only based on the number of hand events successfully detected and not on the actual time required to detect them. A twenty-stage cascade was trained on these samples, yielding an error rate on the training set approaching zero. Section 3.5 presents a quantitative evaluation of the classifier performance.

Nearly all of the children expressed a preference for the passive game mode, where the robot imitates the child's hand actions, and their comments were supported by the significantly greater amount of time each of those children spent in the passive mode compared to the active mode during their respective sessions. These preliminary results as an important proof-of-concept in preparation for controlled tests with children with ASD.

3.6 Field Study with CHARLIE and Eight Children with Autism

Experiment Objectives

Recently, the robotics community has seen a rapid acceleration of research in the development and testing of socially assistive robots (SAR) for children with ASD [27, 28, 68]. A number of qualitative studies describe the therapeutic benefits of using interactive robots in therapy such as increased speech, social interaction, joint and directed attention [6, 33, 36], but few studies exist which quantify observed communication increases using assessment instruments accepted by the autism and speech therapy communities. Even fewer provide a statistical evaluation comparing the benefits received through speech therapy with those obtained through an additional robot intervention. Further, no field studies have employed a robot prototype that is sufficiently robust and reasonably easy to operate for cooperative use by the therapist in the clinic, the family at home, and the special education teacher at school to deliver generalizable results.

We designed and conducted a pilot field study to achieve three primary objectives: (1) to quantitatively assess the effectiveness of a robust, interactive robot named CHARLIE for increasing spontaneous speech, overall communication and social skills in children with autism, (2) to compare communication and social skills increases obtained through therapy(ies) supplemented with a robot-assisted intervention to increases achieved without robot intervention and, (3) to explore a new interactive, human-in-the-loop therapeutic methodology for generalizing learned social behaviors that extend child-to-robot interactions to child-to-co-present-other exchanges.

Institutional Review Board approval was sought to conduct this study and obtained on February 19, 2013 under study identification number Pro00023119, and the study title: "Effectiveness of CHARLIE the Robot for Improving Verbal and Nonverbal Skills in

Children with Autism." The approved study protocol is included in Appendix A.

Target Population

We recruited a study group of eight children, between 3 and 6 years of age, who have been diagnosed with autism and a speech deficiency as confirmed by the Autism Diagnostic Observation Schedule (ADOS) and a speech pathologist, respectively. Children were invited to participate in the study based on the information provided in the Prescreening Questionnaire (Appendix B) and confirmation of autism and speech deficiency diagnoses.

Recruitment Procedures and Experiment Location(s)

Flyers were provided to local autism support groups, physicians and clinicians and posted at local area pediatricians' offices, speech therapy clinics, diagnostic clinics and nearby elementary schools (see Appendix C). Prior approval was sought from each organization or place of business before flyers were posted or distributed. Additionally, in order to recruit as many participants as possible, participant enrollment was conducted on a rolling basis between April and October 2013.

To determine whether each child was a good candidate for the study, the Prescreening Questionnaire was administered to the prospective participant's caregiver over the phone. The questions on the prescreening form were used to confirm that: (1) a formal diagnosis of autism was received (also, who made the diagnosis and when it was made), (2) the child's language ability was delayed for their chronological age (and to what degree), (3) the child's nonverbal communication ability (pointing, shaking/nodding head, etc) was delayed for their chronological age (and to what degree), (4) the child did not have any diagnosed hearing impairment and, (5) the child's current therapy schedule would allow for additional therapy sessions with the robot. Most of the children in the study were concurrently receiving other forms of intervention, but 3 of the 8 children were not receiving

any other intervention at the time of their participation in the study.

If the child did not meet the requirements to participate in the study, the information collected over the phone during the prescreening process was destroyed and the child was not invited to participate in the study. If the child was determined to be a good candidate, a numeric identifier was assigned to each child and the completed Prescreening Questionnaire was stored in a secure office, in a locked drawer. Copies of any prior, formal diagnoses of autism or speech and language delays by a licensed professional were requested upon selection for the study. Upon receipt, these documents were also secured in a locked office cabinet inside a locked office. All identifiers were removed from the data and replaced with a participant number. A password-protected file matching participant names with participant numbers was stored on a password-protected laptop computer.

Once all required documents were obtained and the child was invited to participate in the study, the first face-to-face meeting was scheduled at the USC School of Medicine's Department of Neuropsychiatry and Behavioral Sciences located at 3555 Harden Street Extension, Suite 301, Columbia, SC 29201. The second meeting and subsequent sessions with CHARLIE, the researcher and the speech pathologist took place at the USC Speech and Hearing Research Center located at 1601 Saint Julian Place, Columbia, SC 29204.

Obstacles in recruitment and retention

- **Excessive absences**

  Participants were allotted a total of three allowable misses during the entire six-week intervention. If a participant exceeded three absences in the six-week study period, that participant's scores for the duration of the study were not considered accurate since s/he would have missed more than 25% of the total session time and reasonable consistency between one subject and the next could not be assured. Participants 5

and 6 missed four and five out of the 12 sessions, respectively, and were consequently excluded from the study results.

- **Limited clinic and family availability**

  The second diagnostic meeting and all sessions with CHARLIE were conducted at the USC Speech and Hearing Research Center. Clinic space and participant availability significantly limited the number of children that could concurrently receive therapy during each six-week period. The center has a limited number of therapy rooms, and maintains a schedule of clients who regularly receive speech therapy each week. Additionally, many of the children who were invited to participate in the study were enrolled in school at the time of the study and were not available during daytime hours. Finally, many of the participants' parents worked during the day. To accommodate study participants, their caregivers and the USC Speech and Hearing Center schedule as much as possible, sessions with CHARLIE were scheduled two days a week, in the evenings.

- **Inability to complete study**

  An additional obstacle to retention was the loss of interest in or inability to complete the study. Two prospective participants completed the prescreening quetionnaire but did not follow through by attending the first meeting at the USC Department of Neuropsychiatry and Behavioral Sciences. Four participants completed the prescreening questionnaire and signed the informed consent, but did not continue in the study by attending the first diagnostic meeting at the Speech and Hearing Research Center. Several prospective participants were already receiving multiple therapies at the time of their invitation to the study and it is quite possible that their caretakers chose not to participate due to their already very demanding schedules (especially given that

some of them also had young siblings.)

Finally, Participant 1 completed the 12 sessions with CHARLIE, but did not attend the final evaluation due to the child and his family traveling out of the country for two months before final session data could be collected. Since the final VABS-II and the final MLU were both missing, statistical analyses for this study did not included data collected from Participant 1.

Experiment Protocol

First Meeting The first meeting with the caregiver(s) of each child participating in the study included the completion and signing of the informed consent, signing of releases for medical records form(s) documenting a diagnosis of autism and a speech impairment, completion of the Vineland Adaptive Behavior Scale II (VABS-II) [69], and the Social Communication Questionnaire (SCQ) [70]. To confirm the diagnosis of autism, the results from the Autism Diagnositic Observation Schedule (ADOS) [71] for children (if available) was requested.

The VABS-II is designed to measure personal, communication and social skills especially for special needs populations such as individuals with mental retardation, autism, and attention-deficit/hyperactivity disorder (ADHD). It is comprised of four domains: Communication, Daily Living Skills, Socialization, and Motor Skills. The Communication Domain measures receptive, expressive, and written communication; the Daily Living Skills Domain assesses personal, domestic, and community skills; the Socialization Domain measures interpersonal relationships, play and leisure time, and coping skills; and the Motor Skills Domain measures gross and fine motor skills. Scores from the Communication and Socialization Domains in the VABS-II were used to establish a basal score for each participant and stored for later comparison with scores recorded at the end of the study.

The SCQ is a brief questionnaire which aids in the evaluation of communication skills and social functioning in children who may have autism or autism spectrum disorders [72]. Typically, the SCQ is a fast way to determine if an individual should be referred to a qualified professional for a complete diagnostic evaluation. The SCQ was used in this study as an additional measure to confirm the child's autism diagnosis. A cutoff score of 15 or greater was used as an indication of possible ASD; any participant scoring below 15 on the SCQ was not invited to participate in the study.

The duration of the first meeting was approximately one hour.

Consent and Risk Parents of children with autism identified as potential study participants were asked to read and sign the informed consent at the first meeting. They were provided a copy of the signed Autism Study Consent Form (Appendix D) and a copy of the Study Session Methodology (see Appendix E) which details each step of the study procedure and approach for intervention with the robot. There were very few risks associated with participating in this research except a slight risk of breach of confidentiality.

To assist in documenting each child's progress, sessions were videotaped and catalogued using the date and a unique identifier assigned to the individual child. Video-recorded sessions were highly useful for calculating Mean Length Spontaneous Utterance Determination (MLSUD) [73] measures throughout the six-week study period for each child and for making note of any significant changes in the child's response as new parts of the intervention were introduced. The original informed consent document was revised in order to include language that allows for demonstration of recorded videos for research and educational purposes. The two parents who had signed the original consent form before the revision was made, also agreed to sign the revised consent form and their signatures

were subsequently obtained.

Second Meeting The second meeting took place at the USC Speech and Hearing Research Center where the senior clinical instructor and the researcher conducted three additional screenings. To assess motor imitation ability the Motor Imitation Scale (MIS) [74] and the Unstructured Imitation Assessment (UIA) [75] were administered. The Expressive Vocabulary Test 2 (EVT2) [76] was also administered to assess expressive vocabulary and word retrieval ability. At the conclusion of the second meeting, the video-recorded session was reviewed in order to perform an additional measure of verbal utterances using the MLSUD. The MLSUD provides a total score for spoken meaningful language during the 1.0-1.5 hours assessment period. The MLSUD score is derived by assigning one point for each spoken morpheme divided by the total number of utterances in the session.

Sessions with CHARLIE Following the second meeting, each child received two 30-minute sessions per week for a duration of six weeks, or 12 sessions of intervention with the robot. The room designated for the study sessions included one child-sized table, two child-sized chairs, one or two adult-sized chairs, the robot and several hats and accessories for game play (Figure 3.9). The senior clinical instructor from the USC Speech and Hearing Research Center, Sarah Scarborough, was regularly present (with few exceptions) to provide guidance and intervention expertise throughout each child's therapy. Additionally, the researcher was always present to provide continual monitoring and periodic operation of the robot during each session.

Recognizing that children with autism tend to experience high levels of stress in new situations, we also provided a second room where children could go to jump on a trampoline, roll on a large ball or read a book. The "break" room was used by participants on an

Figure 3.9: Therapy room at the USC Speech and Hearing Research Center.

as-needed basis.

Six phases were initially identified for introducing the child to and engaging the child with the robot. Each phase was designed to address specific therapeutic goals, including increased speech and social skills, for interactions between study personnel, the child and the robot. The original intervention methodology developed for the study is described in Appendix E. However, upon completing sessions with the first two participants, the intervention strategy was revised to achieve improved study outcomes. A detailed discussion of the initial and revised intervention methodologies and the rationale for doing so, is described below.

Approach for Conducting Therapy with CHARLIE

Initial Approach The initial study procedure (Appendix E) was developed over a series of weekly meetings with the researcher, two experts in autism diagnosis and treatment and an expert in speech therapy. Before including activities in the study procedure that promote specific speech and social skills and directly engage a child through interactive play with CHARLIE, the study group agreed that including exercises to facilitate the child's trust of the robot should precede any direct engagement. Therefore, the first two phases detailed in the study procedure describe objectives that encourage the child to manipulate the robot physically and explore controlling CHARLIE's motions through teleoperation. This affords the child the opportunity to observe CHARLIE's range of motion, kinematics and hear the sound(s) of the servo motors in a manner that gives the child control over the robot to the greatest extent possible. The objectives in Phases III-VI focus on promoting foundational skills required for communication and socialization through robot-assisted play.

Once the interventionist deems that a basic level of trust and acceptance has been established through Phases I and II, she will lead the child to begin Phase III which specifically promotes joint attention, imitation and turn-taking. Phase III introduces interactive play between the child and the robot where CHARLIE autonomously plays either "If you're happy and you know it" or "Wheels on the bus" while performing appropriate hand/arm motions. The interventionist and the researcher direct their attention toward CHARLIE, perform the appropriate hand motions and actively encourage the child to imitate as well. If the child indicates a willingness to participate by directing his/her attention and engaging in some hand play with CHARLIE and demonstrates that s/he would like to continue imitative play with the music, the song or songs are played several times (as directed by the child). Should the child indicate that s/he is no longer interested in the music/hand play or the child has already engaged in music play several times, the interventionist offers the remote control and introduces a new game of imitation that builds on the same concepts

played with the music/hand play imitation. In the remote control-based imitation game, the child uses the remote control to move the robot to pose in a particular way and all others in the room play along by imitating the robot's pose.

Phases IV, V and VI focus on introducing two new interactive games for further practicing joint attention, imitation and turn-taking skills. The single-player version of the imitation game, called "Imitate Me, Imitate You", is intended to enable the child to control the robot without a remote. By initiating hand poses and observing CHARLIE imitating the same pose, the child is encouraged to practice fundamental one-to-one turn-taking and imitation skills through the cause and effect nature of the game. The two-player imitation game called "Pass the Pose" requires a more advanced understanding of communication and social interaction. Successful game play in this mode requires that the child is already mastering turn-taking and imitation in a one-to-one scenario and is ready to practice turn-taking and imitation in a triadic exchange. As the child progresses through Phases V and VI, more child-directed play is encouraged and the child is able to choose from a menu of activities and is expected to communicate - either through gesture or speech - of her/his preference.

An important aspect of the study protocol is that each preceding phase provides the opportunity to build the basic skills necessary for succeeding in subsequent phases of the intervention. By first engendering trust and confidence, the exercises which follow can focus on scaffolding increasingly more challenging social and communication skills that rely on a well-established protocol with which the child has already become familiar.

Lessons Learned Early in the field study, study personnel made several significant observations that were contributing to and limiting success with the robot intervention. First,

the music and hand play featured in Phase III did effectively capture the attention of study participants. Most of the children in the study seemed to enjoy the music, directed their attention to CHARLIE and engaged in at least some imitation.

Second, the exercises described in Phases I and II for promoting trust by familiarizing the child with the robot's kinematics were accomplished in a much shorter period of time. Some of the children participating in the study completely bypassed the first two phases and were ready to engage directly with the robot from the outset, while the remaining children progressed through the first two phases within minutes of their first session.

Third, the "Imitate You, Imitate Me" and "Pass the Pose" games, as designed, were ineffective in practice. Motor imitation games without any kind of positive sensory reinforcement were not effective for attracting a child's attention or maintaining his/her interest. Instead, what we discovered is that a key component driving productive, interactive gameplay is the robot's *reaction*. The increases in attention, eye contact, communication, speech and social interaction we observed as a result of the robot reacting in an amusing way were dramatic.

Finally, while other studies have shown that a robot can be used to effectively catalyze communication by encouraging attention, motor imitation and turn-taking, generalizing these behaviors to co-present others has remained a challenge. In our field study, generalizing robot-child games to child-to-co-present others is easily facilitated when others in the room engage in gameplay using the same amusing reaction performed by the robot.

Improved Approach As a result of these important observations, the original study procedure was modified. While some of the phases in the initial procedure remained unchanged, new software was written to deliver a new interactive game, the existing tele-

operated imitation game was modified, less emphasis was placed on Phases I and II, and more emphasis was placed on child-directed play. This last modification effectively led to more child-initiated creative play, where speech introduced by study personnel to play robot-assisted games was often spontaneously generalized by the child to communicate a need or to engage in a different (but related) game created by the child. All modifications were made to introduce or emphasize positive sensory reinforcement received by the child when s/he successfully engages in interactive social play.

The description of Phases I and II include the child making eye contact with and greeting the robot and others in the room, touching and moving the robot's arms. These exercises remain unchanged from the original study procedure. However, instead of devoting one or two sessions to this activity these are all included as part of Phase I (and because the approach is based on scaffolding these skills, are also integrated into each subsequent session). Additionally, the improved approach features the robot responding by saying "hello" or "goodbye" and waving its hand as part of the greeting and parting process. Requiring that the child say "hello" and "goodbye" and/or wave to the robot and co-present others encourages social interaction, motor imitation and, in most cases, verbalization. Several participants enjoyed being able to "cause" the robot to respond in this way so much that they would practice this greeting repeatedly, seemingly enjoying their control over the robot and the predictability of its response.

The revised Phase II also features the addition of the simple game of peek-a-boo. Because this game requires some manipulation of the robot's arms and the game is at or below the developmental level of all the children participating in the study, this is an activity that the participants can master and enjoy early in their intervention; especially, when paired with the amusing response the child receives when the robot says "boo!". Moreover, the game easily generalizes to co-present others and establishes the practice of rehearsing

games with the robot and then immediately generalizing them to others in the room.

Phase III remains unchanged except for the modification of the teleoperated game of imitation. While the original study procedure describes allowing the child to fully control the robot's motions with the remote control to familiarize her/him with the robot's kinematics and sounds and to encourage the child's trust of CHARLIE, Phase III of the improved study procedure places more emphasis on reinforcing motor imitation by providing positive sensory feedback for the child's participation through one-to-many (robot-to-others) group imitation. This revised version of the teleoperated game of imitation promotes joint attention, motor imitation and turn-taking gameplay.

Revisions made to Phases IV, V and VI were significant. The "Imitate You, Imitate Me" and "Pass the Pose" games were excluded from the study procedure and replaced by a new, interactive game called "the Hat Game". Inspired by one of the participants in the study who, when asked a question, enjoyed shaking his head in response, the "Hat Game" mainly features CHARLIE responding in an amusing way to questions posed by the child or a co-present other. For example, when one of several hats is placed on CHARLIE's head and the child or co-present other asks, "CHARLIE, do you like your hat?", the robot responds by shaking its head and saying "Nooooo". This typically elicits surprise and laughter from the child (and co-present others), thereby encouraging her/him to ask again or to try another hat.

Since its inception, the "Hat Game" has been expanded to include an assortment of accessories including several kinds of hats, sunglasses, a scarf, a flower clip and wolf ears and now features an alternative, more positive response of, "Yeah!" As the child becomes more able to ask the appropriate question, even if only in part, the game is generalized to the child and to others in the room. When it is the child's turn, all others in the room direct

their question to the child and wait for him/her to respond. When it is another person's turn, the child is encouraged to face that person and verbalize the question while pointing his/her finger at the person wearing the accessory.

The "Hat Game" is an important part of our intervention strategy which effectively encourages and improves eye contact, directed attention, speech and social interaction by providing a positive sensory response to reinforce each child's efforts to communicate. Moreover, because the game is simple enough to be played in any setting, with almost any accessory, the increases in verbal utterances observed during gameplay in the clinic often generalize to other settings. In fact, several of the participants' caregivers reported that their child began initiating and/or playing the game at home and in the car during their participation in the study.

The last significant change to the original study procedure, was the addition of a visual schedule for many of the children participating in the study. By offering a visual menu of choices from which each child could select an activity, the child is encouraged to actively participate in the direction of the session and to communicate his/her choice for the next exercise. While child-directed activity is part of Phase VI in the original study procedure, the improved version incorporates this aspect using the visual schedule from the very first few phases of the intervention.

3.7 Field Study Results

Five evaluative tests were administered at the outset and at the completion of each child's participation in the study. These tests include the: (1) VABS-II, (2) MLSUD, (3) MIS, (4) UIA and (5) EVT2. A within-subject t-test statistic was performed on scores obtained from evaluative tests and a discussion of results collected is provided in the "Within-study data results" subsection below. Four additional analyses were conducted to compare results

recorded for the within-study group data with results extracted from similar autism intervention studies which used the same clinical instruments to assess improvements in motor imitation, speech, communication and socialization. The comparative analyses cross-evaluate data collected from the: (1) VABS-II, (2) MLSUD and (3) UIA and are included in the subsection entitled, "Comparative data analysis".

Within-study data results

Raw data from the five evaluative tests administered during the field study are presented in Tables 3.2, 3.4, 3.5, 3.6. Pre- and post-communication and pre- and post-socialization composite scores are included in Table 3.2. MLSUD scores recorded during the preliminary session, two intermediate sessions and the final session with participants are presented in Table 3.4 and results from the Unstructured Imitation Assessment - including data for social, requesting and joint attention behavior - are included in Table 3.5. Finally, result from the Expressive Vocabulary Test and Motor Imitation Scale are included in Table 3.6.

T-test statistics. A paired-samples t-test statistic was computed to evaluate the significance of increases reported for each of the seven evaluative categories: (1) VABS-II Communication Domain, (2) VABS-II Socialization Domain, (3) VABS-II Receptive and Expressive Communication V-Scale Scores, (4) MLSUD, (5) UIA social imitation, (6) UIA requesting, (7) UIA joint attention, (8) MIS and (9) EVT2. For each of these categories, the raw data collected for each participant was used to compute a difference score:

$$d_i = X_i^{post} - X_i^{pre} \tag{3.1}$$

where the difference score, d_i, is the reported change between the pre-test, X_i^{pre} and post-test, X_i^{post}, scores. Next, the mean between the difference scores was calculated:

$$\overline{x} = \Big(\sum_i^N d_i\Big)/N \tag{3.2}$$

and used to determine whether the average difference score was large compared to the variability in the difference scores:

$$t = \overline{x}/(\sigma/\sqrt{N}) \tag{3.3}$$

The resulting t value was used as a significance measure of pre- and post-test scores collected for each individual where $\overline{x}$ is the computed mean of difference scores, σ is the standard deviation of collected data and N is the number of sample pairs collected. The t value was then converted into a probability, p, which describes the probability of the difference in collected data being due to sampling error. In other words, the p-value is the area under the null distribution curve that is in bigger disagreement with the null hypothesis than the observed test statistics. A critical t value, t_{crit}, is the cutoff threshold value which determines when samples give cause to reject or fail to reject the null hypothesis of a test statistic and is defined for all t-tests performed herein as $t_{crit} = \pm 2.36$ with seven degrees of freedom and a two-tailed alpha threshold p value of 0.05. Six out of nine of the t-tests performed demonstrate that field tests resulted in statistically significant increases as they exceed the following requirements:

$$t_{crit} \geq 2.36, p < 0.05 \tag{3.4}$$

A detailed discussion of the raw data collected and results from the within-subject t-tests performed are presented, by category, below.

VABS-II. The VABS-II questionnaire was completed by a caretaker for each child in the study group. Composite communication and socialization scores were extracted from each pre- and post-test VABS-II questionnaire to present within-study data for expressive and receptive communication and social skills results.

Table 3.2: **Parent/Caregiver Reported Vineland Adaptive Behavior Scale Results.** Participant identifier and beginning age are followed by (from left to right): (a) Composite communication pre-test score, (b) Composite communication post-test score, (c) Change in composite communication scores, (d) Composite social interaction pre-test score, (e) Composite social interaction post-test score, (f) Change in composite socialization scores.

Part	**Pre-test Comm.**	**Post-test Comm.**	**Percent Change**	**Pre-test Social**	**Post-test Social**	**Percent Change**
2	49	59	**+20.41%**	61	61	**0.0%**
3	42	40	**-4.76%**	51	49	**-3.92%**
4	79	79	**0.0%**	72	79	**+9.72%**
7	83	91	**+9.64%**	83	95	**+14.46%**
9	69	72	**+4.35%**	74	97	**+31.08%**
10	54	67	**+24.07%**	57	86	**+50.88%**
11	57	61	**+7.02%**	61	81	**+32.79%**
12	85	83	**-2.35%**	74	81	**+9.46%**
AVG	64.7	69.0	**+6.65%**	66.6	78.6	**+18.02%**

Field study results show a mean increase of 6.65% in overall communication and a 18.02% mean increase in social interaction skills as reported by caregivers on the Vineland-II Parent/Caregiver Rating Form (Table 3.2). The range of increases in composite communication scores collected was approximately 28.8% and the range of composite socialization scores was approximately 54.8%. Five out of the eight participants demonstrated an increase in the Communication Domain and six out of eight showed an increase in the Socialization Domain.

There are a few possible causes for the lack of improvement in the VABS-II Communication Domain recorded for a few of the participants in our study. First, the VABS-II provides graduated scores based on the chronological age of the child being assessed. It is likely that the participants moved from one age class to another during the course of their involvement in the study and this change in classification may have diminished any increases that would have otherwise been reflected in their scores. Second, the commu-

nication composite scores include 3 areas of communication: (1) receptive, (2) expressive and (3) written. Lack of increases or relative decreases in the written portion of the communication composite may have also diminished actual gains in expressive and receptive communication.

Because our study does not address written communication, results and analysis for the VABS-II V-Scale Scores representing the Receptive and Expressive Communication Domain scores are provided for comparison with the Communication Composite Scores (Table 3.3). As illustrated, average overall increases in composite communication scores were approximately 6.65% while mean increases in the combined receptive and expressive communication domains were slightly higher at approximately 9.84%.

Results from the within-subject t-test show a mean improvement rating of $t(7) = \pm 2.14$, where $p < 0.0699$ for the Communication Domain and a mean improvement rating of $t(7) = \pm 3.06$, where $p < 0.0184$ for the Socialization Domain. These results demonstrate that while increases recorded for the VABS-II Communication Composite Domains do not indicate that the null hypothesis can probabilistically be ruled out, mean increases observed for the Socialization Domain are confirmed to be statistically significant. Additionally, t-test results for the Receptive and Expressive Domain scores show a statistically significant mean gain of $t(7) = \pm 2.51$, where $p < 0.0404$.

MLSUD. A total of four separate MLSUD measures were collected; scores were computed at the first meeting with the child, during sessions 4 and 8 and at the very end of the six-week intervention period (Table 3.4). Only MLSUD scores from the preliminary and final evaluation sessions are included and used to perform comparative assessments and the t statistic analysis of increased speech since intermediate sessions featured a significant amount of rote (non-spontaneous) speech that was used to engage in games with CHAR-

Table 3.3: **Parent/Caregiver Reported VABS-II Combined Receptive and Expressive Communication V-Scale Results.** Participant identifier and beginning age are followed by (from left to right): (a) Total receptive and expressive communication v-scale pre-test score, (b) Total receptive and expressive communication v-scale post-test score, (c) Change in receptive and expressive communication pre-test/post-test v-scale scores and (d) Change in composite communication scores

Part.	Pre-test V-Scale	Post-test V-Scale	Percent Change in Pre/Post-Test V-Scores	Percent Change in Composite Comm Scores
2	12	14	+16.67%	+17.0%
3	8	7	-12.5%	-4.76%
4	21	21	0.00%	0.0%
7	16	19	**+18.75%**	**+9.64%**
9	20	20	0.00%	+4.35%
10	13	17	**+30.77%**	**+24.07%**
11	14	16	**+14.29%**	**+7.02%**
12	18	20	**+11.11%**	**-2.35%**
AVG.	15.25	16.75	**+9.84%**	**+6.65%**

LIE and with co-present others.

A mean increase of 35.3% in spontaneous speech as calculated by the Mean Length Spontaneous Utterance Determination measure (Table 3.4) was observed after a 6-week intervention with CHARLIE. The range of preliminary MLSUD scores was from 0.00 to 2.76 and final MLSUD scores ranged from 0.00 to 3.77. All participants but one demonstrated an increase in MLSUD score. The single participant whose MLSUD did not improve throughout the course of the intervention was nonverbal at the start of his 6-week intervention and did not demonstrate any acquired speech during the final evaluation. Although his scores for the VABS-II and EVT2 also did not improve, the child's scores did improve in each of the three categories included in the UIA evaluation. This might suggest that while the participant did not make any significant gains in communication or speech,

Table 3.4: **Mean Length of Spontaneous Utterance Determination (MLSUD) Results.** Participant identifier and beginning age are followed by (from left to right): (a) Preliminary evaluation MLSUD, (b) Session 4 MLSUD, (c) Session 8 MLSUD, (d) Final evaluation MLSUD, (e) Greatest increase in MLSUD, (f) Change in pre-test/post-test MLSUD.

Part.	**Age**	**Prelim**	**Mid1**	**Mid2**	**Final**	**Greatest change**	**Pre/Post Change**
2	4:10	0.37	1.17	1.52	1.08	**+305.3%**	**+188.0%**
3	4:1	0.0	0.0	0.0	0.0	**0.0%**	**0.0%**
4	3:5	2.76	2.90	3.26	3.77	**+36.6%**	**+36.6%**
7	3:10	1.75	3.24	3.98	1.86	**+127.4%**	**+6.3%**
9	6:2	2.38	3.13	3.21	3.18	**+34.9%**	**+33.6%**
10	6:5	1.55	1.72	1.97	1.83	**+27.1%**	**+18.1%**
11	6:5	1.41	2.42	2.95	2.5	**+109.2%**	**+77.3%**
12	3:3	1.80	2.92	2.25	2.05	**+62.2%**	**+13.9%**
AVG.	4:10	1.50	2.19	2.39	2.03	**+59.3%**	**+35.3%**

Note: Mid1 and Mid2 changes include some rote speech learned for gameplay. Spontaneous speech increases are evaluated using only the Pre/Post percentage changes listed above.

Table 3.5: **Unstructured Imitation Assessment Results.** Participant identifier followed by the percentage of successful imitations for: (1) Preliminary UIA-social, (2) Final UIA-social and, (3) Change between preliminary and final social scores; (4) Preliminary UIA-requesting, (5) Final UIA-requesting and, (6) Change between preliminary and final requesting scores; (7) Preliminary UIA-joint attention, (8) Final UIA-joint attention and, (9) Change between preliminary and final joint attention scores.

Part.	**UIA-Soc(A)**	**UIA-Soc(B)**	**Change**	**UIA-Req(A)**	**UIA-Req(B)**	**Change**	**UIA-JA(A)**	**UIA-JA(B)**	**Change**
2	28.0%	83.0%	**+196.4%**	20.8%	45.8%	**+120.2%**	23.3%	33.3%	**+42.9%**
3	5.0%	22.2%	**+344.0%**	4.2%	25.0%	**+495.2%**	3.3%	23.3%	**+606.1%**
4	50.0%	66.7%	**+33.4%**	33.3%	33.3%	**+0.0%**	30.0%	46.7%	**+55.7%**
7	100%	100%	**0.0%**	20.8%	45.8%	**+120.2%**	13.3%	30.0%	**+125.6%**
9	39.0%	56.0%	**+43.6%**	20.8%	71.0%	**+241.3%**	13.3%	66.7%	**+401.5%**
10	33.3%	28.0%	**-15.9%**	17.0%	41.7%	**+145.3%**	20.0%	40.0%	**+100.0%**
11	28.0%	78.0%	**+178.6%**	29.0%	45.8%	**+57.9%**	40.0%	33.3%	**-16.8%**
12	61.1%	100.0%	**+63.7%**	20.8%	91.7%	**+340.9%**	43.3%	90.0%	**+107.9%**
AVG	42.4%	66.7%	**+57.3%**	20.8%	50.0%	**+140.4%**	22.9%	45.4%	**+98.3%**

Table 3.6: **Expressive Vocabulary Test (EVT) and Motor Imitation Scale (MIS) Results.** Participant identifier followed by: (a) Preliminary EVT(A), (b) Final EVT(B), (c) Percent change in EVT score, (d) Preliminary MIS(A), (e) Final MIS(B), (f) Percent change in MIS score.

Participant	**EVT(A)**	**EVT(B)**	**CHANGE**	**MIS(A)**	**MIS(B)**	**CHANGE**
2	42	73	**+73.8%**	44.0%	44.0%	**0.0%**
3	42	42	**0.0%**	0.0%	0.0%	**0.0%**
4	114	121	**+6.1%**	81.3%	90.6%	**+11.4%**
7	97	114	**+17.5%**	84.0%	94.0%	**+11.9%**
9	111	105	**-5.4%**	100.0%	100.0%	**0.0%**
10	87	69	**-20.7%**	100.0%	88.0%	**-12.0%**
11	94	94	**0.0%**	94.0%	100.0%	**+6.4%**
12	114	119	**4.4%**	87.5%	96.9%	**+10.7%**
AVERAGE	87.6	92.1	**+5.1%**	73.8%	76.7%	**+3.9%**

some fundamental imitation and attention skills - key precursors to communication and speech - were improved throughout the course of the study.

The paired-samples t-test showed a mean improvement of $t(7) = \pm 3.56$, where $p < 0.0092$ for the MLSUD category. These t-test statistics again confirm that the increases observed in the collected raw data are significant for the population of children included in our study.

UIA. The UIA is one of the two assessments we used to evaluate growth in imitation ability in three major areas: (1) social interaction, (2) requesting and (3) joint attention and consists of a total of 24 measures. Pecentages included in Table 3.5 reflect the ratio of the total number of points received by each child to the total number of points achievable on the assessment. Points provided for each measure reflect the number of examples successfully demonstrated by the child during the 60-minute initial and final evaluations. The maximum number of points assigned for each of the 24 measures is “3”, giving a total of

72 points possible. Analysis of the raw UIA data show that mean increases for each of the three UIA areas - social imitation, requesting and joint attention - were 57.3%, 140.4% and 98.3%, respectively.

Within-subject t-test results show a significant increase in the social interaction domain ($t(7) = \pm 3.02$, where $p < 0.0193$), the requesting domain ($t(7) = \pm 3.79$, where $p < 0.0068$) and the joint attention domain ($t(7) = \pm 3.23$, where $p < 0.0145$). Interestingly, these test results reveal that the most significant improvements were achieved in the requesting domain - a primary focus area of the intervention provided. The t-test statistic confirms that the observed increases in UIA pre- and post-test scores for all three UIA domains were each statistically significant.

MIS. The MIS is the only instrument used to assess motor imitation ability and was administered at the beginning of the study period and at the end, for each participant. The MIS evaluates motor imitation using a total of 16 measures which assess a child's ability to imitate meaningful and nonmeaningful actions and body movements. Percentages included in Table 3.6 reflect the ratio of the total number of points received by each child to the total number of points achievable. The maximum number of points assigned for each of the 16 measures is "2", giving a total of 32 points possible for the MIS. A "2" indicates a passing score, a "1" indicates an emerging skill and "0" indicates a failure for that particular skill.

Mean increases in scores for the MIS were marginal, improving by only 3.9%. Since activities included in the intervention did not primarily focus on improving motor imitation skills as much as social interaction, communication and speech, these results were not unexpected. T test results showed a mean improvement of $t(7) = \pm 1.07$, where $p < 0.3182$, which were not statistically significant for this population in this study.

EVT2. The Expressive Vocabulary test was administered primarily to provide additional data regarding each child's progress in word acquisition and retrieval as a secondary measure of communication skill. The EVT2 consists of a total of 190 items and is typically administered by a speech-language pathologist, psychologist or early childhood specialist. The test features a series of pictures depicting objects, people and situations and is administered by the examinee who prompts the child to name or describe a picture after being provided a stimulus question. Scores reported in Table 3.6 are based on the extracted Growth Scale Values (GSVs), a metric used for easily measuring each child's progress over time.

Improvements in the EVT2 were also slight and were not found to be statistically significant for this population of children using this intervention ($t(7) = \pm 0.86$, where $p < 0.4166$). Again, given that the our primary study objectives did not include targeting the acquisition of new vocabulary or improving word retrieval ability, these statistical conclusions are not surprising. Instead, they do provide additional information about other mitigating factors that may contribute to an individual's performance on other tests administered during the study. For example, a child with apraxia of speech may improve marginally on the VABS-II Communication Domain, the MLSUD and the EVT2, but show greater improvements in the VABS-II social domain, the MIS and the UIA. Given that a few of the study participants had other known medical diagnoses, including some that limited the physical ability of the child, a future study with a larger study population and further analysis would shed light on the possible effects these complicating factors.

Comparative data analysis

Results obtained from the VABS-II, UIA and MLSUD were also evaluated in comparison with data reported from other, similar autism intervention studies using these same instruments to measure improvements in communication and socialization. These three in-

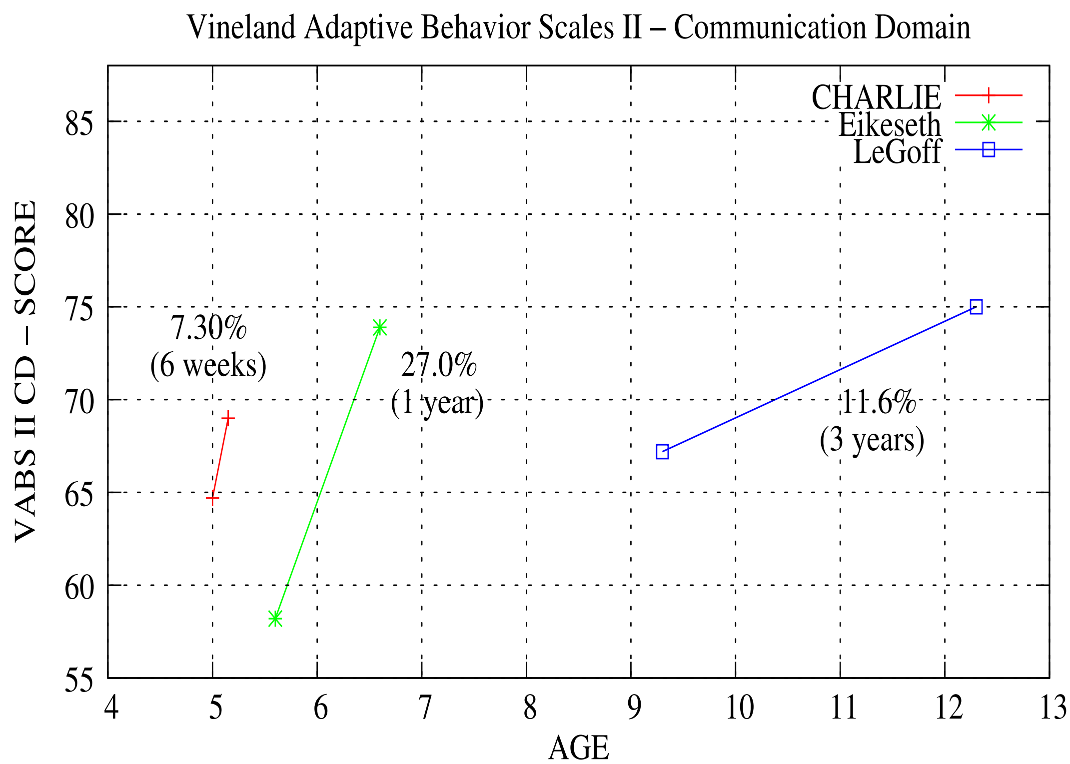

Figure 3.10: Reported communication composite scores for the Vineland Adaptive Behavior Scale obtained through three studies; Red line: CHARLIE, Green line: Eikeseth and Blue line: LeGoff.

struments were selected for comparison primarily because each test sufficiently measures a specific, fundamental communication skill targeted in our intervention from more than one point of view. For example, two composite scores from the VABS-II - Socialization and Communication - were extracted for evaluation and comparison since they directly measure the child's progress in overall expressive and receptive language and socialization from the caregiver's viewpoint. This is an important assessment because caregivers can evaluate their child's improvement given their performance in different settings and with different people. The 3-part UIA assesses three fundamental communication skills (social interaction, requesting and joint attention) that were instrumental for playing several of CHARLIE's interactive games. Finally, the MLSUD is a widely-used tool for measuring progress in speech. Comparative analyses are provided, by category, below.

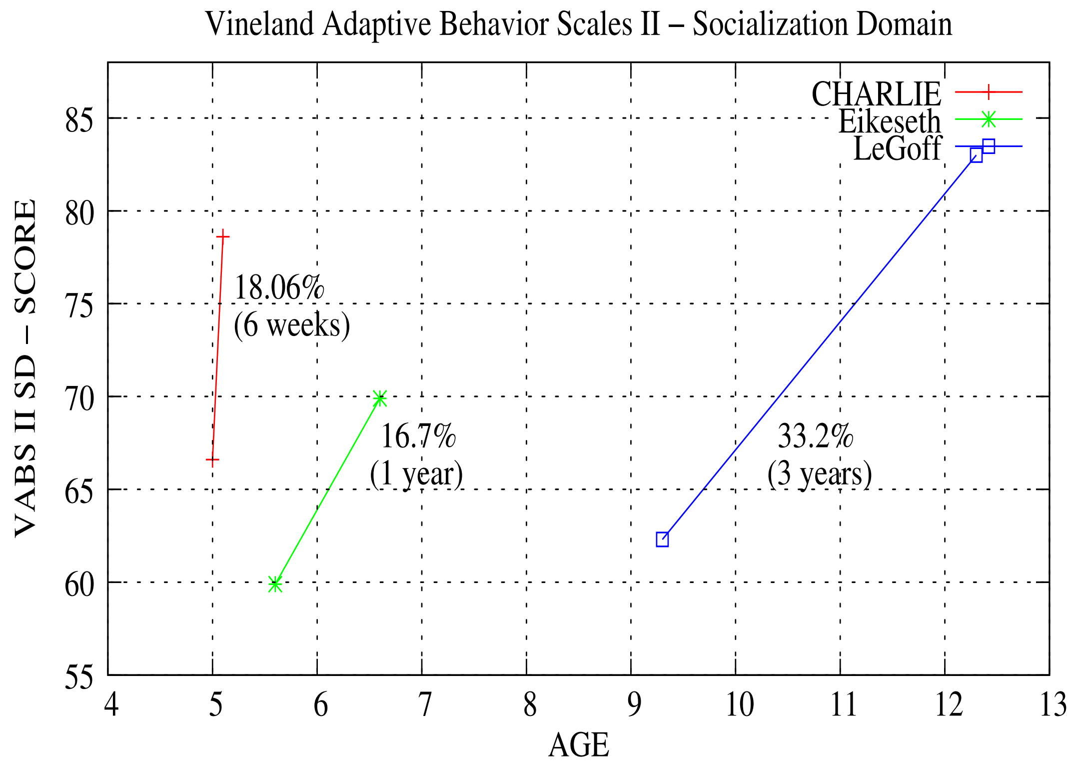

Figure 3.11: Reported socialization composite scores for the Vineland Adaptive Behavior Scale obtained through three studies; Red line: CHARLIE, Green line: Eikeseth and Blue line: LeGoff.

Table 3.7: Comparing average reported points for the VABS communication composite scores increases obtained from the CHARLIE, Eikeseth, LeGoff studies.

Study	**VABS Pre Communication**	**VABS Post Communication**	**Points increased per hour of therapy**
CHARLIE	62.00	66.60	0.767
Eikeseth	58.23	73.93	0.015
LeGoff	67.20	75.00	0.034

Table 3.8: Comparing average reported points for the VABS socialization composite scores increases obtained from the CHARLIE, Eikeseth, LeGoff studies.

Study	VABS Pre Socialization	VABS Post Socialization	Points increased per hour of therapy
CHARLIE	65.50	77.30	1.967
Eikeseth	59.92	69.92	0.010
LeGoff	62.27	82.95	0.090

Comparing VABS-II Results In this subsection, we provide a comparative analysis of the VABS-II results obtained in our study with those obtained from two different research studies.

Eikeseth et al., 2002. A study conducted in 2002 with a population of 25 children with autism, aged 4 to 7 years, examined the efficacy of an intensive behavioral treatment based on the University of California at Los Angeles (UCLA) treatment model [77]. Participants in the study were evaluated by a pediatric neuropsychologist, child psychiatrist, child psychologist and a speech pathologist and confirmed the diagnosis of autism based on results from the administration of the VABS, either the Wechsler Preschool and Primary Scale of Intelligence-Revised (WPPSI-R) or the Wechsler Intelligence Scale for Children - Third Edition (WISC-III) and the Gilliam Autism Rating Scale (GARS). The research compared gains achieved through the behavior modification treatment with gains achieved through an eclectic intervention consisting of sensory-motor therapy, applied behavior analysis and methods derived from the clinician's personal experience. Children who participated in the study received a minimum of 20 hours of treatment per week for the period of one year (or approximately 1,040 hours of intervention). This study demonstrated that for both the Communication and Socialization Domains of the VABS-II, the group receiving behavioral treatment showed the greatest improvements.

We compared the VABS-II Communication and Socialization scores obtained from our study group, which received a robot-assisted intervention for only 6 weeks (or approximately 6 hours), to the behavioral test group, which demonstrated the most increases in these domains. Over the one year study period, children in the Eikeseth et al study showed a mean increase of 16.7% in the VABS-II Socialization Domain, and a 27.0% increase in VABS-II Communication Domain. Comparative graphs, based on VABS-II scores from our study and those from the Eikeseth study are included in Figures 3.10 and 3.11.

LeGoff et al., 2006. A second comparative analysis of recorded VABS-II scores was performed using another research study from 2006 which included 60 children with autism, with a mean age of 9 years at the start of the study [78]. Children participating in the study had a confirmed autism diagnosis based on the results from Autism Diagnostic Interview-Revised (ADI-R), deviation IQ of 50 or more on the WPPSI-R or a ratio IQ of 50 or above on the Bayley Scales on Infant Development-Revised and the absence of major medical conditions other than autism. The study group included in the study received LEGO therapy once per week for 90 minutes, for a total period of 3 years. This equates to approximately 230 hours of intervention. While participants in the study group showed an average increase of 33.2% in VABS-II Socialization Domain they also demonstrated an 11.6% improvement in the VABS-II Communication Domain. Results comparing VABS-II scores from our study with those reported in the LeGoff study are included in Figures 3.10 and 3.11.

Two additional tables are provided to compare point increases observed per hour of intervention for the VABS communication and socialization components for each of the three studies (Tables 3.7 and 3.8).

Comparing UIA Scores. A study published in 2010 from Ingersoll et al ([75]), explores the effects of Reciprocal Imitation Training (RIT) for object and gesture imitation on language behavior. Specifically, researchers sought to determine whether adding ges-

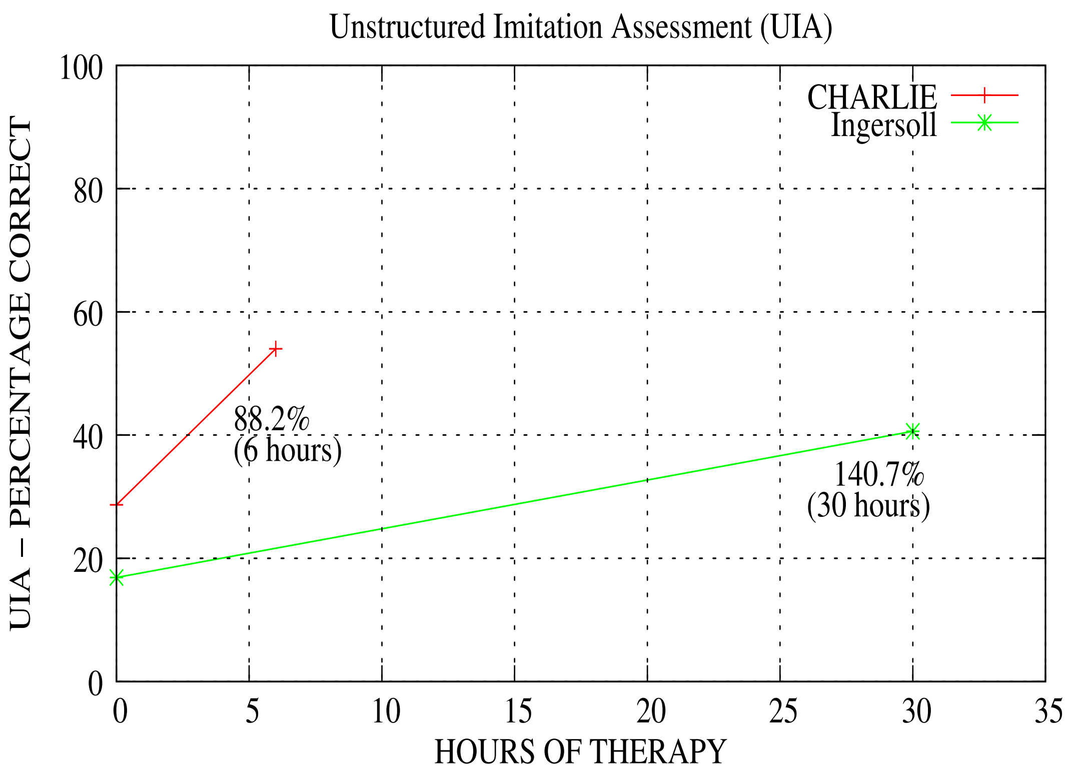

Figure 3.12: Unstructured Imitation Assessment scores for two studies; Red line: CHARLIE, Green line: Ingersoll.

Table 3.9: Comparing average reported points for the Unstructured Imitation Assessment from the CHARLIE and Ingersoll studies.

Study	**UIA Pretest Average Score**	**UIA Posttest Average Score**	**Points increased per hour of therapy**
CHARLIE	28.7	54.0	4.22
Ingersoll	16.9	40.6	0.79

ture imitation training improved the overall rate of appropriate language use and whether children were more likely to engage in verbal imitation during object or gesture imitation training using RIT. The study included four children, between 35 and 47 months of age who received 1 hour of intervention each day, 3 days per week for a total period of 10 weeks (or approximately 30 hours of therapy). Children showed a mean increase of 140.7% in UIA scores in pre- and post-intervention scores. Figure 3.12 compares cumulative UIA increases using the number of intervention hours received and the pre-test and post-test scores recorded for each participant in each study.

An additional table demonstrating score increases (per hour of intervention) for each of the two studies is included for comparison (Table 3.9).

Comparing MLSUD Scores. A final comparative analysis was performed using a study published in 2002 by Hancock et al [79]. The research study included four children between 2.5 and 5.0 years of age who received an intervention based on Enhanced Milieu Training (EMT). Children who participated in the study received the EMT intervention two times per week for a total of 24 15-minute sessions (equivalent to approximately 6 hours of total intervention). The primary goal of the research was to examine the effects of Enhanced Milieu Teaching on the social communication skills of preschool children with autism. Mean Length Spontaneous Utterance Determination scores for participants in the study showed a mean increase of 30.7% over the study period and are included along with MLSUD scores collected for participants in our study in Figure 3.13.

Reported MLSUD scores for the Hancock study and our study demonstrate comparable results obtained from two different but fundamentally similar interventions. EMT is described as a set of language tools that facilitates communication in children. Five of the key strategies of EMT include:

- Setting up an Interactive Context

- Responsive Interaction
- Modeling and Expanding Play
- Modeling Communication Targets
- Environmental Arrangement Strategies

Although the EMT model does not include a robot, several of the intervention strategies listed above are incorporated as part of the intervention used in our study. By introducing a non-threatening robot “playmate”, an interactive context is facilitated in which opportunities for communicating and connecting socially with each child are created. Teaching and reinforcing basic communication and social skills are further enhanced through responsive, fun interactions with the robot and co-present others. Finally, modeling and expanding play by varying communication targets is an integral part of our intervention. For example, the “Hat Game” relies on an interventionist modeling behavior and speech for the child to imitate and then expanding play through turn-taking with co-present others. However, the differences between our study and an EMT intervention are many.

While EMT relies on a naturalistic, conversation-based approach that expands on a child’s interests and initiations as opportunities to model and prompt language in everyday contexts, our intervention strategy with CHARLIE provides a more structured environment that doesn’t require initiation on the child’s behalf. This is especially useful for children who don’t often initiate communication or explicitly demonstrate interests and in situations when time is limited. Further, all participants for the Hancock study were screened to ensure that they were already verbally imitative and possessed an expressive vocabulary of at least 10 spontaneous words. Neither of these requirements were used to screen participants in our study. In fact, two of the eight children in the CHARLIE study were nonverbal at the outset of their participation in the study.

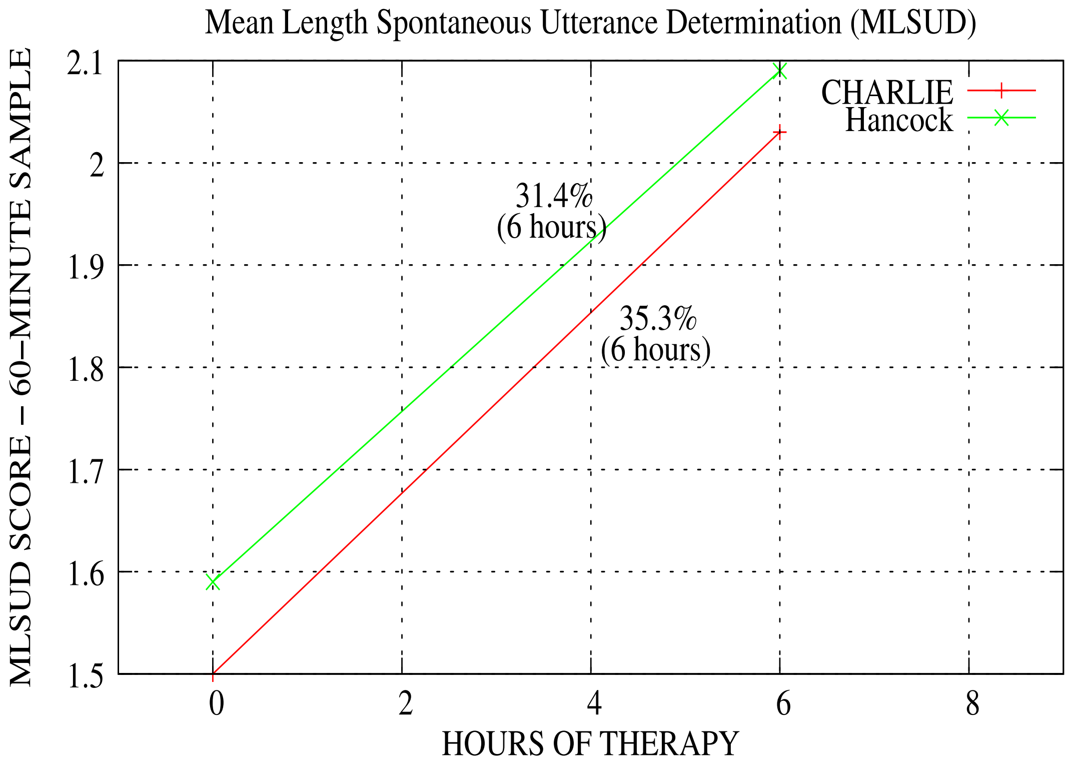

Figure 3.13: Mean Length Spontaneous Utterance Determination scores for two studies; Red line: CHARLIE, Green line: Hancock.

3.8 Interactive Robot Conclusions

Conclusions and contributions This portion of our research resulted in the design and development of a low-cost, adaptive robot and several interactive games for use in robot-assisted autism therapy. One of the aims of our research is to create a robot that is financially accessible to a greater population of therapists and families with children with ASD in order to broaden the impact of traditional therapies. Another objective was to develop a hand detector enabling a larger scope of interactive games in which the robot can engage a child autonomously. Achieving this second objective allows for real-time collection of important user interaction information specific to the preference and progress of each child undergoing autism therapy. Collectively, these contributions produce a new robot which is designed to be child-centered and adaptive to user preference, while fulfilling a

key supportive role for therapists by automatically generating user progress reports.

Additionally, measuring improvements in communication, social interaction and speech through the combined use of a manipulable robot with a specific sequence of interactive games provides critical insight as to the viability of using simple robots like CHARLIE as a widely-accessible tool for autism therapy. Our intervention promotes speech, communication and social interaction by scaffolding basic proficiencies through a series of imitation and turn-taking games. Our technique also employs a robot that can withstand some physical manipulation by children which encourages child-initiated interactions and ultimately leads to the generalization of acquired skills with co-present others.

While foundational skills such as motor imitation and turn-taking are key to facilitating communication and socialization, these skills must be practiced in a context that is conducive to interaction and that provides responses which make the learning fun and interesting. Providing an environment that promotes imitation and turn-taking by rewarding the child with a fun response is an effective strategy for increasing speech, communication and social interaction. Further, our study demonstrates that children with autism can learn to effectively generalize learned behaviors from robot-child interactions to human-human interactions

Finally, although there has recently been growing interest in increasing the autonomy of robots in robot-assisted autism interventions, a semi-autonomous robot may actually be preferable for certain activities. For example, rewarding a child appropriately during social interactions and after child-initiated speech is of critical importance to reinforcing socialization and communication. A human expert who can discern between a child who is having trouble articulating a request from a child who is not making the effort can appropriately provide the robot's positive response, when it is warranted. Our human-in-the-loop method emphasizes the importance of leveraging the expertise of therapists, teachers and

caregivers to direct the course of the robot intervention. A crucial characteristic of the method is the way the robot responds when prompted. By relying on the discretion of a human operator to assess a child's ability to perform a particular task (on a particular day) s/he can control the progression of gameplay and determine when a sensory reward is appropriate.

CHAPTER 4

REMOTE STRESS DETECTION

Numerous robotics applications stand to benefit from remote stress detection. Stroke and post-cardiac surgery patients undergoing physical therapy may be asked to engage in exercises that repeatedly work a particular limb or muscle. A robot that is able to monitor the patient's heart and breathing rate during therapy can adjust the workload based on those physiological indicators. Search and rescue robots designed to measure stress remotely can potentially assess the physical condition of victims found in an emergency or disaster scenario and relay this information back to a base station. Assistive robots employed to help the elderly live independently can monitor breathing and heart rate on a regular basis and detect possible warning signs that signal an impending health crisis. Remote stress detection can also be an important tool for robots interacting with people with developmental and physical disabilities by providing critical physiological information to teachers, therapists and caregivers in many aspects of their lives.

A variety of methods have been used to collect data about a user's emotional or stress state including measuring the amount of eye contact, body pose, number, quality and content of verbal utterances, and several physiological indicators such as galvanic skin response, EEG, breathing and heart rate. Galvanic skin response measures changes in the electrical conductance of skin [80] while EEG is used to measure the voltage fluctuations resulting from ionic current flows within the neurons of the brain [81]. Breathing is a physiological indicator which has been referred to as the “neglected vital sign" and is used as a critical measure of a user's psychophysiological state [82, 83].

Two basic modalities have been employed to capture breathing events: contact and non-

contact. Contact approaches have used wearable sensors such as thermistors, respiratory gauge transducers and acoustic sensors. These devices typically deliver accurate breathing data, but are not suitable for mobile applications or when the use of wearable sensors is otherwise impractical. In addition, although solutions exist using non-contact methods such as infrared video cameras, radar and doppler modalities, these approaches rely on high-cost equipment and collecting and analyzing very large amounts of data at a high processing cost. Some existing non-contact modalities may be suitable for medical applications such as polysomnography which require in-depth recovery of very specific breathing information, but the computational and equipment cost is not reasonable for most human-robot interaction (HRI) applications where recovery and analysis of breathing information must occur in real-time and sensor costs are just one component of the total system expense. Another important distinction between existing non-contact methods and our approach is that we are interested in recording changes in basic breathing and heart rate as an indicator of a user's psychological or physical state, not in diagnosing medical conditions.

One of the novel contributions of our research is a simple, autonomous and low cost system for the real-time collection and monitoring of breathing and heart rate. Our research presents a new non-contact breathing and heart rate measurement technique suitable for most HRI applications. This is accomplished by continuously targeting a high precision infrared (IR) sensor, tracking changes in the sub-nasal skin surface temperature, curve-fitting a sinusoidal function to extract the breathing rate and performing a Discrete Wavelet Transform (DWT) to automatically compute the heart rate from recorded temperatures.

4.1 Methodology and Approach

Our research presents a new technique for remotely capturing and measuring breathing and heart rates to ultimately deduce a user stress state and use this state information to adapt the behavior of an interactive robot. Although resting breathing and heart rates may vary from one individual to the next, healthy adults have a typical resting breathing rate between

8-16 breaths per minute [84] and a resting heart rate of 60-120 beats per minute [85].

Normal breathing consists of three phases: inspiration, expiration and a postexpiratory pause. Inspiration occurs when the diaphram contracts, creating negative pressure inside the chest cavity and the passive process of expiration follows as a function of the elastic recoil property of the lungs [86]. For this research, one complete breathing cycle is measured as the interval between the beginning of the expiration phase and the beginning of the next expiration phase.

Each beat of the heart consists of a series of deflections reflecting the time evolution of electrical activity in the heart that is responsible for initiating muscle contraction. A single heartbeat is typically decomposed into five constituent parts labeled : P, Q, R, S, and T. The largest-amplitude portion of the ECG is the QRS complex, caused by currents generated when the ventricles depolarize prior to their contraction. We are most interested in measuring the QRS component of the cardiac cycle where one heart beat is measured from the beginning of one QRS cycle to the beginning of the next.

The breathing and heart rate measurement system employs a single-point infrared sensor (see Figure 2.1). To measure breathing and heart rate with this sensor, the system: (1) aims the sensor at a pre-defined sub-nasal target region using the location of the nose as extracted from the most recent video frame and, (2) extracts the temperature information provided by the sensor analog signal. To achieve these objectives, a specific combination of hardware and software was included in the overall system design.

Remote Breathing and Heart Rate Detection Hardware

The remote breathing and heart rate measurement method presented herein relies on collecting temperatures using the same custom-built actuated platform, on which a non-contact infrared temperature sensor and a camera are mounted. The primary hardware components of this system are enumerated below.

1. *Infrared Sensor:* The infrared sensor is a FAR infrared (spectral range $8 - 14$ mi-

Figure 4.1: Single-point FAR infrared sensor.

crons) sensor, manufactured by Micro-Epsilon, model “thermoMETER CX-SF15-C8.” The sensor is capable of reading a range of temperatures between $30 - 150$°C. This single-point sensor has an optical resolution of $15 : 1$ with a reading precision or temperature resolution of 0.025°C and an accuracy of approximately ±1°C or 1% of the reading. Power is supplied to this sensor using an Arduino microcontroller, and digitizes its output analog signal using a standard 6.5 digit bench multimeter.

2. *Camera:* A consumer-grade USB camera is mounted below the IR sensor to assist maintaining a correct aim of the IR sensor at the user’s sub-nasal region.

3. *Pan-tilt platform:* Both the IR sensor and the camera are mounted on a direct-drive pan-tilt platform, actuated by two titanium gear servos. The servos are powered by a 7.4V, 2100mA lithium polymer battery, and controlled over USB via an 8-channel servo controller board.

Remote Breathing and Heart Rate Detection Software

The software that was developed for measuring breathing and heart rate achieves five main objectives: (1) infrared sensor positioning, (2) infrared temperature collection, (3) data pre-

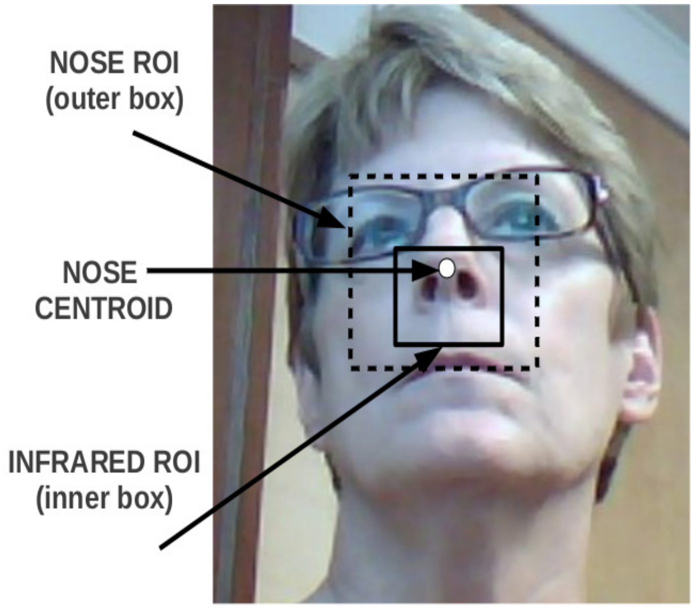

Figure 4.2: Nose and infrared sensor regions of interest.

processing, (4) breathing rate calculation and, (5) heart rate calculation. Sensor positioning is accomplished through repeated nose detection and automatic adjustment of the infrared sensor's pan and tilt angles in order to maintain the region of interest within a pre-defined target. Temperatures are obtained by sampling the IR sensor and processing the signal to convert to a Fahrenheit temperature reading. Finally, the data set is smoothed with a low-pass filter, curve-fit to extract the breathing rate and transformed using a DWT to calculate the heart rate.

Positioning the IR sensor

A camera is used to assist with the positioning of the infrared sensor. Because the camera is mounted on the same pan-tilt platform as the infrared sensor, the center of the sensor's target region remains at a constant (x, y) offset from the center of each camera image. In our research, the sensor's target region is referred to as the "infrared region of interest" (irROI.) For the sensor to be positioned properly, a corresponding target point on the image of the subject's face is defined and aligned with the irROI from one image to the next.

The nose classifier from the freely-available OpenCV library [11] is used to perform nose detection, extract the nose region of interest (ROI) and compute the (x, y) coordinates of the nose centroid. Positioning the sensor to point at the sub-nasal target area is accomplished by maintaining the nose centroid within the irROI in each image video frame.

Algorithm 1 Track Infrared ROI

```
repeat
    get image frame
    find nose in image
    while nose not found do
        get another image frame
        check for nose in the image
    end while
    // obtain centroid for nose ROI
    centroid.x ← (noseROI(xpos)+noseROI(width))/2
    centroid.y ← (noseROI(ypos)+noseROI(height))/2

    // Check sensor's position
    if centroid.x < irROI.lowerLeft.x then
        pan the platform left 1
    end if
    if centroid.x > irROI.upperRight.x then
        pan the platform right 1
    end if
    if centroid.y < irROI.lowerRight.y then
        tilt the platform down 1
    end if
    if centroid.y > irROI.lowerRight.y then
        tilt the platform up 1
    end if
until program exit
```

The irROI is large enough to accommodate variances in physical features, but small enough to ensure that temperature fluctuations occurring during the breathing cycle will be detected. In this system, the irROI covers about 5% of the total image area. Since the infrared sensor reading reflects the average temperature of its target surface area, the primary requirement is that at least a sufficiently large portion of the target contains the sub-nasal area.

Continuous positioning of the infrared sensor is driven by repeated nose detections which provide the (x, y) coordinates of the centroid of the nose ROI. This centroid position must be maintained within the rectangular irROI in order to keep the sensor in a stable position. Should the nose centroid move outside of the irROI, incremental pan and tilt commands are automatically generated and executed until the sensor returns to a stable position. Algorithm 1 illustrates the process for positioning the IR sensor.

Infrared Temperature Collection

Collecting data for breathing rate extraction The infrared sensor is initially sampled for 30 seconds, the temperature data set is stored and an initial breathing and heart rate are computed. Subsequent data sets consist of the last 25 seconds from the previous data set and the next 5-second window of temperature data. Samples are collected at a rate of approximately 20-30 samples per second and each is recorded along with a corresponding timestamp. The breathing rate is recalculated each time a full 5-second window of breathing data is collected. This “sliding window" approach enables the system to detect subtle changes in the breathing rate quickly since small increases or decreases in breathing begin to affect the overall breathing rate within a few seconds (Figure 4.3).

Collecting data for heart rate extraction The infrared sensor is continuously sampled until a window of 32 time-stamped samples or approximately 1.6 seconds of temperature data has been collected. Various window sizes were tested in order to evaluate the system’s performance during periodic heart rate fluctuations. Although larger window sizes provide higher stability in computed heart rates, they are prone to excessive smoothing and reduce the system’s ability to detect short-lived heart rate increases or descreases. Further, while small window sizes are susceptible to being dominated by relatively small errors that can be introduced when temperatures are collected during re-targeting, they provide more resilient and responsive heart rate detection overall. Two representative samples of collected raw IR data are included in Figure 4.4.

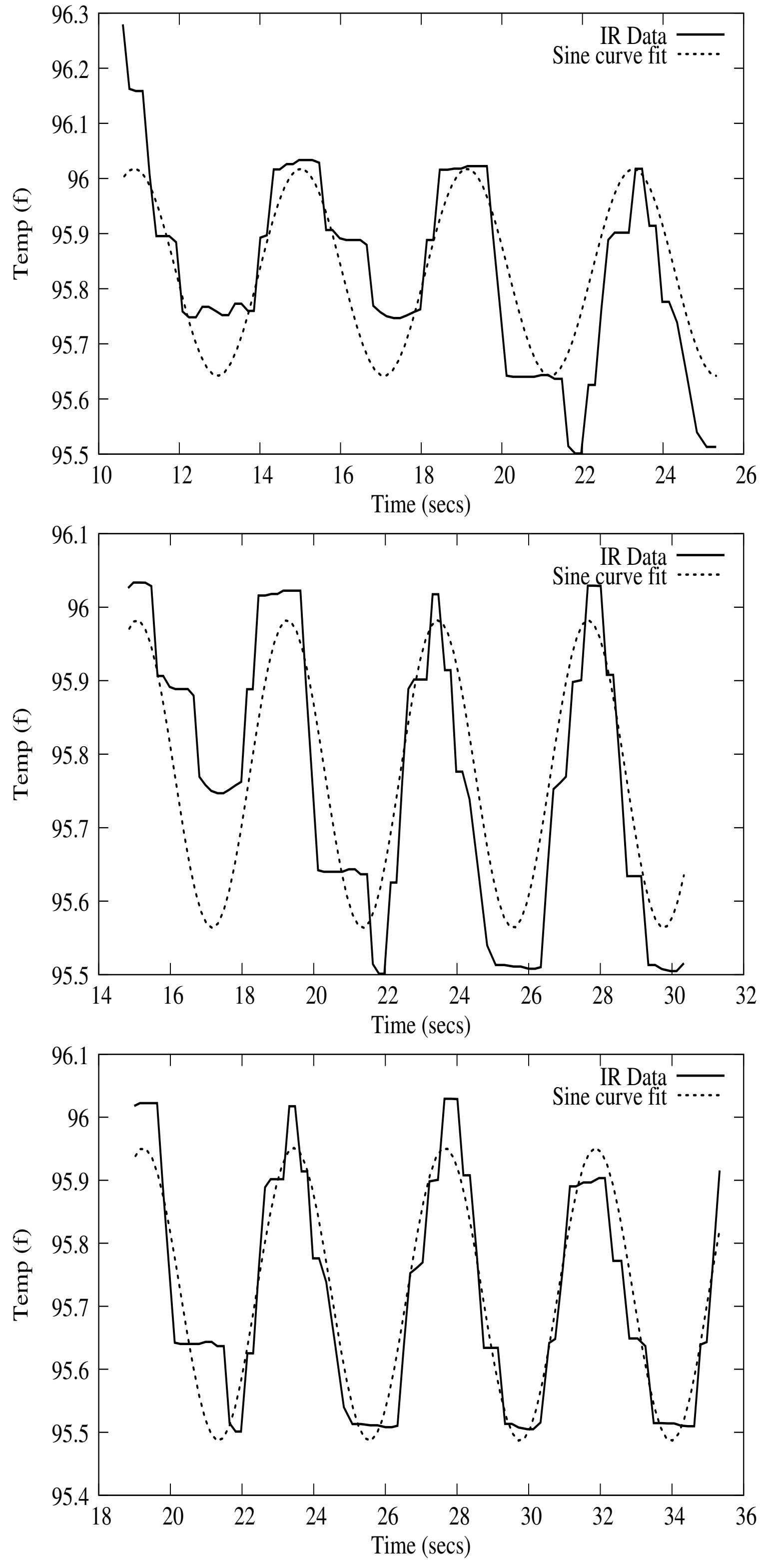

Figure 4.3: Successive sliding windows of IR sensor readings. Solid line corresponds to infrared temperature data. Dashed line corresponds to the curve-fit results.

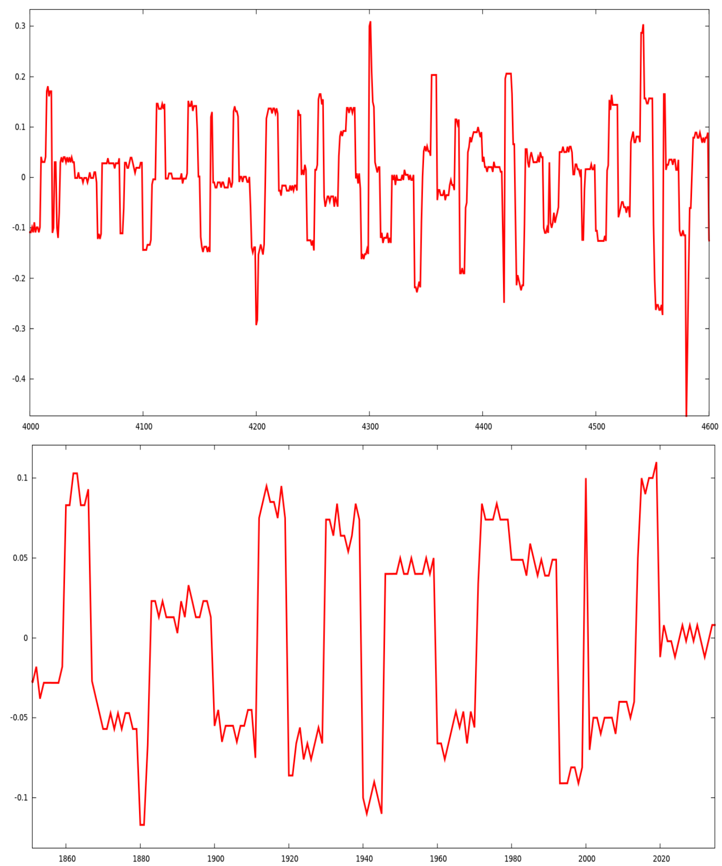

Figure 4.4: 30-second window of unprocessed infrared data (top). Detailed 8-second view of raw data (bottom).

Data Pre-Processing

Data pre-processing mitigates minor errors occurring from temperatures collected during sensor re-positioning and occasional noise. Temperatures are continuously sampled from the infrared sensor regardless of whether the sensor is stable or not. For this reason, readings too low to be considered human body temperature are assumed to be room temperature or another non-human source and are excluded from the collected data set. Additionally, in order to smooth out occasional noise from the sensor signal, a low-pass filter is applied to each set of data collected by the infrared sensor. Finally, to make the IR data suitable for processing with a DWT and extracting the heart rate, the 0-mean is computed for all the samples in each window.

Breathing Rate Calculation

We obtain individual breathing rates for the infrared sensor data sets by fitting a sinusoidal curve to the infrared data. The best curve-fitting results were observed using four basic fitting parameters: period T, mean B, amplitude A, and offset C.

$$FittedSine(x) = A \sin(2\pi/Tx + C) + B \tag{4.1}$$

Based on the results of this curve-fitting operation, the value assigned to the variable T is used as the breathing rate output of the system. We performed curve fitting using the curve-fitting command provided by the freely-available graphing utility, gnuplot [87]. Gnuplot uses an implementation of the nonlinear least-squares Marquardt-Levenberg algorithm, where a user-defined function is fit to a set of input data. After each iteration of the algorithm, the quality of the fit is determined by the sum of the squared differences or “residuals" between the input data points and the function values, evaluated at the same places. Each iteration of the algorithm attempts to minimize the residuals, and terminates only when a specific residual minimum or limit is reached. In the current implementation, data sets are curve-fit in real-time so it is important to set the residual low enough for an accurate fit but not so low that computation causes delay in processing each consecutive data set. After testing several values, we found the best residual limit to be 10^{-15}. The function we defined to fit the collected temperature data sets is a sinusoidal curve, with several fitting parameters added.

To help in assessing the quality of the curve-fit and the resulting breathing rate output, a running average of the residuals was computed for each data set and an error threshold was defined. Residuals in excess of five times the residual average for each data set were not considered successful and were classified as “no response." This technique allows the system to determine when it has succeeded in fitting a curve to the infrared data and to avoid generating erroneous results when this is not the case.

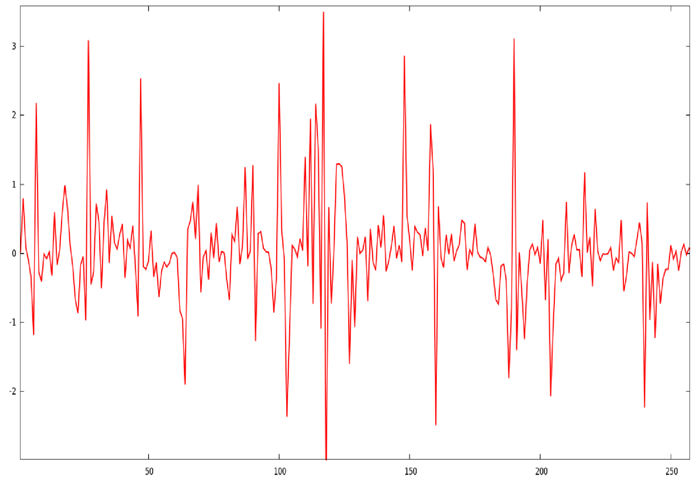

Figure 4.5: Level 3 DWT coefficients.

Heart Rate Calculation

Heart rates are computed using a DWT [88] on each window of collected infrared data. A DWT is used to process the IR data for two fundamental reasons : (1) heart rates are not stationary since they have varying frequency components at different time intervals and, (2) we are interested in the temporal information associated with each reading. Unlike Fast Fourier Transforms (FFTs), DWTs are capable of extracting specific frequencies occuring at particular time intervals.

The DWT first sends samples through a low pass filter which yields approximation coefficients and a high pass filter which results in one or more detail coefficients. The outcome of this filtering technique is that the component signal frequencies are cut in half and according to Nyquist's rule, half the samples can be discarded. Although this process halves the time resolution and each output has half of the input frequency band (since only half of each filter output characterizes the entire signal), the frequency resolution is effectively doubled with each decomposition.

The decomposition process is recursively repeated in order to increase the frequency resolution until no further decompositions are possible. Once the decomposition is completed, a set of coefficients are output that were produced at various scales and at different

time intervals of the signal. The coefficients can then be analyzed to extract frequency information for particular time intervals or for the signal in its entirety.

Due to the nature of this technique, the number of samples processed in a given data set by the DWT must be in powers of two. Pre-processed IR data sets are padded in order to meet this requirement. Our system uses the daubechies(6) wavelet to perform the transform and collects temperature readings at a rate of 20 samples per second so the highest frequency that can be extracted is 10 samples per second or 10 Hertz (Hz). The range of frequencies in which we are most interested for this research are 0.8-1.90 Hz because they correspond to heart rates between 48 bpm and 114 bpm. The DWT levels of decomposition which contain the detail coefficients within that frequency range are found at levels 3 and 4 and represent frequencies between 1.25-2.5 Hz. (level 3) and 0.625-1.25 Hz. (level 4). Figure 4.5 illustrates a representative level 3 coefficient file produced for one set of pre-processed IR data.

Finally, the heart rate is extracted by (1) computing and comparing the average amplitude of the detail coefficients at levels 3 and 4, (2) selecting the level with the largest average amplitude, (3) counting the number of zero crossings for the coefficients at the selected level and, (4) multiplying that number by 37.5.

4.2 Remote Breathing and Heart Rate Testing

Breathing Detection Tests

We conducted experiments to test and measure the effectiveness of a single-point infrared sensor for monitoring breathing rate. In order to evaluate the performance of various parts of the system, preliminary tests were conducted before formal tests were carried out.

Informal Preliminary Tests

Early tests underscored various limitations in nose detection, curve-fitting and self-reporting of breathing rates using a push button to record expiration. The OpenCV nose classifier used in the nose detection system is prone to drifting caused by excessive ambient light and changes in the angle of the nose. Additionally, because typical breathing tends to vary in terms of breath length and frequency of breaths, temperature data collected do not always conform to a sinusoidal wave which is characteristically consistent within each data set. Finally, collecting the ground truth for breathing using manually reported data is susceptible to inconsistencies caused by participants failing to report breaths.

To reduce complications with the nose detection system, we subsequently conducted formal testing in a temperature-controlled setting with controlled lighting conditions. In order to manage errors resulting from a poor fit of the sine wave function to the breathing data, the sum of squares residuals were extracted from curve-fitting results in order to assess the quality of the fit between each data set and the curve generated. Inaccuracies in manually reported data were mitigated by removing instances when it was clear that the study participant failed to report breathing events. It would be impossible to remove all inaccurately reported data since a participant may fail to report just one or two breaths per minute for a given data set. However, data sets in which manually reported breaths were 0.00 per minute for a given window were discarded. Even though the error contributed by each of these factors was minimized, it was not completely eliminated. Consequently, we believe that it is likely that formal test results were still somewhat influenced by errors from limitations in nose detection, curve-fitting and manually reported data and resulted in higher residual values and lower values for accuracy.

Formal tests

We conducted formal experiments with ten study participants, four females and six males, between the ages of 18 and 60. Individuals who participated in this study were not tak-

Data Sets	Typical	Anomaly	All
Successful response	74.8%	16.4%	68.8%
No response	25.2%	83.6%	31.2%

Figure 4.6: Response rate across data sets.

ing medication which could interfere with their breathing at the time of the experiment. Each participant sat in a chair that was placed approximately 30 inches from a rolling table equipped with the infrared sensor system and a laptop computer. Each study subject watched a video playing on the laptop computer for approximately 10 minutes. The primary purpose of the video was to maintain the participant's attention in a forward-facing, relatively still position. In addition, participants were provided a push-button sensor to self-report their breaths. This self-reported data was used as ground truth to evaluate our system's performance. Each individual was asked to breathe naturally through the nose and to push the button through the entire expiration phase.

As with the infrared temperature collection, manual reports of each expiration were collected in 15-second windows and subsequent data sets consist of the last 10 seconds of the previous window and the next 5-second window. Each data entry in the window includes a time stamp and a corresponding "high" (expiration) or "low" (other) temperature. Breathing rates based on manually reported expirations are calculated by dividing the number of complete breaths recorded in a given window by the total number of seconds elapsed between the beginning of the first complete expiration and the beginning of the last reported expiration.

Experiment Results

Ten test sets, consisting of approximately 120 infrared and 120 manually reported breathing rates each, were collected and analyzed. Of those 10 sets, three were identified as anomalous due to frequent nose detection problems observed while the test was being con-

Data Sets	Typical	Anomaly	All
< 4 bpm	94.1%	76.9%	91.2%
< 3 bpm	86.9%	71.2%	83.2%
< 2 bpm	70.9%	53.8%	63.5%
< 1 bpm	42.4%	38.5%	37.4%

Figure 4.7: Successful response results in breaths per minute.

ducted or a large number of missing manually reported entries. The other 7 test sets contain typical data collected when the nose detection and tracking was working properly and there were few, if any, missing manually reported entries. Approximately 75% of the typical test sets yielded breathing rates at or below the residual threshold (five times the residual average) and were classified as “successful" compared to approximately 25% of the anomalous test sets. Data from both typical and anomalous test sets exceeding the error threshold were classified as “no response." A summary of the successful and “no response" rates for typical and anomalous data sets is shown in Figure 4.2.

Accuracy was evaluated by computing the difference between breathing rates detected by the infrared sensor and breathing rates reported by study participants. Typical and anomalous test sets were analyzed separately and accuracy was measured in breaths per minute (bpm). Breathing rate entries were classified into one of four basic categories: (1) under 4 bpm, (2) under 3 bpm, (3) under 2 bpm and, (4) under 1 bpm, as illustrated in Figure 4.2.

An important factor in assessing the accuracy of test data is the error threshold for determining which breathing rates were fitted successfully with the curve-fitting function and which were not. As the error threshold is increased, the number of breathing rates in the “no response" category decreases along with the number of breathing rates in the “successful" category. Conversely, with a very low error threshold, the number of successful breathing rates increases as does the number of breathing rates classified as “no response." A relatively low error threshold was selected so as to evaluate only the data which most

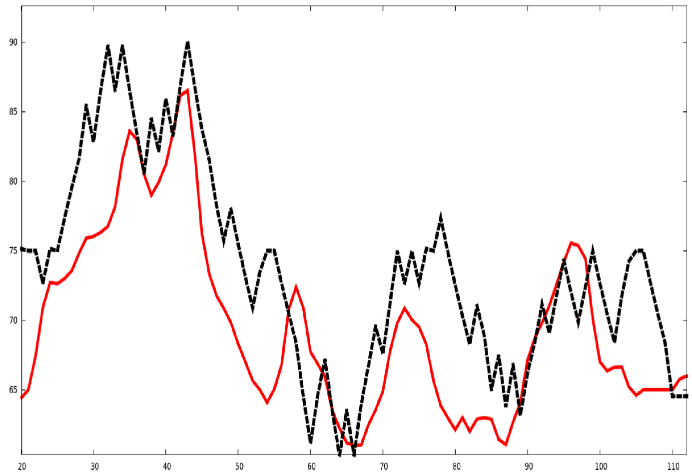

Figure 4.8: IR heart rate (dashed line) and ECG heart rate window (solid line).

accurately reflects the effectiveness of the infrared sensor. If a higher error threshold were used, the number of breathing rates with a poor fit would be increased and the resulting accuracy would reflect more about the performance of the curve-fit function and less about the sensor's ability to detect temperature changes corresponding to breathing.

Heart Rate Detection Tests

Experiments were conducted to measure the effectiveness of the single-point infrared sensor for detecting heart rates remotely. A representative graph of extracted heart rates as detected by the IR sensor and by the ECG illustrates typical results over a period of approximately 90 seconds in Figure 4.8.

Formal Tests

Experiment setup

The proposed system's accuracy were quantitatively measured by collecting temperature data with the infrared sensor, computing the heart rate and comparing the results with heart rate data obtained from an ECG. Additionally, anti-aliasing testing was performed to ensure

that fluctuations detected by the IR sensor were not due to resolution limitations of the sensor. The sensor sampling rate, independent from any other processing, is approximately 300 samples per second. The sensor was targeted at a surface with a constant temperature and flucutations were measured for over 10 seconds. Temperatures for the entire 10 second period fluctuated $0.001°$F.

For the ECG data collection, participants were each fitted with 3 electrodes attached to a bioradio which continuously transmitted heart rate data to a nearby computer. ECG information is collected at approximately 600 samples per second and a heart rate is computed for each 960 samples, or 1.6-seconds of ECG data, so that IR and ECG heart rates can be easily processed and compared.

Range of RMSE in bpm	**Percentage of All Cases**
< 4	25.0%
< 9	66.7%
< 14	83.4%
< 19	91.8%
< 24	95.9%
< 29	100.0%

Figure 4.9: All test set results by root-mean-square error.

Range of RMSE in bpm	**Percentage of Typical Tests**
< 4	27.3%
< 9	72.7%
< 14	90.9%
< 19	100.0%

Figure 4.10: Typical test set results by root-mean-square error.

Because the ECG data collected during experiments consists of a heart rate without a time-stamp, part of the system performance analysis includes an auto-correction for the

temporal alignment of data between ECG heart rates and IR-derived heart rates by comparing the root-mean-square errors (RMSE) of various offsets for each window of coefficients computed.

Experiments included 24 study participants, 17 females and 7 males, between the ages of 18 and 35. Individuals who participated in this study were not taking medication which could interfere with their heart rate at the time of the experiment. Each participant was asked to sit in a chair that was situated approximately 1 meter from a rolling table equipped with the infrared sensor system and a laptop computer. During the course of each 10-minute test session study subjects watched a video playing on the laptop computer. The primary purpose of the video was to maintain the participant's attention in a forward-facing, relatively still position.

Experiment Results

Twenty four test sets, each consisting of approximately 10 minutes of heart rate data were collected and analyzed. Of those 24 sets, two were identified as anomalous due to persistent nose detection problems observed while the test was being conducted. The remaining 22 test sets contain typical data collected when the nose detection and tracking was working properly. Approximately 73% of the typical test sets averaged heart rates within 0-9 beats per minute compared to average heart rates produced by the ECG over the entire 10-minute test set (Figure 4.2).

Overall system accuracy was measured by computing the difference between the reported ECG heart rate and the IR detected heart rate for each 1.6-second window (Figure 4.2). Typical and anomalous test sets were analyzed separately and accuracy was assessed in beats per minute (bpm). Six categories were used to classify our results: (1) 0-4 bpm, (2) 5-9 bpm, (3) 10-14 bpm, (4) 15-19 bpm, (5) 20-24 and (6) 25 and higher bpm.

An additional consideration in the assessment of system performance is the system's ability to effectively track increases and decreases in heart rate even when the baseline is

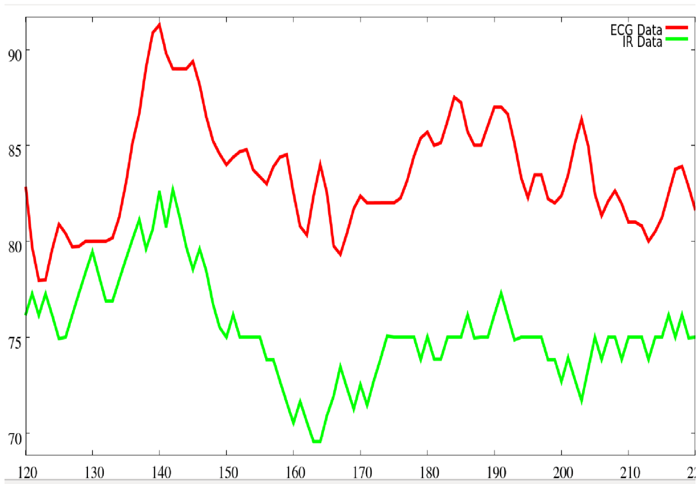

Figure 4.11: Offset of IR heart rate (lower line) and ECG heart rate (upper line).

shifted by an offset as shown in Figure 4.11. Test sets that mirror heart rate fluctuations as reported by the ECG but are offset by a certain amount will produce higher RMSE scores on average even though increases and decreases in heart rate are accurately detected. Future work will include an evaluation of these cases to determine if they can still be used to provide valuable information pertaining to changes in heart rate that are indicative of stress state.

4.3 Remote Breathing and Heart Rate Detection Conclusions

Our research presents a new non-contact technique for monitoring changes in the sub-nasal skin surface temperature to calculate the breathing and heart rate of a user. Heart rate and breathing rate information may be useful for deducing the user's stress state and adapting the behavior of an interactive robot. Our main research objectives were to examine the effectiveness of using a non-contact, computationally lightweight and low cost sensor for accurately measuring breathing and heart rate. Overall, the results obtained from the first sets of formal experiments for collecting and measuring breathing and heart rates are very promising. Data from typical test sets clearly demonstrate that a single-point infrared sen-

sor, when accurately positioned, can detect the subtle temperature changes corresponding to respiration and cardiac pulse. Given the small size of the sensor and the minimal computation required to perform non-contact monitoring (as compared to existing methods), this research demonstrates the usefulness of this sensing modality for a variety of HRI applications.

The first round of testing for monitoring breathing rate highlighted limitations with the methods used to position the sensor, collect ground truth and automatically compute breathing rates using a curve-fitting approach. Several enhancements were made to the system software for the heart rate detection tests which followed. These improvements made it possible to collect more samples per second which, in turn, allowed us to use the more accurate DWT method for the automatic calculation of heart rate.

Positioning the IR sensor properly and precisely is based on the detection of a subject's nose using a nose classifier which is "trained" with various samples of noses. However, feature detectors that are based on trained classifiers are susceptible to differences in illumination, scale and rotation. The nose detection system used in our research can be improved in one of several ways. First, the nose classifier can be replaced with one that we train to detect noses in a greater range of poses, scales and lighting conditions. Second, the camera was originally positioned approximately 6-8 inches below the subject's nose. By adjusting the position of the camera in the heart rate detection tests, so that it is at the approximate level of the user's face caused a significant reduction in drifting errors that were caused by the system's inability to detect a nose. Although breathing tests were conducted with the camera positioned below the user's face, in subsequent tests for calculating heart rate the camera was placed at the approximate height of the user's nose. As a result, we experienced a much lower rate of drifting and false positive nose detections.

Collecting ground truth via self-reporting introduces two undesirable factors. First, the "true" breathing rate reported by the user may be unreliable since the experiment is conducted while the subject is engaged in watching a short film and s/he may become

distracted and forget to depress the button to report a breath. Second, the natural pattern of breathing is disrupted when a subject is intentionally conscious of their breathing. This may lead to an unnaturally longer or anomalous inspiration/expiration pattern and produce results which are not reproducible when the subject is unaware that breathing is being measured. During our second set of experiments, ground truth was automatically collected using a respiratory belt tranducer and ECG making it much less susceptible to human error.

The sinusoidal curve-fitting function that was used to automatically compute breathing rates, sometimes yields inconsistent or inaccurate results due to variances in the typical wave period from one user to the next and anomalies in breathing patterns that commonly occur within a single session. In order to obtain a more robust data set and improve the accuracy of the automatic breathing rate computation, the software is now multi-threaded so that the two processes managing face detection/tracking and temperature collection/analysis are now executed separately. This simple modification resulted in an increase to the collected samples per second (sps) from 6 sps to 20-30 sps. Additionally, a Fast Fourier Transform (FFT) was implemented for extracting the frequency of the temperature intensities collected. Due to the significant increase in the number of sps collected and the way in which FFT analyzes the data, the estimate of breaths and pulses per minute was consistently more accurate.

CHAPTER 5

RESEARCH PLAN

Table 5.1: Dates of completion for research goals

Timeline	**Description of Work Completed**
January 2011	Robot prototype
January 2011	Face detection and tracking
January 2011	New hand classifier / hand detection
June 2011	Single-player interactive game
June 2011	Tests for single-player game
July 2011	Teleoperation and Two-player interactive game
August 2011	Infrared sensor system built
October 2011	Remote breathing detection
November 2011	Tests for breathing detection
March 2012	Remote heart rate detection
March 2012	Tests for heart rate detection
July 2012	Proposal defense
August 2012	Identify team of autism, psychology, speech, hearing experts
October 2012	Develop field study protocol, test design for field study
January 2013	Obtain IRB approval, Advertise field study
February 2013	Order and obtain testing materials for field study
February 2013	Recruitment of study participants
April 2013	Initiate field tests
November 2013	Complete field tests
December 2013	Dissertation defense

My dissertation defense was approved on December 12, 2013.

BIBLIOGRAPHY

[1] Gabriel Barbash and Sherry Glied. New technology and health care costs - the case of robot-assisted surgery. *The New England Journal of Medicine*, 303:701–704, August 2010.

[2] Peter S. Lum, Charles G. Burgar, Peggy C. Shor, Matra Majmundar, and Machiel Van der Loos. Robot-assisted movement training compared with concentional therapy techniques for the rehabilitation of upper-limb motor function after stroke. *Archives of Physical Medicine and Rehabilitation*, 83(7):952–959, July 2002.

[3] Fumiaki Hiramatsu S. Ishii, S. Tanaka. Meal assistance robot for severely handicapped people. In *Proceedings IEEE International Conference on Robotics and Automation*, volume 2, pages 1308–1313, May 1995.

[4] Kaoru Inoue, Kazuyoshi Wada, and Yuko Ito. Effective application of pao : Seal type robots for disabled people in according to ideas of occupational therapists. In Klaus Miesenberger, Joachim Klaus, Wolfgang Zagler, and Arthur Karshmer, editors, *Computers Helping People with Special Needs*, volume 5105 of *Lecture Notes in Computer Science*, pages 1321–1324. 2008.

[5] Kerstin Dautenhahn. Robots as social actors: Aurora and the case of autism. In *Proc. Cognitive Technology Conference*, pages 359–374, 1999.

[6] Karla Conn Welch, Uttama Lahiri, Zachary Warren, and Nilanjan Sarkar. An approach to the design of socially acceptable robots for children with autism spectrum disorders. *International Journal of Social Robotics*, pages 391–403, 2010.

[7] Prevalence of autism spectrum disorders - Autism and developmental disabilities monitoring network, United States, 2006. `http://www.cdc.gov/mmwr/preview/mmwrhtml/ss5810a1.htm`, January 2011. Accessed: 01-21-11.

[8] Chrystopher Nehaniv. Synchrony and turn-taking as communicative mechanisms. In C. Nehaniv and K. Dautenhahn, editors, *Imitation and Social Learning in Robots, Humans and Animals*. Springer, 2007.

[9] Jeroen C.J. Brok and Emilia I. Barakova. Engaging autistic children in imitation and turn-taking games with multiagent system of interactive lighting blocks. In *Proceedings of the 9th international conference on Entertainment Computing*, pages 115–126, 2010.

[10] Laura Boccanfuso and Jason M. O'Kane. CHARLIE: An adaptive robot design with hand and face tracking for use in autism therapy. *International Journal of Social Robotics*, 3:337–347, 2011.

[11] Gary R. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, pages 120–126, NOV 2000.

[12] Laura Boccanfuso and Jason M. O'Kane. Remote measurement of breathing rate in real time using a high precision, single-point infrared temperature sensor. In *Biomedical Robotics and Biomechatronics (BioRob), 2012 4th IEEE RAS EMBS International Conference on*, pages 1704–1709, 2012.

[13] Laura Boccanfuso, Eva Juarez Perez, Myra Robinson, and Jason M. O'Kane. Collecting heart rate using a high precision, non-contact, single-point infrared temperature sensor. In *Proceedings of the 4th International Conference on Social Robotics*, ICSR'12, pages 86–97. Springer-Verlag, 2012.

[14] Colwyn Trevarthen. Intrinsic motives for companionship in understanding: Their origin, development, and significance for infant mental health. *Infant Mental Health Journal*, 22(1–2):95–131, 2001.

[15] Hideki Kozima, Cocoro Nakagawa, and Hiroyuki Yano. Can a robot empathize with people? *Artificial Life and Robotics*, 8(1):83–88, 2004.

[16] Autism Spectrum Disorders - Diagnostic Criteria. `http://www.cdc.gov/ncbddd/autism/hcp-dsm.html`, December 2011. Accessed: 12-06-11.

[17] Autism Society of America. `http://www.autism-society.org/about-autism/`, June 2011. Accessed: 06-29-11.

[18] 'Baldi' helps kids with autism - virtual talking head teaches autistic children language skills. `http://www.webmd.com/brain/autism/news/20080215/baldi-grows-on-kids-with-autism`, June 2011. Accessed: 06-29-11.

[19] Meredith Ringel Morris, Anthony Cassanego, Andreas Paepcke, Terry Winograd, Anne Marie Piper, and A. Huang. Sides: A cooperative tabletop computer game.

In *Proceedings of the 2006 20th anniversary conference on Computer Supported Cooperative Work (CSCW)*, pages 1–10, 2006.

[20] Andrea C. Pierno, Morena Mari, Dean Lusher, and Umberto Castiello. Robotic movement elicits visuomotor priming in children with autism. *Neuropsychologia*, 46:448–454, 2008.

[21] Sylvia Weir and Ricky Emanuel. Using LOGO to catalyse communication in an autistic child. *Technical Report, DAI Research Report*, 15, 1976.

[22] Iain Werry and Kerstin Dautenhahn. Applying mobile robot technology to the rehabilitation of autistic children. In *Proceedings of the Symposium on Intelligent Robotics Systems. Coimbra, Portugal*, pages 265–272, 1999.

[23] François Michaud, Tamie Salter, Audrey Duquette, Henri Mercier, Hèléne Larouche, and François Larose. Assistive technologies and children-robot interaction. In *Multidisciplinary Collaboration for Socially Assistive Robotics, Association for Advancement of Artificial Intelligence (AAAI) Spring Symposium*, pages 45–49, 2007.

[24] Ben Robins, Farshid Amirabdollahian, Ze Ji, and Kerstin Dautenhahn. Tactile interaction with a humanoid robot for children with autism: A case study analysis involving user requirements and results of an initial implementation. In *RO-MAN, 2010 IEEE*, pages 704–711, Sept 2010.

[25] Adriana Tapus, Andreea Peca, Amir Aly, Cristina Pop, Lavinia Jisa, Sebastian Pintea, Alina S. Rusu, and Daniel O. David. Children with autism social engagement in interaction with NAO, an imitative robot - a series of single case experiments. *Interaction Studies*, 13(3):315–347, 2012.

[26] Luthffi Idzhar Ismaila, Syamimi Shamsudina, Hanafiah Yussofa, Fazah Akhtar Hanapiah, and Nur Ismarrubie Zaharid. Robot-based intervention program for autistic children with humanoid robot NAO: Initial response in stereotyped behavior. In *Procedia Engineering*, volume 41, pages 1441–1447, 2012.

[27] Jillian Greczek, Amin Atrash, and Maja Matarić. A computational model of graded cueing: Robots encouraging behavior change. volume 374 of *Communications in Computer and Information Science*, pages 582–586, 2013.

[28] Esubalew Bekele, Julie A Crittendon, Amy Swanson, Nilanjan Sarkar, and Zachary E Warren. Pilot clinical application of an adaptive robotic system for young children with autism. *Autism*, 2013.

[29] Esubalew Bekele, Uttama Lahiri, Amy Swanson, Julie A. Crittendon, Zachary E. Warren, and Nilanjan Sarkar. A step towards developing adaptive robot-mediated intervention architecture (aria) for children with autism. *IEEE Transactiona on Neural Systems and Rehabilitation Engineering*, 21:289–299, Mar 2013.

[30] Patrizia Marti, Alessandro Pollini, Alessia Rullo, and T. Shibata. Engaging with artificial pets. In *Proc. Conference on European Association of Cognitive Ergonomics*, pages 99–106, 2005.

[31] Audrey Duquette, Henri Mercier, François Michaud, and Université De Sherbrooke. Investigating the use of a mobile robotic toy as an imitation agent for children with autism. In *In Proceedings International Conference on Epigenetic Robotics: Modeling Cognitive Development in Robotic Systems*, page 77, 2006.

[32] Kerstin Dautenhahn and Iain Werry. Issues of robot-human interaction dynamics in the rehabilitation of children with autism. In *Proc. International Conference on the Simulation of Adaptive Behavior*, pages 519–528, 2000.

[33] Ben Robins, Kerstin Dautenhahn, and Paul Dickerson. From isolation to communication: A case study evaluation of robot assisted play for children with autism with a minimally expressive humanoid robot. In *Advances in Computer-Human Interactions, 2009. ACHI '09. Second International Conferences on*, pages 205–211, 2009.

[34] Hideki Kozima, Marek Michalowski, and Cocoro Nakagawa. Keepon. *International Journal of Social Robotics*, 1:3–18, 2009.

[35] Hideki Kozima. Infanoid. In Kerstin Dautenhahn, Alan Bond, Lola CaÃśamero, Bruce Edmonds, and Gerhard Weiss, editors, *Socially Intelligent Agents*, volume 3 of *Multiagent Systems, Artificial Societies, and Simulated Organizations*, pages 157–164. Springer US, 2002.

[36] David Feil-Seifer and Maja Matarić. Robot-assisted therapy for children with autism spectrum disorders. In *Proceedings of the 7th international conference on Interaction design and children*, IDC '08, pages 49–52, New York, NY, USA, 2008. ACM.

[37] Francesca Caprino, Elena Laudanna, and Serenella Besio. Guidelines for using robots in education and play therapy sessions for children with disabilities. `http://www.iromec.org/fileadmin/pdf/D5.3_final.pdf/`. Accessed: 01-18-10.

[38] Patrizia Marti, Claudio Moderini, Leonardo Giusti, and Alessandro Pollini. A robotic toy for children with special needs: from requirements to design. In *Proc. IEEE International Conference on Rehabilitation Robotics*, pages 918–923, 2009.

[39] Ester Ferrari, Ben Robins, and Kerstin Dautenhahn. Therapeutic and educational objectives in robot assisted play for children with autism. In *Robot and Human Interactive Communication (RO-MAN 2009). The 18th IEEE International Symposium on*, pages 108–114, 2009.

[40] Ioannis A. Kakadiaris, George Passalis, Theoharis Theoharis, George Toderici, Ioannis Konstantinidis, and Najam Murtuza. Multimodal face recognition: Combination of geometry with physiological information. In *Proceedings, IEEE Conference on Computer Vision and Pattern Recognition*, pages 1022–1029, 2005.

[41] Tsuyoshi Matsukawa and Kiyoko Yokoyama. Visualizing physiological information based on 3DCG. *Forma*, 25:11–14, 2010.

[42] Armando Barreto, Jing Zhai, and Malek Adjouadi. Non-intrusive physiological monitoring for automated stress detection in human-computer interaction. In *Human-Computer Interaction*, volume 4796, pages 29–38. 2007.

[43] Ramya Murthy, Ioannis Pavlidis, and Panagiotis Tsiamyrtzis. Touchless monitoring of breathing function. In *Proceedings 26th Annual International Conference IEEE Engineering in Medical Biology Society, San Francisco, CA*, pages 1196–1199, 2004.

[44] Hamlyn G. Jones, Rachid Serraj, Brian R. Loveys, Lizhong Xiong, Ashley Wheaton, and Adam H. Price. Thermal infrared imaging of crop canopies for the remote diagnosis and quantification of plant responses to water stress in field. *Functional Plant Biology*, 36:978–989, 2009.

[45] Dana Kulic and Elizabeth Croft. Physiological and subjective responses to articulated robot motion. *Robotica*, 25(1), 2007.

[46] Wenxi Chen, X. Zhu, T. Nemoto, Y. Kanemitsu, K. Kitamura, and K. Yamakosh. Involuntary monitoring of breath rate and heart rate during sleep. In R. Magjarevic, J.H. Nagel, and Ratko Magjarevic, editors, *World Congress on Medical Physics and Biomedical Engineering 2006*, volume 14 of *IFMBE Proceedings*, pages 4050–4053. Springer Berlin Heidelberg, 2007.

[47] John T. Cacioppo, Louis G. Tassinary, and Gary G. Berntson. The respiratory system. In *Handbook of psychophysiology*, pages 235–244. Cambridge University Press, 2007.

[48] Phil Corbishley and Esther Rodriguez-Villegas. Breathing detection: Towards a miniaturized, wearable, battery-operated monitoring system. *Biomedical Engineering, IEEE Transactions on*, 55(1):196–204, Jan 2008.

[49] Paul Saulnier, Ehud Sharlin, and Saul Greenberg. Using bio-electrical signals to influence the social behaviours of domesticated robots. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*, HRI 2009, pages 263–264, New York, NY, USA, 2009. ACM.

[50] Laurianne McLaughlin. Engineering meets emotion: New robots sense human distress. In *http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3055998/Intelligent Systems, IEEE*, volume 18, pages 4–7, 2003.

[51] June Groden, Joseph Cautela, Stacey Prince, and Jennifer Berryman. The impact of stress and anxiety on individuals with autism and developmental disabilities. In *Behavioral Issues in Autism*, chapter 9. 1994.

[52] Olga Bogdashina. In *Sensory Perceptual Issues in Autism and Asperger Syndrome*, page 217. Jessica Kingsley Publishers : London UK, 2003.

[53] John W. Rohrbaugh, Erik J. Sirevaag, Naveen Singla, Joseph A. O'Sullivan, Mei Chen, Sean D. Kristjansson, Po-Hsiang Lai, and Alan D. Kaplan. Laser doppler vibrometry measures of physiological function: evaluation of biometric capabilities. *IEEE Transactions on Information Forensics and Security*, 5:449–460, 2010.

[54] A mini-robot with a long standoff capability to detect motion and breathing inside a compound. `http://www.tialinx.com/cougarfamily.html`, September 2011. Accessed: 09-01-11.

[55] Pramila Rani, Jared Sims, Robert Brackin, and Nilanjan Sarkar. Online stress detection using psychophysiological signals for implicit human-robot cooperation. *Robotica*, 20(6):673–685, 2002.

[56] Eugene Ferguson Greneker. Radar sensing of heartbeat and respiration at a distance with security applications. In *Proceedings SPIE, Radar Sensing Technology*, pages 22–27, 1997.

[57] Jin Fei and Ioannis Pavlidis. Thermistor at a distance: unobtrusive measurement of breathing. In *IEEE Transactions on Biomedical Engineering*, pages 988–998, 2010.

[58] Tom T. Shimabukuro, Scott D. Grosse, and Catherine Rice. Medical expenditures for children with an autism spectrum disorder in a privately insured population. *Journal of Autism and Developmental Disorders*, 38(3):546–552, 2008.

[59] Ben Robins, Kerstin Dautenhahn, Rene Te Boekhorst, and Aude Billard. Robots as assistive technology - does appearance matter? In *Proc. IEEE International Workshop on Robot and Human Interactive Communication*, pages 277–282, 2004.

[60] Rainer Lienhart and Jochen Maydt. An extended set of Haar-like features for rapid object detection. In *Proc. IEEE International Conference on Image Processing*, pages 900–903, 2002.

[61] Robert E. Schapire. The boosting approach to machine learning: An overview. In D. D. Denison, M. H. Hansen, C. Holmes, B. Mallick, and B. Yu, editors, *Nonlinear Estimation and Classification*. Springer, 2003.

[62] Gary R. Bradski. Computer vision face tracking for use in a perceptual user interface, 1998.

[63] Paul A. Viola and Michael J. Jones. Rapid object detection using a boosted cascade of simple features. In *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 511–518, 2001.

[64] Paul A. Viola and Michael J. Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154, 2004.

[65] Andre L. C. Barczak and Farhad Dadgostar. Real-time hand tracking using a set of co-operative classifiers based on haar-like features. *Res. Lett. Inf. Math. Sci.*, 7:29–42, 2005.

[66] Patrick Buehler, Mark Everingham, Daniel Huttenlocher, and Andrew Zisserman. Long term arm and hand tracking for continuous sign language tv broadcasts. In *Proceedings of the 19th British Machine Vision Conference. BMVC 2008*, pages 1105–1114, Sept 2008.

[67] Juan Wachs. Enhanced human computer interface through webcam image processing library - aGest.xml. `http://code.google.com/p/ehci/source/browse/trunk/data/aGest.xml?r=223`. Accessed: 03-14-11.

[68] Brian Scassellati, Henny Admonil, and Maja Matarić. Robots for use in autism research. *Annual Review of Biomedical Engineering*, 14:275–294, May 2012.

[69] Administration for children and families archives. `http://archive.acf.hhs.gov/programs/opre/ehs/perf_measures/reports/resources_measuring/res_meas_cdijj.html/`. Accessed: 10-06-13.

[70] Stephan Ehlers, Christopher Gillberg, and Lorna Wing. A screening questionnaire for asperger syndrome and other high-functioning autism spectrum disorders in school age children. *Journal of Autism and Developmental Disorders*, 29(2):129–141, 1999.

[71] Catherine Lord, Michael Rutter, Susan Goode, Jacquelyn Heemsbergen, Heather Jordan, Lynn Mawhood, and Eric Schopler. Austism diagnostic observation schedule: A standardized observation of communicative and social behavior. *Journal of Autism and Developmental Disorders*, 19(2):185–212, 1989.

[72] Social communication questionnaire. `http://portal.wpspublish.com/portal/page?_pageid=53,70432&_dad=portal&_schema=PORTAL/`. Accessed: 10-06-13.

[73] Roger Brown. *A First Language: The Early Stages*. Harvard University Press, 1973.

[74] Wendy L. Stone, Opal Y. Ousley, and Cynthia D. Littleford. Motor imitation in young children with autism: What's the object? *Journal of Abnormal Child Psychology*, 25(6):475–485, 1997.

[75] Brooke Ingersoll and Katherine Lalonde. The impact of object and gesture imitation training on language use in children with autism spectrum disorder. *Journal of Speech Language Hearing Research*, 53(4):1040–1051, 2010.

[76] Kathleen T. Williams. *AGS Expressive Vocabulary Test*. Expressive Vocabulary Test. American Guidance Service, 1997.

[77] Svein Eikeseth, Tristram Smith, Erik Jahr, and Sigmund Eldevik. Intensive behavioral treatment at school for 4-7-year-old children with autism: A 1-year comparison controlled study. *Behavior Modification*, 26:49–68, January 2002.

[78] Daniel B. LeGoff and Michael Sherman. Long-term outcome of social skills intervention based on interactive LEGO play. *Autism: The International Journal of Research and Practice*, 10:317–329, 2006.

[79] Terry B. Hancock and Ann P. Kaiser. The effects of trainer-implemented enhanced milieu teaching on the social communication of children with autism. *Topics in Early Childhood Special Education*, 22:39–54, 2002.

[80] J.D. Montagu and E.M. Coles. Mechanism and measurement of the galvanic skin response. *Psychological Bulletin*, 65(5):261–279, 1966.

[81] Paul D. Tyson. Task-related stress and EEG alpha biofeedback. *Biofeedback and Self-Regulation*, 12(2):105–119, 1987.

[82] Maria Vittoria Pitzalis, Filippo Mastropasqua, Francesco Massari, and Andrea Passantino et. al. Effect of respiratory rate on the relationships between RR interval and systolic blood pressure fluctuations: a frequency-dependent phenomenon. *Cardiovascular Research*, 38(2):332–339, 1998.

[83] Yuri Masaoka and Ikuo Homma. Anxiety and respiratory patterns: their relationship during mental stress and physical load. *International Journal of Psychophysiology*, 27(2):153–159, 1997.

[84] Michelle A. Cretikos, Rinaldo Bellomo, Ken Hillman, and Jack Chen et. al. Respiratory rate: the neglected vital sign. *The Medical Journal of Australia*, 188(11):657–659, 2011.

[85] Paolo Palatini. Need for a revision of the normal limits of resting heart rate. *Hypertension*, 33(2):622–625, 1999.

[86] Peter T. Macklem. The mechanics of breathing. *American Journal of Respiratory and Critical Care Medicine*, 157(4), 1998.

[87] Thomas Williams, Colin Kelley, and et al. Gnuplot 4.5: an interactive plotting program. Accessed: 06-07-11.

[88] C. Sidney Burrus, Ramesh A. Gopinath, and Haitao Guo. *Introduction to Wavelets and Wavelet Transforms: A Primer*. Prentice Hall, 1997.

APPENDIX A

STUDY PROTOCOL

Effectiveness of CHARLIE the Robot for Improving Verbal and Nonverbal Skills in Children with Autism

Principal Investigator: Laura Boccanfuso

Faculty Mentor: Jason M. O'Kane

Table of Contents

1 Abstract

This research presents a new, autonomous, interactive game-playing robot named CHARLIE. The robot is equipped with a head and two arms, each with two degrees of freedom, and a camera. We trained a human hands classifier and used this classifier along with a standard face classifier to create two autonomous interactive games : single-player ("Imitate Me, Imitate You") and two-player ("Pass the Pose".) Further, we implemented a third setting in which the robot is teleoperated by remote control. This research will makes three novel contributions: (1) A new low-cost robot design for use in autism therapy, (2) three autonomous, interactive robot games and (3) a new hand classifier created for performing hand detection and trained for use with the interactive games.

CHARLIE is physically designed to be aesthetically appealing to young children between three and six years of age. The hard, wood and metal robot body is covered with a bright green, fuzzy material and additional padding so that it appears toylike and soft. Additionally, several structural features were included to ensure safety during interactive play and to enhance the robustness of the robot. Because children with autism spectrum disorder (ASD) often enjoy exploring new or interesting objects with their hands, the robot must be able to withstand a moderate amount of physical manipulation without causing injury to the child or damaging the robot or its components. CHARLIE plays three distinct interactive games that are designed to be entertaining to young children, appeal to children with varying levels of communication and social skill and promote two fundamental requirements for communication : imitation and turn-taking.

2 Background and Significance

The use of robots for cooperative work with humans is becoming increasingly pervasive across an ever widening range of disciplines. Medical procedures using robots are reported to be less invasive, result in faster recovery times and are estimated to have nearly tripled from 2007 to 2010 [1]. They have been employed for use in post-stroke rehabilitation [2], as assistive feeding systems for the physically handicapped [3] and in therapeutic roles such as robotic pets for the elderly in nursing homes [4]. Engineers across multiple disciplines have capitalized on the unique qualities of robots to perform autonomously and predictably and repeat mechanical tasks consistently. These characteristics also make robots well suited as part of an early intervention strategy for many autistic children who tend to perceive them as nonthreatening and intrinsically interesting.

Robots have been used to effectively engage autistic children in interactive game playing and research has demonstrated that robot-assisted autism therapy promotes increased speech and increased child-initiated interactions in children with Autism Spectrum Disorder (ASD) [5, 6]. However, more research is needed to develop definitive paradigms for autonomous robot designs that will most benefit children with autism.

According to a 10-state study conducted by the Center for Disease Control (CDC) [7] the number of diagnosed autism cases increased an average of 57% from 2002 to 2006. Not only does this translate to a growing population of autistic children but it also means that existing resources used to treat and care for children with autism are under greater strain. Further, because of the added expense of therapy and specialized medical care, the cost of raising an autistic child in the United States is estimated to be between 8.5 to 9.5 times greater than raising a typically developing child. This additional financial burden may mean that some families have to choose whether to incur significant debt to get the proper care for their child or limit the amount of therapy their child receives. Even though robots have been proven to be effective for promoting communication with some autistic children, there are few existing robots currently in use for autism therapy and those that do exist are cost prohibitive for widespread use.

3 Specific Aims

This research focuses on the design and development of a low-cost, interactive robot that is simultaneously capable of playing imitative games autonomously with a child and recording vital interaction information for each session. The two most important questions motivating this research are: (1) What kinds of simple, low-cost robot designs are effective for promoting human-to-human interaction with autistic children? and, (2) How much improvement in verbal communication and motor imitation can be observed after the use of a robot during autism therapy?

Figure 3.1: CHARLIE poses. From left to right : Left hand high. Right hand high. Both hands high. Neutral. Peek-a-boo.

Basic turn-taking and imitation skills are imperative for effective communication and social interaction [8]. Research has shown that interactive games using turn-taking and imitation have yielded positive results with autistic children who have impaired communication or social skills [9]. In [10] we present research in which we designed and built a toy-like robot with face and hand detection capabilities to autonomously engage autistic children in interactive games using imitation and turn-taking skills. Examining the effectiveness of the robot and game designs will provide critical insight as to the viability of using this robot as a widely-accessible tool for autism therapy.

4 Research Design, Methods and Data Analysis

The physical design of CHARLIE addresses three major objectives. First, the outward appearance of the robot was designed to be pleasant and cute so as to invite the attention of young children with ASD and avoid being intimidating to the greatest extent possible. Second, we carefully designed the robot structure so as not to allow the robot, nor its constituent parts, to harm the child interacting with it. Third, we made the robot more robust by adding features to protect its mechanical components and allow children to explore and interact more freely with the robot without excessive concern for the physical integrity of the robot.

Our approach to the design of CHARLIE's interactive games is based on the integration of robot and game designs that are known to be effective with autistic children. Each game was designed to be entertaining to young children and to promote two fundamental requirements for communication : imitation and turn-taking. Further, we created three different types of games to appeal to children with varying levels of communication and social skill. For children who are reluctant to play with a completely autonomous robot or for those who would benefit from a period of exploration before they begin playing the interactive games, the robot can be teleoperated using a simple remote control. For those who are ready to play directly with the robot, but who are not necessarily ready to play a cooperative game with another child, the single-player interactive game is available. Finally, a two-player interactive game was created to appeal to those children who have established some level of simple imitation and turn-taking but could use more practice with these skills using the robot as a social mediator.

4.1 Recuitment procedures and location of study

Recruitment will take place at local elementary schools and area speech therapy clinics. The administration of pre-test questionnaires and data collection will be conducted at two sites: (1) Med Park 15 will be used to administer tests to measure cognitive ability (Vineland Adaptive Behavior Scale) and to confirm autism diagnosis (Social Communication Questionnaire), and (2) Tests to motor imitation ability (Motor Imitation Scale) and language ability (Unstructured Imitation Assessment, Expressive Vocabulary Test and Mean Length of Spontaneous Utterance Determination) and all therapy with the robot will be conducted at the USC Speech and Hearing Research Center.

4.2 Physical design

We deliberately designed the outward appearance of CHARLIE with the end-user in mind. Recent research has shown that robots with a simple interface are generally better received initially by children with autism, than robots with a more realistic, human-like appearance [11]. The implication is that low-tech robots, when designed appropriately for the particular needs of the child(ren) with ASD they will serve and the context in which they will be

Figure 4.1: [top] Snap off arm. [bottom] Snap off head.

used, can be used effectively to teach and promote social skills. In addition to the low-cost design, CHARLIE'S physical appearance is intended to be toylike to create a friendly and approachable outward appearance and to more easily attract the attention of a child.

4.2.1 Basic hardware components

CHARLIE's hardware includes 6 servos, 3 pan-tilt platforms, an 8 channel servo controller, a consumer-grade web cam, and 2 D-cell battery packs. The robot's body is padded for safety, and its outer surfaces are covered with a bright green, fur-like material to achieve a non-threatening appearance. During active game play the child's attention is typically focused near CHARLIE's hands, so one LED is embedded in each of the hands to provide positive feedback during interactive games. A speaker is also included in the CHARLIE's body in order to provide optional auditory instructions for playing interactive games and positive feedback. Exclusive of the computing hardware, the retail cost of the robot's components is approximately 200 USD. In a production version of this robot, a computer could be integrated into the robot's body, or users could connect via USB to a standard laptop or desktop PC.

4.2.2 Features for robustness and safety

In general, children are curious about robots and many enjoy exploring the physical features of the robot as much as interacting with it. This can present hazards to both the child and to the robot's mechanical hardware. In order to minimize potential hazards and to improve the robustness of the robot, we included two characteristics in the robot's design. First, the body of the robot is secured to a platform that may be strapped to a desk or table. Immobilizing the robot in this way prevents the child from being able to pick up

the robot and potentially harm him/herself, others in the room or the robot itself. Second, the arms and head of the robot are attached to the robot's body using snap fasteners so that excessive force will not cause damage to the servo motors, but will instead allow that piece to snap off. Furthermore, allowing the arms and head of the robot to detach, affords the child more continuous free play since there will be less concern over the child's safety and the integrity of the robot's hardware. As described in the IROMEC study [12], while the adult must fulfill a more active role for promoting play skills with children with ASD, "much of the literature on childhood play emphasizes the importance of free play and the need to interfere as little as possible in the child's actions, thus underscoring the creative aspects that in essence cannot be controlled or oriented." We expect that longer, uninterrupted interactions will maximize the opportunity for each child to benefit from each session.

4.3 Interactive software design

We used the Open Source Computer Vision Library (OpenCV) [13], a cross-platform library for real-time computer vision applications, for training the hand classifier and for the implementation of hand and face detection. OpenCV provides a facility for object detection based on an extended set of Haar-like features [14]. Informally, this method works by screening small portions of an image for visual characteristics of the target object. To train a classifier to identify a specific class of objects, OpenCV uses Adaptive Boosting (AdaBoost) [15] to create a cascade of boosted classifiers defined over these features. We then included the resulting hand classifier along with a standard OpenCV face classifier to detect user hands, track the user's face and provide position information for managing three interactive games. In the first game the robot waits for the child to initiate an interaction by raising one or both hands. In the second game, the robot initiates interactions. The primary objective of our game designs is to increase attention, promote turn-taking skills and encourage child-led verbal and non-verbal communication through simple imitative play.

4.3.1 Face detection and tracking

We relied upon the frontal face classifier provided by OpenCV (more specifically, a cascade of boosted classifiers working with Haar-like features) for face detection. Haar-like features are used as an abstraction of RGB pixel values for object detection since image intensities are computationally expensive to work with. Each feature type is used to screen a given portion of an image for different characteristics of the target object. The extended sets of rectangular Haar-like features used for the face and hand detectors described in this paper are applied to assess whether a particular rectangular portion of a video frame contains a face or hand by summing the pixels contained within the rectangle and determining whether it matches the characteristics of the target object as defined by the classifier.

To make the overall program as efficient as possible, we implemented a face tracking algorithm instead of repeating the computationally intensive detection process for each

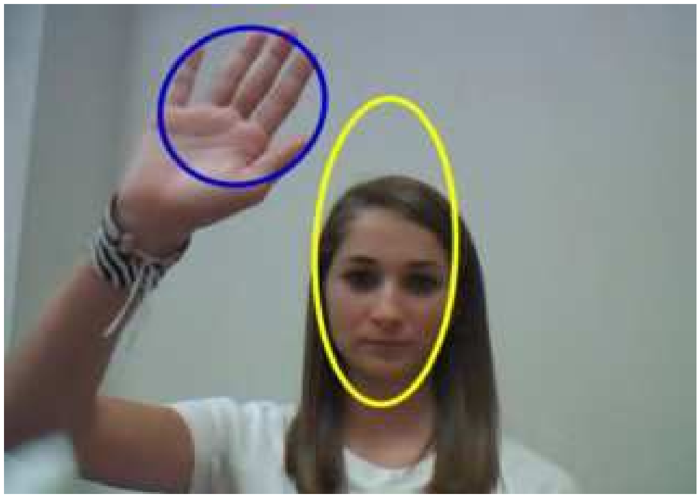

Figure 4.2: Face and hand detection.

frame. Face tracking was accomplished using the Continuously Adaptive Mean Shift (CAMSHIFT) algorithm [16]. CAMSHIFT incorporates the MEANSHIFT algorithm which is based on a nonparametric technique for climbing density gradients to find the peak of the probability distribution of the position of a given target object. For face tracking, this translates to identifying the center of the target color distribution in a given video frame. In order to make face tracking fast and relatively robust (and appropriate for use in real-time tracking applications), we used the CAMSHIFT technique. This tracking method improves performance by eliminating the need to repeat the face detection for each frame of the video. To overcome errors resulting from drift in the CAMSHIFT algorithm, the robot periodically repeats the full face detection process. In the event that the robot cannot detect the face, the robot head is reset to a neutral position and searches outward in an increasingly larger area.

Whereas face detection is a well-studied problem [17, 18], and effective face classifiers are freely available through OpenCV, robust and real-time hand detection in diverse environments is a topic of continuing research.

4.3.2 Hand classifier and hand detection

Numerous approaches for developing robust hand detectors have been explored [19, 20], but the resulting classifiers have not been made available to the research community. Further, some hand classifiers that are freely available such as the gesture letter "A" detector by Juan Wachs from the Ben Gurion University of the Negev, Israel and Washington Hospital Center [21], are too narrow in scope for use in this context and others are not accurate or efficient enough for our application. In order to implement a hand detector suitable for our purposes, a new hand classifier was trained to detect hands in various lighting conditions,

Figure 4.3: Sample images used to train the hand detector. [top] Positive examples. [bottom] Negative examples.

rotations, scales and finger positions. Approximately 750 positive hand images of various size, color and position and approximately 3300 negative images were collected and cropped to a uniform pixel size of 40x40. Representative examples are shown in Figure 4.3. To create additional positive training samples representing variations in lighting, rotation and scale, ten distortions were applied to 100 of those samples, yielding a total of approximately 1750 positive hand samples. The resulting vector files were then merged and the AdaBoost training procedure was initiated using the combined vector file representing all positive hand samples and the complete set of negative samples. A twenty-stage cascade was trained on these samples, yielding an error rate on the training set approaching zero.

4.3.3 Interactive game design

Research in robot-assisted autism therapy typically emphasizes specific objectives for ideal human-robot interaction including an increased attention span, eye contact, proactive interaction with the robot initiated by the child, verbal and non-verbal cues, turn-taking, imitative game playing and overall use of language.

First, we defined the play scenario in terms of : (1) a main target group, (2) a play type, (3) actors involved, (4) a setting, and (5) the duration of the play activity. The main target group consists of a small group of children ages 3 to 6 who have been diagnosed with autism and have documented communication deficiencies. The play type consists of a very simple game of imitation with a basic set of rules and is designed to engage one teacher or one child at a time. The tests take place in a closed classroom, where both the child and teacher are seated across from the robot and the robot will be seated atop and securely attached to a nearby desk so that the robot's head is at approximately the same height as the

child's. The duration of the play activity is variable. The length of a typical session with the robot is based on the normal amount of session time allotted for that particular child, the perceived benefit of the robot to the child's development and the child's interest in the robot.

Second, we prepared a detailed description of how CHARLIE is introduced to each child and how play proceeds during the first and subsequent sessions. Prior to introduction, a baseline for communication skills and developmental ability is established for each child using assessment information provided by the child's teacher. At the first meeting, the teacher introduces CHARLIE and explains and/or demonstrates how to play the imitation game. The teacher then invites the child to play with robot and provide guidance, when necessary. For children who prefer to examine the robot and learn about its capabilities independently, the teacher assumes a more passive role, as an observer and guide.

Third, we identified measures of success using the baseline communication skills identified prior to the child's first session. Initially, the child's level of interest in CHARLIE is noted in addition to any specific robot characteristics that are especially interesting to the child. During each session, communication between the child and robot, and the child and teacher is documented by the teacher or researcher. Because the robot measures successful imitations between the robot and child it is not necessary to document these interactions, but other nonverbal and verbal communication occurring during the session is noted for subsequent analysis. Measures of success and user information collected during an interactive game can be used to assess the child's readiness for more advanced, child-initiated games such as collaborative group play and story-telling.

Ultimately, we designed and implemented two additional interactive games to appeal to children with ASD of a wider range of ability and skill. The original game developed is a single-player game which engages a child in a game called "Imitate Me, Imitate You". In this game, the child may either initiate a pose for the robot to imitate ("Imitate Me") or the child may follow the robot's pose ("Imitate You"). The single-player game is intended for the child with ASD who is comfortable interacting with an autonomous robot but who may not be ready for turn-taking with another child.

Single-player "Imitate Me, Imitate You" The "Imitate Me, Imitate You" game is detailed in Figure 4.4 and consists of two primary modes: passive and active. Within each of the two modes, there are five poses: neutral (both hands down), left hand raised, right hand raised, both hands raised and peek-a-boo, as shown in Figure **??**. In order to give the child initial control over the robot's actions, the default robot state is the passive game mode. Once the robot detects and begins tracking the child's face and hands, the robot indicates that it is ready to interact by moving to the neutral pose and blinking the LEDs in its hands three times. The robot then immediately enters the passive game mode and waits for the child to initiate a game by raising one or both hands. As the child's hand movements are detected, the robot responds by imitating the child's hand positions and lighting the LED in the corresponding hand while simultaneously detecting any additional hand movements. If ten seconds elapse without any detected hand movement, the robot will transition to the

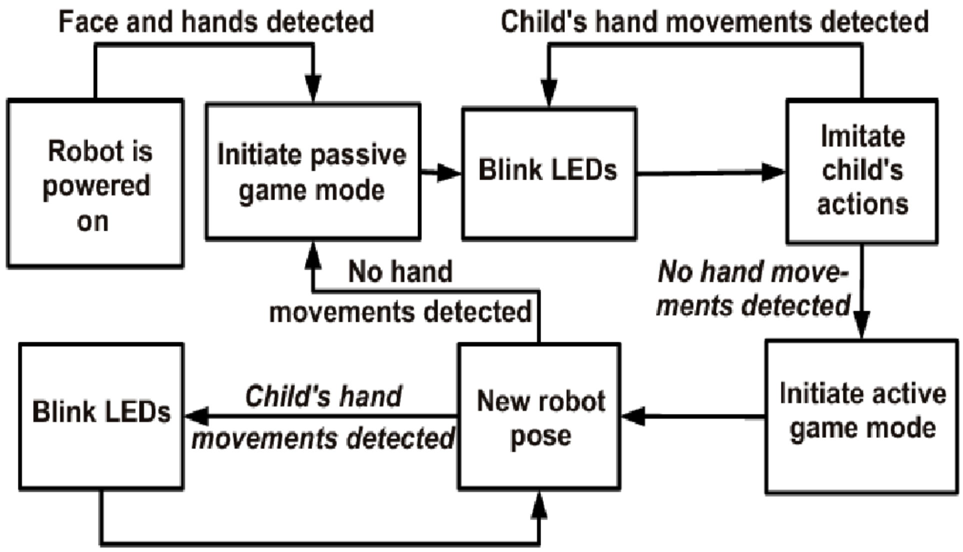

Figure 4.4: State diagram for CHARLIE's "Imitate Me, Imitate You" autonomous interactive game.

active game mode.

During the active game mode, the robot initiates a new game and attempts to engage the child by raising or lowering one or both arms, or beginning a game of peek-a-boo. Each pose assumed by the robot in the active game state is selected randomly in order to avoid repetitive patterns of poses. When a positive outcome is detected (the child successfully imitates the robot's pose), positive sensory feedback is generated by the robot. A positive sensory response entails the robot lighting a small LED in the hand corresponding to the raised hand or hands of the imitated pose. As with the passive game mode, the robot will wait ten seconds for the child's response. If ten seconds elapse and a positive response has not been detected, the robot will transition back to the passive game mode, waiting again for the child to initiate a new game.

Two-player "Pass the Pose" The second interactive game is a two-player game described in Figure 4.5 called "Pass the Pose". In this game, two players interact directly with the robot and indirectly with one another. With the optional sound enabled, the "Pass the Pose" game works as follows: Game play begins with CHARLIE describing how to play "Pass the Pose" and asking the first player (seated to the right of the robot) to assume a pose. Once she has detected the pose, CHARLIE indicates that she has learned the pose by saying "Ok, I got it. Now let me try", turns to the second player (seated to the left of the robot), asks the child to follow her and then assumes the same pose learned from the first player. If the second player successfully imitates the pose assumed by CHARLIE, she responds by saying "You got it!", claps her hands and giggles. If the player does not imme-

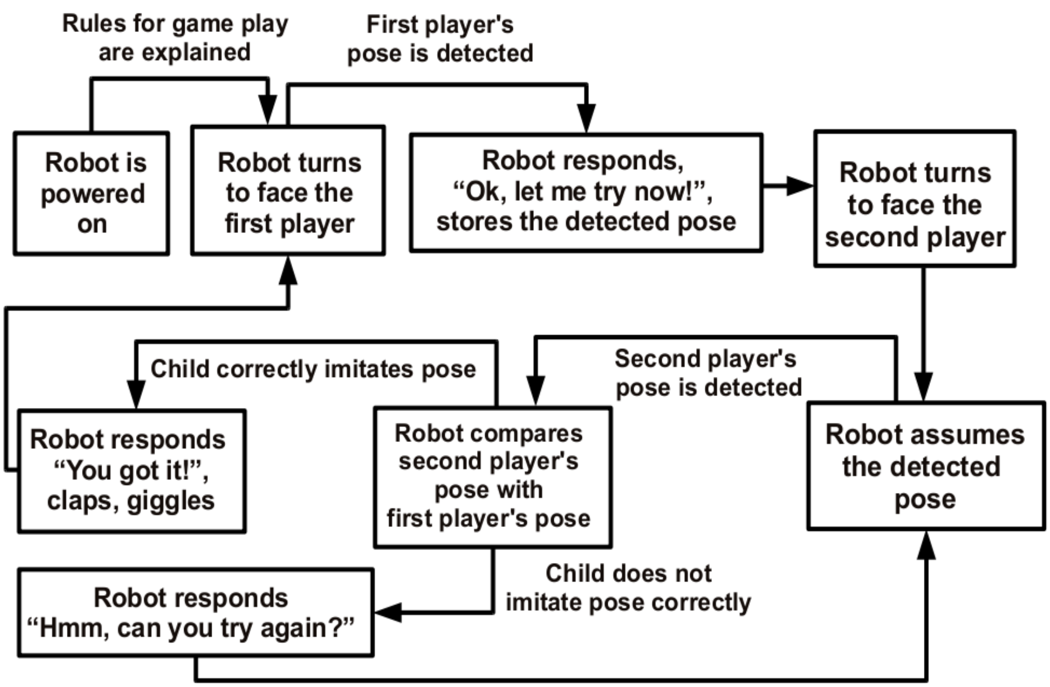

Figure 4.5: State diagram for CHARLIE's "Pass the Pose" autonomous interactive game.

diately imitate the correct pose, CHARLIE will ask the child to try again. If the child does not correctly assume the pose after three tries, the robot asks the current player to initiate a new pose and the game continues, this time with the second player initially "passing" the pose to the robot.

If the sound is disabled, we expect that the teacher, therapist or parent will describe how to play the "Pass the Pose" game. When the players are ready the teacher will start the game and CHARLIE will turn to the first player and wait for the child to assume a pose. Once CHARLIE has detected the pose, she turns to the second player and assumes the same pose. If the child correctly imitates the pose, CHARLIE claps her hands and waits for the second player to initiate the next pose. If the second player does not correctly imitate the pose, CHARLIE lowers her head and shakes it slowly from side to side. Should the child fail to imitate the pose correctly after three tries, CHARLIE resumes a neutral position and waits for the second player to start a new game. This two player game is ultimately designed to promote shared attention and cooperative play. We anticipate that the "Pass the Pose" game will be most useful for children who have already demonstrated some level of proficiency with turn-taking and imitation and who are able to play a game with a simple set of rules.

Teleoperation In addition to the two autonomous games, we developed and implemented software that allows for the robot to be teleoperated so that when a button is pushed on the remote, the player is given complete control over CHARLIE's limbs and head. While each of the four push buttons on the remote correspond to specific pre-programmed poses, the two joystick buttons provide continuous control for the movement of each arm and a single directional button allows for continuous control of the head. We expect this game play to be useful for the child with ASD who may be initially wary or hesitant to interact with the robot. By temporarily disabling the robot's autonomous actions, the child is given the freedom to learn about CHARLIE's various capabilities at his or her own pace.

4.4 Study procedures

If a child meets the criteria to participate in the study, there are two initial meetings that will be scheduled prior to the beginning of the study:

(a) The first meeting will take place at MedPark 15. The parent will be asked to complete a series of questionnaires without the child present. At that time, the parent will complete three measures that will help to establish the baseline for the child:

- The Vineland Adaptive Behavior Scales (assesses the child's behavior and estimate the child's cognitive age based on the parent's report of their performance)

- The Social Communication Questionnaire (SCQ)

(b) During the second meeting (which will take place at the USC Speech and Hearing Research Center), the parent and the child will be introduced to CHARLIE. We will evaluate the child using the following two tests:

- Unstructured Imitation Assessment (Version A)

- Motor Imitation Scale: 16-item assessment of motor imitation, especially designed for children with autism (Version A)

- Expressive Vocabulary Test (Second edition) (Form A)

- Completion of the Mean Length of Spontaneous Utterance Determination

Experiment procedures

Following the initial evaluations, we will schedule 30-minute sessions, twice a week (semi-weekly) for a total of 6 weeks.

(a) Each session will involve the speech therapist, a graduate student and the robot, CHARLIE.

(b) Introducing the robot to each child and will follow these general guidelines and will incorporate the following therapy goals:

Phase I: The robot will be situated in the room where therapy will take place. For the first few sessions (and at the therapist's discretion) the robot will be placed in stationary mode to allow the child the opportunity to physically explore the robot and its components before introducing movement.
(1) On arriving and leaving each session, the subject will briefly make eye contact with clinician, CHARLIE and the researcher as part of his/her greeting. The child will be ready for Phase II after the child has been observed to:
(2) Approach CHARLIE
(3) Touch CHARLIE
(4) Move CHARLIE's arms

Phase II: The child, with the therapist's guidance (if required) will be given the opportunity to control the robot's arms and head with a remote control.
(1) On arriving and leaving each session, the subject will briefly make eye contact with clinician, CHARLIE and the researcher as part of his/her greeting.
(2) During follow directions task, subject will point/operate remote/follow direction in order to lead CHARLIE through an activity at least once during the session. We will be ready to approach Phase III after the child has been observed to:
(3) Use the remote control to make CHARLIE move

Phase III: The child, with the therapist's guidance (if required) will be given the opportunity to play music and practice movement along with the robot. The robot can play If you're happy and you know it and The wheels on the bus with hand and head movements.
(1) On arriving and leaving each session, the subject will briefly make eye contact with clinician, CHARLIE and the researcher as part of his/her greeting.
(2) During follow directions task, subject will point/operate remote/follow direction in order to lead CHARLIE through an activity at least once during the session.
(3) During song activity, subject will participate in fingerplays/gestures with CHARLIE for 80% of opportunities.
(4) Once CHARLIE has imitated subject's movement, subject will continue to move/interact with CHARLIE through X turns (X to be determined from performance on baseline/previous session). The child will be ready to approach Phase IV after the child has been observed to:
(5) Respond to song with appropriate fingerplay/gesture
(6) Move in response to CHARLIE's prompt/action

Phase IV: The child, with the therapist's guidance (if required) will be given the opportunity to play imitation games with the robot. One-on-one games include just the child and the robot. The Pass the pose game includes the therapist, the child and the robot.
(1) On arriving and leaving each session, the subject will briefly make eye contact with clinician, CHARLIE and the researcher as part of his/her greeting.
(2) During follow directions task, subject will point/operate remote/follow direction in order to lead CHARLIE through an activity at least once during the session.
(3) During song activity, subject will participate in fingerplays/gestures with CHARLIE for 80% of opportunities.
(4) Once CHARLIE has imitated subject's movement, subject will continue to move/interact with CHARLIE through X turns (X to be determined from performance on baseline/previous session).
(5) When offered a choice of activities, subject will clearly make his/her performance known to others in session for 80% of trials.
(6) Throughout the therapy session, subject will cooperate with a turn-taking task with CHARLIE, parent and/or clinician through (2) turns (change this number as client progresses) each. We will be ready to approach Phase V after the child has been observed to:
(7) Imitate the robot movements on 80% of trials
(8) Imitate movements with another person in the intervention room

Phase V: The child will be given the opportunity to select from various modes of play with the robot.
(1) On arriving and leaving each session, the subject will briefly make eye contact with clinician, CHARLIE and the researcher as part of his/her greeting.

(2) During follow directions task, subject will point/operate remote/follow direction in order to lead CHARLIE through an activity at least once during the session.
(3) During song activity, subject will participate in fingerplays/gestures with CHARLIE for 80% of opportunities.
(4) Once CHARLIE has imitated subject's movement, subject will continue to move/interact with CHARLIE through X turns (X to be determined from performance on baseline/previous session).
(5) When offered a choice of activities, subject will clearly make his/her performance known to others in session for 80% of trials.
(6) Throughout the therapy session, subject will cooperate with a turn-taking task with CHARLIE, parent and/or clinician through (2) turns (change this number as client progresses) each.
(7) During interactive games and songs with CHARLIE, subject will participate in a structured reciprocal play routine for (2) minutes on (3) occasions (change number as client progresses).

Phase VI: If the child moves smoothly through the previous 5 sessions, then on the 6th session s/he will be given the opportunity to select from various modes of play with the robot and any member of the research staff.
(1) On arriving and leaving each session, the subject will briefly make eye contact with clinician, CHARLIE and the researcher as part of his/her greeting.
(2) During follow directions task, subject will point/operate remote/follow direction in order to lead CHARLIE through an activity at least once during the session.
(3) During song activity, subject will participate in fingerplays/gestures with CHARLIE for 80% of opportunities.
(4) Once CHARLIE has imitated subject's movement, subject will continue to move/interact with CHARLIE through X turns (X to be determined from performance on baseline/previous session).
(5) When offered a choice of activities, subject will clearly make his/her performance known to others in session for 80% of trials.
(6) Throughout the therapy session, subject will cooperate with a turn-taking task with CHARLIE, parent and/or clinician through (2) turns (change this number as client progresses) each.
(7) When presented with communication opportunities by clinician, subject will use gestures, vocalizations, or verbalizations for a variety of communicative intents on 80% of opportunities presented.
(8) When subject desires to initiate, change or discontinue activities within the last session, subject will make eye contact with appropriate clinician, the researcher or parent before communicating the message.

4.5 Data collection

There are two distinct types of data that will be collected, pertaining to nonverbal and

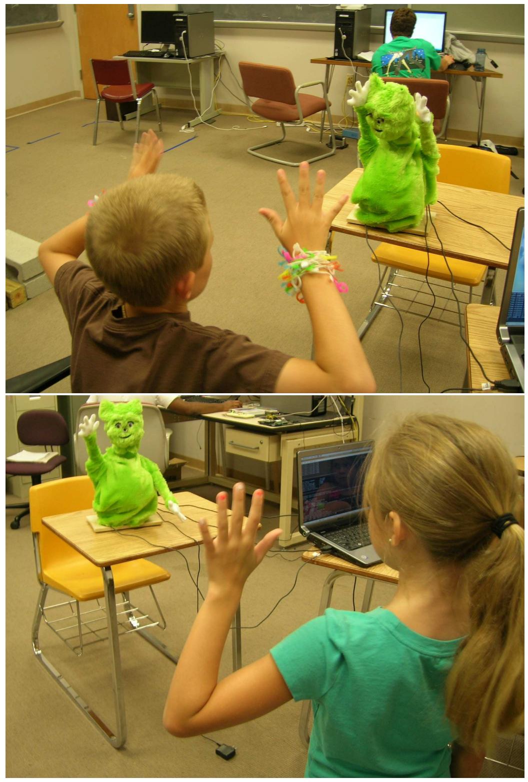

Figure 4.6: Children Interacting with CHARLIE.

verbal ability. Nonverbal ability will be measured using two sets of the Motor Imitation Scale in order to measure the motor imitation skill improvement from the beginning of the experiment to the post-test period. Verbal ability will be measured using each of the following tests: (1) Unstructured Imitation Assessment, (2) Expressive Vocabulary Test and, (3) Mean Length of Spontaneous Utterance Determination. One set of the verbal assessments will be administered prior to therapy with the robot and a second set will be used to measure improvements gained during the course of therapy.

4.6 Data analysis

Once the study is complete, a statistical analysis will be completed using the pre- and post-test nonverbal and verbal measurements. This study will last approximately six months. We intend to recruit 25 study participants, each of whom will receive therapy with the robot twice a week for six weeks. Since one therapy room at the USC Speech and Hearing Center is only available for 3 hours twice a week, the study will likely require a full six months to complete.

5 Human Subjects

Target Population We are looking for children between 3 and 6 years old who have a formal diagnosis of autism by a qualified professional. Additionally, candidates for the study will have speech or language delays and deficiencies as documented by a speech assessment or therapist. We expect to include up to 25 children in the study. Selected study participants will be chosen based on the information provided in prescreening questionnaire (Appendix A) as it complies with the requirements above.

Recruiting Plans Flyers will be posted at local area pediatrician's offices, speech therapy clinics and diagnostic clinics (Appendix B.) Additionally, a take home flyer will be provided to local area elementary schools (Appendix C.) Prior approval will be sought from each organization or place of business before flyers are posted or distributed.

In order to determine whether the child is a good candidate for the study, the researcher will administer the Prescreening Questionnaire to the prospective participant's parent over the phone. Upon hearing about the study (via flyer, teacher, therapist or physician), interested parents will contact the researcher by phone and the researcher will ask the questions outlined in the prescreening questionnaire.

The questions on the prescreening form will be used to confirm that (1) a formal diagnosis of autism has been received, (2) the child's language ability is delayed for their chronological age and to what degree (3) the child's nonverbal ability is delayed for their chronological age and to what degree, (4) the child does not have any diagnosed hearing disorder and, (5) the child's current therapy schedule will allow for additional therapy sessions with the robot.

If the child is not a good candidate, the information collected over the phone during the prescreening process will be destroyed (shredded). If the child is a good candidate, a numeric identifier will be assigned to each child and the completed prescreening questionnaire will be stored in a secure office, in a locked drawer, to which the researcher has exclusive access.

Existing Data/Samples Any prior, formal diagnoses of autism or speech and language delays by a licensed professional will be requested upon selection for the study. Copies of these documents will be stored in a locked office in a locked drawer which is only accessible by the researcher. All identifiers will be removed from the data and replaced with a participant number. A password-protected file with the corresponding names and participant numbers will be stored on the researcher's laptop computer.

Consent Parents of autistic children identified as potential study participants will be provided a letter, including the consent form (Appendix D) and the the Overview of Study Procedures (Appendix E). Parents will be asked to read the consent at the first face-to-face meeting and sign the form.

Potential Risks There are very few risks associated with participating in this research except a slight risk of breach of confidentiality, which remains despite steps that will be taken to protect your child's privacy. Each videotaped session will be catalogued using the date and a unique identifier assigned to the individual child. The real names of children par-

ticipating in the study and their respective identifiers will be kept on a password-protected computer in a password-protected file.

Because autistic children are individuals with widely varying interests, abilities and personalities, it is possible that while some children may find the robot cute and interesting, others may be fearful of it or act aggressively towards it. If a child should act/show fear or distress, use of the robot will be discontinued and/or removed from the room.

Potential Benefits Taking part in this study may benefit your child directly. The robot will be used as a tool during play to engage the autistic child and is intended to promote verbal and nonverbal human-to-human communication. Results from this research may also help us understand how to better design robots that can be used by therapists, teachers and parents to help promote communication skills in autistic children.

Confidentiality Participation will be confidential. A unique identifier will be assigned to each participant at the beginning of the project. This number will be used on project records rather than your name, and no one other than the researcher will be able to link your information with your name. Study records/data will be stored in a password protected laptop computer in a password protected file. The results of the study may be published or presented at professional meetings, but your identity will not be revealed.

In rare cases, a research study may be evaluated by an oversight agency, such as the USC Institutional Review Board or the U.S. Office for Human Research Protections. If this occurs, records that identify you and the consent form signed by you may be inspected so that they may evaluate whether the study is properly conducted and the rights of participants were adequately protected.

Excerpts from the video may be used for demonstrating the viability of the robot as a therapeutic tool at research conferences and still frames selected from specific sessions may be used in a publication of this research in conference proceedings or journal article(s), but the name(s) of those children appearing in the photograph(s) will be kept strictly confidential. If the parent wishes to include the child in the study, but exclude him/her from appearing in any video excerpt or photograph to be made publicly available, they will be given the choice of opting out.

Compensation There will be no compensation for participation in this study.

Withdrawal Participation in this study is voluntary. The parent is free to decline participation or to withdraw their child at any time, for whatever reason, without negative consequences. In the event that the parent does withdraw their child from this study, the information already provided will be kept in a confidential manner.

There are a few circumstances under which the subjects participation may be terminated without his/her consent.

Your child may be dismissed from the study without your consent for various reasons, including the following: Your child is continually disinterested in playing with the robot Your child experiences increased or significant distress or anxiety because of the robot Your child is physically aggressive toward the robot or another person present because of the robot The interventionist deems that the robot is no longer beneficial to your child If the study sponsor decides to stop or cancel the study. If the investigators otherwise believe

that it is not in your child's best interest to continue in the study.

References

[1] Gabriel Barbash and Sherry Glied. New technology and health care costs – the case of robot-assisted surgery. *The New England Journal of Medicine*, 303:701–704, AUG 2010.

[2] P.S. Lum, C.G. Burgar, P.C. Shor, M. Majmundar, and M. Van der Loos. Robot-assisted movement training compared with concentional therapy techniques for the rehabilitation of upper-limb motor function after stroke. *Archives of Physical Medicine and Rehabilitation*, 83(7):952–959, July 2002.

[3] F. Hiramatsu S. Ishii, S. Tanaka. Meal assistance robot for severely handicapped people. In *Proceedings IEEE International Conference on Robotics and Automation*, volume 2, pages 1308–1313, May 1995.

[4] Kaoru Inoue, Kazuyoshi Wada, and Yuko Ito. Effective application of pao : Seal type robots for disabled people in according to ideas of occupational therapists. In Klaus Miesenberger, Joachim Klaus, Wolfgang Zagler, and Arthur Karshmer, editors, *Computers Helping People with Special Needs*, volume 5105 of *Lecture Notes in Computer Science*, pages 1321–1324. 2008.

[5] K. Dautenhahn. Robots as social actors: Aurora and the case of autism. In *Proc. Cognitive Technology Conference*, pages 359–374, 1999.

[6] Karla Conn Welch, Uttama Lahiri, Zachary Warren, and Nilanjan Sarkar. An approach to the design of socially acceptable robots for children with autism spectrum disorders. *International Journal of Social Robotics*, pages 391–403, 2010.

[7] Prevalence of autism spectrum disorders — autism and developmental disabilities monitoring network, United States, 2006, January 2011.

[8] C. Nehaniv. Synchrony and turn-taking as communicative mechanisms. In C. Nehaniv and K. Dautenhahn, editors, *Imitation and Social Learning in Robots, Humans and Animals*. Springer, 2007.

[9] J.C.J. Brok and E.I. Barakova. Engaging autistic children in imitation and turn-taking games with multiagent system of interactive lighting blocks. In *Proceedings of the 9th international conference on Entertainment Computing*, pages 115–126, 2010.

[10] Charlie : An adaptive robot design with hand and face tracking for use in autism therapy. *International Journal of Social Robotics*, pages 1–11.

[11] B. Robins, K. Dautenhahn, R. Boekhorst, and A. Billard. Robots as assistive technology - does appearance matter? In *Proc. IEEE International Workshop on Robot and Human Interactive Communication*, pages 277–282, 2004.

[12] Guidelines for using robots in education and play therapy sessions for children with disabilities, January 2010.

[13] G. Bradski. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, pages 120–126, NOV 2000.

[14] R. Lienhart and J. Maydt. An extended set of Haar-like features for rapid object detection. In *Proc. IEEE International Conference on Image Processing*, pages 900–903, 2002.

[15] R. E. Schapire. The boosting approach to machine learning: An overview. In D. D. Denison, M. H. Hansen, C. Holmes, B. Mallick, and B. Yu, editors, *Nonlinear Estimation and Classification*. Springer, 2003.

[16] G. R. Bradski. Computer vision face tracking for use in a perceptual user interface. *Intel Technology Journal*, 1998.

[17] P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 511–518, 2001.

[18] Viola and Jones. Robust real-time face detection. *International Journal of Computer Vision*, 57(2):137–154, 2004.

[19] A.L. C. Barczak and F. Dadgostar. Real-time hand tracking using a set of co-operative classifiers based on haar-like features. *Res. Lett. Inf. Math. Sci.*, 7:29–42, 2005.

[20] P. Buehler, M. Everingham, D. Huttenlocher, and A. Zisserman. Long term arm and hand tracking for continuous sign language tv broadcasts. In *British Machine Vision Conference*, 2008.

[21] Juan Wachs. Enhanced human computer interface through webcam image processing library - aGest.xml, March 2011.

APPENDIX B

PRESCREENING QUESTIONNAIRE

PRESCREENING QUESTIONNAIRE

[1]	**Date information collected**	______
[2]	**Name of parent/guardian**	______
[3]	**Parent contact information**	
	Address	______
	City/State	______
	County	______
	Phone (H)	______
	Phone (C)	______
	Other	______
[4]	**Father's occupation**	______
[5]	**Highest grade completed**	______
[6]	**Mother's occupation**	______
[7]	**Highest grade completed**	______
[8]	**Marital status of parent**	______
[9]	**Other languages spoken at home**	______
[10]	**Name of child**	______
[11]	**Date of birth**	______
[12]	**Age of child... years/months**	______
[13]	**Race/ethnic background of child**	______
[14]	**Gender**	______
[15]	**Age (years) and sex of siblings**	______
	Sibling 1	Age Sex
	Sibling 2	Age Sex
	Sibling 3	Age Sex
	Sibling 4	Age Sex
[16]	**Has child had a hearing evaluation?**	______
		Newborn Hearing Screening'

[17] **Name of audiologist** ______________________

[18] **Date of evaluation** ______________________

[19] **Reported results?** ______________________

[20] **Has s/he received diagnosis of autism?** Yes No

[21] **When was s/he diagnosed?** ______________________

[22] **Where was s/he diagnosed?** ______________________

[23] **Who made the diagnosis?** ______________________

[24] **If you do not have the results of the evaluation, are you willing to sign a release for the results?** ______________________

[25] **Are there any other relatives with a diagnosis of an autism spectrum disorder?** ______________________

If yes, please list

Maternal relatives ______________________

Paternal relatives ______________________

[26] **Does your child have any other diagnoses?** ______________________

Medical diagnoses ______________________

Behavioral diagnoses ______________________

Therapy

[27] Is s/he currently receiving therapy? ______________________

If so, what kind(s)? ______________________

Therapy #1

(a) When did s/he start? ______________________

(b) How long has s/he been in therapy? ______________________

(c) How often is s/he seen? ______________________

– which days per week ______________________

(d) How long is the therapy session? ______________________

(e) Who is providing the therapy? ______________________

Therapy #2

(a) When did s/he start? ______________________

(b) How long has s/he been in therapy? ______________________

(c) How often is s/he seen? ______________________

– which days per week ______________________

(d) How long is the therapy session? ______________________

(e) Who is providing the therapy? ______________________

Therapy #3

(a) When did s/he start? ______________________

(b) How long has s/he been in therapy? ______________________

(c) How often is s/he seen?

--- which days per week ________

(d) How long is the therapy session? ________

(e) Who is providing the therapy? ________

[28] Has s/he received a speech evaluation? ________

If yes, do you have the results from the assessmt? ________

If not, are you willing to sign a release for the results from the assessmt? ________

If not, can you obtain the results from the assessmt? ________

[29] How would you describe your child's language ability?

No words used? ________

Echoed language only? ________

Spontaneous single wds? ________

Spontaneous phrases? ________

Limited for age? ________

Fluent nonmeaningful language? ________

Fluent meaningful language? ________

Can you give me an

example of something s/he might say? ______

[30] How would you describe your child's nonverbal ability?

Does your child point? ______

Does your child lead you by the hand? ______

Does your child have other nonverbal comm such as:

(a) shakes head for no ______

(b) nods head for yes ______

(c) waves goodbye ______

(d) points ______

(e) hand out (give me that) ______

(f) shoulder shrug ______

(g) other ______

Any use of sign language? ______

(a) If yes, how many signs? ______

(b) What are they? ______

[31] Please describe some play activities your child enjoys doing (cars, books, etc) ______

APPENDIX C

RECRUITMENT: FUN WITH CHARLIE THE ROBOT FLYER

(INDIVIDUAL)

FUN WITH CHARLIE THE ROBOT!

USC SC Autonomous Robots & Research Lab
together with the USC Speech and Hearing Center and
The USC School of Medicine
are looking for

CHILDREN WITH AUTISM

DESCRIPTION

We are currently conducting a research study to evaluate the effectiveness of a robot named CHARLIE to improve the verbal and nonverbal communication skills in children with autism

DURATION

The duration of the study will be approximately 8 weeks.

ELIGIBILITY

We are looking for children:

- Between 3 and 6 years old
- Diagnosed with autism
- Speech/communication difficulties

Participation in this study is FREE of charge!

For more information, please call:

803-237-7598

Laura Boccanfuso
University of South Carolina
USC SC Autonomous Robots & Research Lab

APPENDIX D

AUTISM STUDY CONSENT FORM

Department of Computer Science and Engineering

Consent Form

Adaptive Robot Design with Hand and Face Tracking for Use in Autism Therapy

Laura Boccanfuso

<u>Information Statement</u>

Your child is invited to participate in a research study conducted by Laura Boccanfuso, Sarah Scarborough, Ruth Abramson and Harry Wright. Laura Boccanfuso is a Ph.D. candidate in the Computer Science and Engineering Department at the University of South Carolina (USC) and she is conducting this study as part of the requirements for her Ph.D. degree in Computer Science. Sarah Scarborough, M.A., is a senior clinical instructor in the USC, Department of Communication Sciences and Disorders, and a therapist at the USC, Speech and Hearing Research Center. Ruth Abramson, Ph.D. and Harry Wright, M.D., are faculty members of the University of South Carolina, School of Medicine, and possess expertise in autism diagnosis and treatment.

Illustration: CHARLIE

We are looking for approximately 25 children between 3 and 6 years of age and we are inviting your child to participate. Sessions will take place in the presence of CHARLIE (the robot used as a tool to increase interaction) and an interventionist. The purpose of the study is to improve communication skills in children diagnosed with Autism Spectrum Disorders (ASD). This form explains what you will be asked to do if you decide to participate in this study. feel free to ask any questions you like before you make a decision about participating.

Description of Study Procedures

The first visit will consist of a parent interview and will take place without the child present. You will be asked to sign a release of medical records, including any speech/language services s/he has received. We would like to know more about your child's skills at the present time and will give you two questionnaires: (1) the Social Communication Questionnaire (SCQ) and, (2) the Vineland Adaptive Behavior Scale. The first visit will likely take approximately one hour.

On your second visit, you and your child will be introduced to CHARLIE the robot and the interactive games it can play will be demonstrated. The robot can engage the child in a number of interactive games and songs. The robot uses a camera to track the child's face and detects the position of the child's hands. Use of the robot will not continue should your child demonstrate any distress or fear of the robot. If the child shows interest in the robot, game play will continue until (a) the therapist initiates a new task, (b) the session time ends or (c) the child shows significant signs of distress/fear. The second visit will take approximately 30 minutes.

This study will take place over the period of six weeks, with 30-minute sessions scheduled two times per week, and will involve both you and your child. Each session will be videotaped and catalogued for subsequent analysis. Access to the videotaped sessions will be limited exclusively to Laura Boccanfuso, as the primary researcher, Sarah Scarborough, Ruth Abramson and Harry Wright.

Risks of Participation

There are very few risks associated with participating in this research except a slight risk of breach of confidentiality, which remains despite steps that will be taken to protect your child's privacy. Each videotaped session will be catalogued using the date and a unique identifier assigned to the individual child. The real names of children participating in the study and their respective identifiers will be kept on a password-protected computer in a password-protected file.

Because autistic children are individuals with widely varying interests, abilities and personalities, it is possible that while some children may find the robot cute and interesting, others may be fearful of it or act aggressively towards it. If a child should act/show fear or distress, use of the robot will be discontinued and/or removed from the room.

Benefits of Participation

Taking part in this study may benefit your child directly. The robot will be used as a tool during play to engage the autistic child and is intended to promote verbal and nonverbal human-to-human communication. Results from this research may also help us understand how to better design robots that can be used by therapists, teachers and parents to help promote communication skills in autistic children.

Costs
There will be no costs to your family for participating in this study.

Circumstances for Dismissal from the Study
List the circumstances, if any, under which the subject's participation may be terminated without
his/her consent.

Your child may be dismissed from the study without your consent for various reasons, including the following:

- Your child is continually disinterested in playing with the robot
- Your child experiences increased or significant distress or anxiety because of the robot
- Your child is physically aggressive toward the robot or another person present because of the robot
- The interventionist deems that the robot is no longer beneficial to your child
- If the study sponsor decides to stop or cancel the study.
- If the investigators otherwise believe that it is not in your child's best interest to continue in the study.

Confidentiality of Records
Participation will be confidential. A unique identifier will be assigned to each participant at the beginning of the project. This number will be used on project records rather than your name, and no one other than the researcher will be able to link your information with your name. Study records/data will be stored in a password protected computer in a password protected files at the University of South Carolina. The results of the study may be published or presented at professional meetings, but your identity will not be revealed.

In rare cases, a research study may be evaluated by an oversight agency, such as the USC Institutional Review Board or the U.S. Office for Human Research Protections. If this occurs, records that identify you and the consent form signed by you may be inspected so that they may evaluate whether the study is properly conducted and the rights of participants were adequately protected.

Excerpts from the video may be used for demonstrating the viability of the robot as a therapeutic tool at research conferences and still frames selected from specific sessions may be used in a publication of this research in conference proceedings or journal article(s), but the name(s) of those children appearing in the photograph(s) will be kept strictly confidential. If you wish to include your child in the study, but exclude him/her from appearing in any video excerpt or photograph to be made publicly available please initial here _______.

Contact Persons
For more information concerning this research, or if you believe you may have suffered a

research related injury, you should contact Laura Boccanfuso at 803.237.7598 or boccanfu@email.sc.edu, or Sarah Scarborough at 803.777.2622 or scarbosc@mailbox.sc.edu.

If you have any questions about your rights as a research subject contact, Lisa Marie Johnson, IRB Manager, Office of Research Compliance, University of South Carolina, 901 Sumter Street, Byrnes 515, Columbia, SC 29208, Phone: (803) 777-7095 or LisaJ@mailbox.sc.edu. The Office of Research Compliance is an administrative office that supports the USC Institutional Review Board. The Institutional Review Board (IRB) consists of representatives from a variety of scientific disciplines, non-scientists, and community members for the primary purpose of protecting the rights and welfare of human subjects enrolled in research studies.

Voluntary Participation

Participation in this study is voluntary. You are free to decline participation or to withdraw at any time, for whatever reason, without negative consequences. In the event that you do withdraw from this study, the information you have already provided will be kept in a confidential manner.

Email Communication

I understand that I may request to be contacted via email for the purpose of scheduling appointments with your child. Please note that most standard email does not provide a secure means of communication. There is some risk that any protected health information contained in email may be disclosed to, or intercepted by, unauthorized third parties. Use of more secure communications, such as phone or fax is always an alternative that is available to you. We will not give out your email address to third parties. If you consent for us to communicate with you via email, please provide your email address below.

____

Email address

Signatures /Dates

I have read (or have had read to me) the contents of this consent form and have been encouraged to ask questions. I have received answers to my questions. I give my consent for my child to participate in this study, although I have been told that I may withdraw my child at any time without negative consequences. I have received (or will receive) a copy of this form for my records and future reference.

Parent/Legal guardian signature Date

As a witness, I attest that the consent form was read by (or to) the subject, the research purpose, procedures, risks, and benefits were explained to the subject, questions were solicited and if the subject had any questions, they were answered to the subject's satisfaction. In my judgment, the subject voluntarily agreed to participate in the study.

Researcher Date

Appendix E

Study Session Methodology

Overview of experiment procedure

I. Pre-test Procedure

If your child meets the criteria to participate in the study, there are two initial meetings that will be scheduled prior to the beginning of the study:

(a) The first meeting will take place at MedPark 15. You will be asked to complete a series of questionnaires without your child present. At that time, you will complete three measures that will help to establish the baseline for your child:

- The Vineland Adaptive Behavior Scales (assesses your child's behavior and estimates your child's cognitive age based on your report of their performance)
- The Social Communication Questionnaire (SCQ)

** Please plan to spend about 1 hour for this meeting.

(b) During the second meeting (which will take place at the USC Speech and Hearing Center), you and your child will be introduced to CHARLIE. We will evaluate your child using the following two tests:

- Unstructured Imitation Assessment (Version A)
- Motor Imitation Scale: 16-item assessment of motor imitation, especially designed for children with autism (Version A)
- Expressive Vocabulary Test (Second edition) (Form A)
- Completion of the Mean Length of Spontaneous Utterance Determination

** Please plan to spend about an hour and a half for this meeting.

Illustration : CHARLIE

II. Test procedure and SLP Goals

Following the initial evaluations, we will schedule 30-minute sessions, twice a week (semi-weekly) for a total of 6 weeks.

(a) Each session will involve the speech therapist, a graduate student and the robot, CHARLIE.
(b) Introducing the robot to each child and will follow these general guidelines and will incorporate the following therapy goals:

- **Phase I** : The robot will be situated in the room where therapy will take place. For the first few sessions (and at the therapist's discretion) the robot will be placed in stationary mode – to allow the child the opportunity to physically explore the robot and its components before introducing movement.

  SLP GOALS:
  (1) On arriving and leaving each session, the subject will briefly make eye contact with clinician, CHARLIE and Laura as part of his/her greeting. *We will be ready to approach Phase II after the child has been observed to:*
  (2) Approach CHARLIE
  (3) Touch CHARLIE
  (4) Move CHARLIE's arms

- **Phase II**: The child, with the therapist's guidance (if required) will be given the opportunity to control the robot's arms and head with a remote control.

  (1) On arriving and leaving each session, the subject will briefly make eye contact with clinician, CHARLIE and Laura as part of his/her greeting.
  (2) During follow directions task, subject will point/operate remote/follow direction in order to lead CHARLIE through an activity at least once during the session. *We will be ready to approach Phase III after the child has been observed to:*
  (3) Use the remote control to make CHARLIE move

- **Phase III**: The child, with the therapist's guidance (if required) will be given the opportunity to play music and practice movement along with the robot. The robot can play “If you're happy and you know it” and “The wheels on the bus” with hand and head movements.

  (1) On arriving and leaving each session, the subject will briefly make eye

contact with clinician, CHARLIE and Laura as part of his/her greeting.

(2) During follow directions task, subject will point/operate remote/follow direction in order to lead CHARLIE through an activity at least once during the session.

(3) During song activity, subject will participate in fingerplays/gestures with CHARLIE for 80% of opportunities.

(4) Once CHARLIE has imitated subject's movement, subject will continue to move/interact with CHARLIE through # turns (# to be determined from performance on baseline/previous session). *We will be ready to approach Phase IV after the child has been observed to:*

(5) Respond to song with appropriate fingerplay/gesture

(6) Move in response to CHARLIE's prompt/action

- **Phase IV**: The child, with the therapist's guidance (if required) will be given the opportunity to play imitation games with the robot. One-on-one games include just the child and the robot. The "Pass the pose" game includes the therapist, the child and the robot.

(1) On arriving and leaving each session, the subject will briefly make eye contact with clinician, CHARLIE and Laura as part of his/her greeting.

(2) During follow directions task, subject will point/operate remote/follow direction in order to lead CHARLIE through an activity at least once during the session.

(3) During song activity, subject will participate in fingerplays/gestures with CHARLIE for 80% of opportunities.

(4) Once CHARLIE has imitated subject's movement, subject will continue to move/interact with CHARLIE through # turns (# to be determined from performance on baseline/previous session).

(5) When offered a choice of activities, subject will clearly make his/her performance known to others in session for 80% of trials.

(6) Throughout the therapy session, subject will cooperate with a turn-taking task with CHARLIE, parent and/or clinician through (2) turns (change this # as client progresses) each. *We will be ready to approach Phase V after the child has been observed to:*

(7) Imitate the robot movements on 80% of trials

(8) Imitate movements with another person in the intervention room

- **Phase V**: The child will be given the opportunity to select from various modes of play with the robot.

(1) On arriving and leaving each session, the subject will briefly make eye contact with clinician, CHARLIE and Laura as part of his/her greeting.

(2) During follow directions task, subject will point/operate remote/follow

direction in order to lead CHARLIE through an activity at least once during the session.

(3) During song activity, subject will participate in fingerplays/gestures with CHARLIE for 80% of opportunities.
(4) Once CHARLIE has imitated subject's movement, subject will continue to move/interact with CHARLIE through # turns (# to be determined from performance on baseline/previous session).
(5) When offered a choice of activities, subject will clearly make his/her performance known to others in session for 80% of trials.
(6) Throughout the therapy session, subject will cooperate with a turn-taking task with CHARLIE, parent and/or clinician through (2) turns (change this # as client progresses) each.
(7) During interactive games and songs with CHARLIE, subject will participate in a structured reciprocal play routine for (2) minutes on (3) occasions (change #'s as client progresses).

- **Phase VI**: If the child moves smoothly through the previous 5 sessions, then on the 6th session s/he will be given the opportunity to select from various modes of play with the robot and any member of the research staff.

  (1) On arriving and leaving each session, the subject will briefly make eye contact with clinician, CHARLIE and Laura as part of his/her greeting.
  (2) During follow directions task, subject will point/operate remote/follow direction in order to lead CHARLIE through an activity at least once during the session.
  (3) During song activity, subject will participate in fingerplays/gestures with CHARLIE for 80% of opportunities.
  (4) Once CHARLIE has imitated subject's movement, subject will continue to move/interact with CHARLIE through # turns (# to be determined from performance on baseline/previous session).
  (5) When offered a choice of activities, subject will clearly make his/her performance known to others in session for 80% of trials.
  (6) Throughout the therapy session, subject will cooperate with a turn-taking task with CHARLIE, parent and/or clinician through (2) turns (change this # as client progresses) each.
  (7) When presented with communication opportunities by clinician, subject will use gestures, vocalizations, or verbalizations for a variety of communicative intents on 80% of opportunities presented.
  (8) When subject desires to initiate, change or discontinue activities

within the last session, subject will make eye contact with appropriate clinician, Laura or parent before communicating the message.

(c) Upon completion of the six weeks of treatment, follow-up assessments (to measure progress achieved during the 6-week course of therapy) will include:

- Unstructured Imitation Assessment (Version B)
- Motor Imitation Scale (Version B)
- Expressive Vocabulary Test (Second edition) – (Form B)
- Completion of the Mean Length Utterance Determination

Please plan to spend approximately an hour and a half for this meeting.