# Multimodal Response Analysis for Behavioral Intervention of Children with ASD

By

Zhaobo K. Zheng

Dissertation

Submitted to the Faculty of the

Graduate School of Vanderbilt University

In partial fulfillment of the requirements

For the degree of

DOCTOR OF PHILOSOPHY

in

Mechanical Engineering

Decmber 18, 2021

Nashville, Tennessee

Approved:

Nilanjan Sarkar, Ph.D.

Amy Weitlauf, Ph.D.

Eric Barth, Ph.D.

Thomas Withrow, Ph.D.

Maithilee Kunda, Ph.D.

ACKNOWLEDGEMENTS

First of all, I want to thank Prof. Nilanjan Sarkar, who has been my mentor here at Vanderbilt University. Over the years, he spent uncountable time teaching me, including technical aspects, collaboration skills, academic writing and presentation, as well as the details to be an independent researcher in general. He put his trust in me for me to carry out projects confidently but whenever I came across obstacles, he stood in front of me to fix them. Every skill I learnt, every good habit I started, and every paper I published, I owe it to him. I want to give my special thanks to Dr. Amy Weitlauf, Dr. Zachary Warren and Mr. John Staubitz, without whom this dissertation would not be possible. They have helped me a great deal on my research and they have supported me unwaveringly. I have worked very closely with them and it has been a magnificent pleasure. I want to thank my PhD committee members, Dr. Maithilee Kunda, Dr. Thomas Withrow and Dr. Eric Barth. They have not only provided many valuable inputs to my dissertation, but also have taught me courses or collaborated with me.

Then, I want to thank my colleagues from the Robotics Autonomous Systems Lab. Special thanks to Dr. Dayi Bian, who have taught me many technical skillsets hand-in-hand for 4 years. I am very grateful for the wonderful working experiences with Jing, Huan, Nandan and Nibraas. Furthermore, I want to thank Esube, Josh, Lian, Guangtao, Ashwaq, Akshith, Sidharth, Cadence, Deeksha, Ritam, Miroslava, Mahrukh and Abigale. Our lab has a very nice atmosphere where we can discuss academically and chat casually all the time, and I will surely miss that. Ms. Amy Swanson has helped me on human subject recruitments about a hundred times and I much appreciate it. I want to thank our devoted behavioral analysts, Lauren Shibley, Michelle Hopton, David Reichley, and William Martin. I also would like to thank all my lovely friends at Vanderbilt University for our friendship. I was so lucky to meet you and I really enjoyed it.

I want to give my special thanks to all the families and children that participated in our studies in the Great Nashville area.

Finally, I would like to express my eternal gratitude to my family. My parents, Guiying Yang and Wanyong Zheng, have supported, cared and encouraged me throughout my life. I met my girlfriend, Dr. Simeng Zhao, here at Vanderbilt and she has offered me insights and happiness that only she could provide. Research presented in this dissertation did not always go smoothly and they were always there to cheer me up. Nothing has humbled me so much as their love and I will devote the rest of my life to repay them.

TABLE OF CONTENTS

## List of Tables

## List of Figures

# Chapter 1: Introduction

## 1.1 Research Goals

Autism Spectrum Disorders (ASD) is a common and complex neurodevelopmental disorder [1]. It affects one in every 54 children in the US [2]. Although ASD affects people across the lifespan, literature has shown that behavioral and educational intervention can make a positive impact on individuals and families. [3][4]. However, conventional behavioral assessments and interventions require highly experienced therapists and sometimes suffer from lack of motivations from the children with ASD [5][6], making it costly, inaccessible and less engaging. It is estimated that the average lifetime cost for individuals with ASD is over $1.4 million to get necessary supports [7]. In recent years, data-driven assessment tools and technology-assisted intervention systems for children with ASD are being considered to be complementary alternative intervention modality for their appeal to children, quantitative measurement and low-cost [8]. In particular, generating multimodal data acquisition systems and analytical frameworks for predictive models can have various clinical and behavioral applications for children with ASD. Thus the goal of this dissertation is to investigate how analytical models can be designed and applied to enhance the overall quality of life of children with ASD. We have chosen 3 different task domains to demonstrate the power of analytical models: 1) Early detection of ASD; 2) Problem behavior precursor detection; and 3) Daily living skills training. These three tasks are not closely related but they span important aspects of assessment and intervention for children with ASD. The background and importance of each task will be addressed below.

The diagnosis of ASD is a "milestone" for many children with ASD and their families so that they can benefit more fully from early supports and interventions. The current reliable diagnosis of ASD require vocal and cognitive skills, which need time to develop. As a result, children can only receive their reliable diagnosis of ASD after the age of two [9]. Furthermore, due to the limited resources and financial conditions, many children do not receive a final diagnosis until much older and some people are not diagnosed until they are adolescents or even adults [10]. Consequently, most children do not receive early interventions in the first year of life, which is the optimal time period for enhancing developmental outcomes due to neural plasticity [11]. Despite the fact that cognitive and vocal skills are present later in infancy, human brains respond much earlier to more basic stimulus including voices, smells and images [12]. Existing studies have provided evidence that infants with ASD have different behavioral responses to audio and visual stimuli [13]. Thus, quantitative data acquisition on behavioral responses in infancy has shown the potential to identify the different patterns between the high-risk and low-risk of ASD groups at much earlier age. However, existing research has mainly focused on audio and visual stimuli because they are easier to deliver. Tactile perception, or the sense of touch, develops in the first few months [14] in infancy but has remained largely unexplored in current research. Literature has shown that children with ASD have significantly different responses to tactile stimuli as compared to neurotypical children [15]. Therefore, the behavioral response patterns of infants at high-risk of ASD to multisensory stimuli, including tactile stimuli, has the potential to better predict risk of ASD but are yet to be explored. Thus, one goal of this research is to design a multimodal data acquisition system with a novel tactile stimulator to create a classification model for early detection of ASD.

Two thirds of the overall population with ASD display frequent problem behaviors [16]. Common problem behaviors that co-occur with ASD include self-injury, aggression and elopement [17]. These behaviors severely impede involvement of children in community and educational activities [18] and can put children and their caregivers at risk of potential physical harm [19]. Persistent problem behaviors can prevent children from learning new skills [20], excluding them from school services and community opportunities and aggravating financial burden on caregivers [21]. Parents and experienced therapists can observe precursors of such problem behaviors and intervene in time to mitigate behavior topographies or even prevent problem behavior episodes. However, these precursors can be very insidious [22] so that some of the problem behaviors seem to be “out of the blue”, making the observation difficult. It is also not practical for parents to always keep an eye on their children due to time constraint. Affective computing on aggression and stress among children with ASD has shown potential to have an automatic and precise alternative alert mechanism for parents and caregivers. The current literature has explored classifying self-injury behaviors and aggression with physiological and motion data and the precision is encouraging [23][24]. However, most current studies are limited to single channel data analysis. Multimodal data streams contain more observations of problem behaviors that can be useful. At the same time, most studies focus on the problem behavior episode themselves, leaving the prediction of precursors unexplored. More importantly, all the current studies focus on offline analysis when a real-time alert system can be more useful to help parents and caregivers. Therefore, one goal of my research is to build a machine learning predictive model of precursors of problem behaviors for children with ASD with a customized multimodal data acquisition system.

Many children with ASD experience difficulties in learning good daily living activity skills potentially due to their fine motor dexterity and learning impairments [25]. In particular, many children with ASD have problems with maintaining good oral health [26][27] and find difficulties in learning how to brush their teeth, at times needing assistance when brushing [28][29]. Lack of toothbrushing skills can lead to increased dental visits, which is already extremely traumatizing for children and even adults with ASD because of their non-compliance and sensitivity [30]. Besides health implications, poor daily living activity skills may also impede one’s social interactions across their life-span [31]. Conventional daily living skill training require repetitive demonstration by parents or caregivers and children usually have limited motivations during these training [32]. Considering the fact that many children with ASD show a natural affinity for computer technologies, skill training through computer-based interventions may lead to a higher level of engagement and fewer disruptive behaviors [33]. Virtual-reality (VR) and augmented-reality (AR) are technologies that allow users to actively participate in the interactive simulated situations and in recent years, they have been applied to provide an attractive, replicable, quantitatively measurable, and controlled intervention environment with real-time feedback [34]. A few VR-based systems have been developed to investigate and teach important skills such as driving skills [35], social skills [36] and motor skills [37]. The results suggest that children were able to appropriately understand, use and react to virtual environments with the possibility of transferring these skills to real life. Despite the success of VR-based skill training intervention systems, there is not yet an intervention system that teaches daily living skills, in particular, toothbrushing skills. AR is a more immersive environment that projects real-world surroundings with the virtual objects, allowing better chance of relating teaching sessions to real-life scenarios. Therefore, one of my research goals is to design an interactive AR coaching system for toothbrushing skills in children with ASD.

The research presented here focuses on the design, development and application of multimodal data acquisition systems and analytical response models for children with ASD. This research aims to: 1) develop a multimodal data capture system and tactile stimulator to investigate multisensory response on children with high and low risks of ASD; 2) build a machine learning model to classify imminent precursors of problem behaviors with multimodal data from children with ASD and a real-time alert system to notify their parents and caregivers; and 3) develop an interactive AR-based intervention system to coach toothbrushing skills to children with ASD with real-time feedback, guidance and multimodal measurements. This research generates multimodal classification models to assist conventional ASD assessments and interventions, reduce the burden of parents as well as therapists, and expand accessibility to effective ASD interventions.

The rest of this chapter is organized as follows. Section 1.2 reviews the literatures in the related fields including analytical models for early detection of ASD, affective computing on children with ASD, and VR-based intervention systems for children with ASD. Section 1.3 addresses the opportunities for research contributions. Section 1.4 summarizes my completed research and presents my proposed future work.

## 1.2 Data-Driven Analytical Systems for Children with ASD

Since the conventional assessments and interventions for ASD are limited by resource constraint [38] and emerging technologies and sensors have become more versatile, precise, and autonomous, research about quantitative analytical systems for children with ASD has gained momentum in recent years, which has several advantages [39]. First, behavioral response data captured by sensors are more objective, automatic and precise. As behavioral responses can be characterized through multiple channels such as eye gaze, facial expressions, body movements and voices, they are difficult for therapists to monitor at the same time. Furthermore, some of these responses are undetectable by human such as physiological and brain electrical activity signals but can be measured by sensors. Advanced sensors accelerate processes and collect process data in real time with the possibility of detection of more subtle responses, making observations more accurate, reliable and continuous. Thus, they allow precise and synchronized response capture on a broad range of observations for behavioral analytical models. Second, intelligent systems can deliver engaging visual, auditory and tactile stimuli in a controlled and precise way. Children with ASD often have difficulty filtering primary information from complex and fast-changing scenarios. Intelligent systems break down these processes so they can better understand. Also, their responses to different stimulus can be compared. Intelligent systems can also provide real-time encouraging feedback to the children, which will likely be motivating. In addition, these analytical models are quantitative, making them significantly scalable and easier to transfer between caregivers and therapists. With these benefits, data-driven analytical systems have great potential to assist behavioral assessments and interventions for children with ASD.

This work aims to design multimodal data acquisition systems and apply them as assessment and intervention tools for children with ASD. Therefore, the literature reviews in the following subsections focus on the existing technologies and studies related to these topics. First, we begin with the quantitative behavioral response studies for early detection of ASD. Next, the affective computing on aggression related emotions on children with ASD is introduced. Finally, we introduce the existing intervention systems using VR and AR technologies.

### 1.2.1 Quantitative Analytical Models of Early Detection of ASD

In the past decade, quantitative data collection has been assisting early detection of ASD. Quantitative data make it easier to investigate the brain responses to more basic stimuli. Those responses develop long before the observable behavioral and communication symptoms of ASD become apparent [40]. There are strong evidences of more basic sensory responses to lower level stimulus developing earlier in infancy. A study [12] assessed sensory processing differences between 24-month infants at high-risk of ASD using a parent-reported measure. Analysis showed that high-risk infants diagnosed later with ASD have more difficulty with auditory processing and lower eye fixation compared to controls. The results of the study supported that behavioral responses to sensory input represent early risk markers of ASD, particularly in high-risk infants. Several studies have investigated visual attention to faces and other social stimuli on children at high-risk of ASD. In a longitudinal study, researchers delivered audio-visual stimuli to the infants and mark the eye fixations and the attention to eyes was shown to be present but in decline in 2-6-month-old infants who were later diagnosed with ASD [13]. The results showed that infants later diagnosed with ASD exhibited mean decline in eye fixation while the later typically developed (TD) infants did not show this pattern. These observations mark the earliest known indicators of social impairment in infancy of children with ASD. Another longitudinal study described the development of components of visual attention, including engaging, sustaining, and disengaging attention in infants at high-risk of ASD [41]. Infants with and without ASD were filmed as they engaged in play with small, easily graspable toys. They measured the durations of time spent looking at toy targets before moving the hand and after they grasped the toys. The results showed that infant who later got diagnosed with ASD were distinguished by prolonged latency to disengage by 12 months of age. To examine the gaze patterns of 6-month-old infants at high risk and low risk for developing ASD, a study showed videos of faces including still, moving and expressing positive affect, and speaking to infants [42]. Infants who later developed ASD spent less time looking at the presented scenes in general than other infants. When these infants looked at faces, their looking toward the socially informative features of faces such as eyes and mouths decreased compared to the TD group only when the presented face was speaking.

There are different methods to observe the behavioral responses of infants and common measurement methods for behavioral responses include electroencephalogram (EEG), eye tracking and video modeling. Local neural network connectivity undergoes rapid change during early development and it may reflect in the EEG signals. Researchers have utilized modified multiscale entropy (mMSE) and time asymmetry index of EEG to classify TD and high-risk groups with a support vector machine [43]. The classification was computed separately within each age group from 6 to 24 months. Multiscale entropy appeared to go through a different developmental trajectory in infants at high risk of autism and differences were greatest at ages 9 to 12 months. Infants were classified with over 80% accuracy into control and high risk groups at the age of 9 months. It is also worth mentioning that classification accuracy for boys were close to 100% at 9 months of age and remained high while for girls, the classification accuracy was higher at 6 months of age and declines afterwards. Another study also investigated the effectiveness of EEG signal detecting ASD [44]. The work presented a new algorithm to derive unprocessed EEG trace to get the multiscale and ranked organizing maps (MS-ROM). With that feature, the study collected EEG data on fifteen subjects with ASD and ten TD subjects and built a machine learning model based on their data. The overall predictive accuracy of the machine learning model in sorting out autistic cases was 84%-92.8%. The results did not vary on the age group, which suggested the machine learning model did not utilize age-related EEG patterns. The feature generation of EEG data was also investigated. A study investigated fractal dimension of EEG of children with ASD using complexity and chaos theory to discover a nonlinear feature space [45]. A wavelet-chaos-neural network methodology was presented for automated EEG-based diagnosis of ASD and the model was tested on a database of nine children with ASD and eight TD children. The classifier reached an accuracy of 90% based on the most significant features with $p < 0.001$.

In order to analyze the eye gaze pattern of infants with high risk of ASD, a study used cameras to record the face-to-face interactions of 32 infants with their parents and manually coded the videos to mark where the infant was looking in each frame [46]. Then they created variable-order Markov models to compare TD infants with infants with ASD. The models correctly classified infants who did develop ASD with an accuracy of 93.75%. Another study investigated and modeled the eye gaze behaviors of TD children and children with ASD while socially interacting with a humanoid robot [47]. A variable-order Markov model was built to represent and analyze the dynamics of eye gaze of these two groups. The experimental results demonstrate that the model can describe and discriminate between the gaze patterns of ASD and TD groups in speaking and listening contexts. The TD group has different gaze patterns in speaking and listening context and their gaze responses would vary significantly depending on their role in conversational context. However, the gaze responses of the ASD group did not vary much with their interaction roles. Another study combined eye gaze with demographic feature descriptors of age and gender to classify autism [48]. They tested the constructed feature descriptors using the eye gaze information from the National Database for Autism Research with three different classifiers including random regression forests, decision tree and partial decision tree. The presented method for classifying autism resulted in a top classification accuracy of 96.2%.

Video modeling is also an important tool for diagnosis of ASD. A study assessed the feasibility of applying a gold-standard diagnostic instrument to brief and unstructured home videos and tested whether video analysis can enable more rapid detection of the core features of autism [49]. Four non-clinical raters independently scored all videos using the autism diagnostic observation schedule-generic (ADOS). The classification accuracy was 96.8% with 94.1% sensitivity and 100% specificity. The results indicate that it is possible to achieve high classification accuracy, sensitivity, and specificity as well as clinically acceptable inter-rater reliability with nonclinical personnel. Machine learning has also been combined with video modeling for early detection of ASD. A study explored machine learning analysis on home videos to speed up the diagnosis [50]. They analyzed item-level records from 2 standard diagnostic instruments to construct machine learning classifiers optimized for sparsity, interpretability, and accuracy. They tested feature extraction by blinded non-expert raters from 3-minute home videos of children with and without ASD to arrive at a rapid and accurate machine learning based autism classification. The classifier reached an accuracy of more than 90%.

### 1.2.2 Assessing Aggression and Affective Computing on Children with ASD

The current state-of-the-art practice for addressing problem behaviors or aggression is Practical Function Analysis (PFA), which is a widely-used practice that focuses on the precursors of problem behaviors [51]. The PFA has demonstrated clinical utility when identifying and measuring precursor behaviors, which are observable behaviors – such as changes in body movement, affect, or vocalizations – that reliably precede the onset of problem behaviors. In fact, it has been shown that precursors are functionally directly related to dangerous problem behaviors [52]. Because of this, assessors can use precursors as safe proxies for problem behaviors within the assessment context to reduce the potential for unsafe behavioral escalation. However, despite the advantages of PFA, it is impractical and difficult for caregivers to monitor children with ASD at all times in order to watch for sometimes very subtle precursors of impending problem behaviors [22]. The current PFA approach to annotate behavioral observations of children is typically using paper and pen while timestamping events via a stopwatch. This method requires significant attention of therapists and offers standardized methodologies, but limited precision, with regard to recording the onset time of precursors. There have been some attempts recently to automate this process. A computerized behavioral data program "BDataPro" allows real-time data collection of multiple frequency and duration-based behaviors [53]. Catalyst, another software for behavioral assessment, allows collection and management of a wide variety of data for behavioral intervention [54]. An annotation tool for problem behaviors for people with ASD was also developed to log data more conveniently [55]. However, there are several limitations of their use in the PFA and many practitioners continue to collect data via paper-and-pen modality [56].

The field of affective computing is an emerging field that aims to enable intelligent system to recognize, infer and interpret human emotions and mental states. To address the current disadvantages of assessing aggression, affective computing has been successfully applied on children with ASD to infer emotional and behavioral states with various sensory data. Peripheral physiological responses such as heart rate (HR) and galvanic skin response (GSR) have been used to predict imminent aggression [23]. The study used an E4 wristband, a physiological biosensor to wirelessly record responses on 20 adolescents with ASD over 69 independent naturalistic observations. Using ridge-regularized logistic regression, results demonstrate that, on average, their models are able to predict the onset of aggression 1 minute before it occurs with an area under curve (AUC) of 0.84 for person-dependent models. Skin conductance and respiration were used to build an ensemble of classifiers to differentiate the arousal level and valence in children with ASD [57]. They recorded the HR, GSR, respiration and body temperature signals of 15 children with ASD when they were viewing different pictures. Then machine learning models were built to classify high or low arousal and positive or negative valence with a mean accuracy of 84.5%. The results suggest the feasibility of objectively discerning affective states in children with ASD using physiological signals. With regard to behavior recognition from body motion, accelerometer data was used to recognize stereotypical hand flapping and body rocking behaviors, which may occur to some children with ASD [58]. The study involved 6 children and each child wore two wristbands on the wrists and another wristband around the chest area. Then 3 non-expert observers used a phone application to annotate stereotypical motor movements of participants for classifier training. The machine learning models achieved an overall recognition accuracy of 88.6%. Stereotypical motor movements in ASD were detected using deep learning and resulted in a significant increase in classification performance relative to traditional classification methods [59]. The results with convolutional neural networks provided the preliminary evidence that feature learning and transfer learning embedded in deep architectures can provide accurate stereotypical motor movement detectors in longitudinal scenarios. Given the success of existing data-driven frameworks for affective computing, a recent study has explored whether they can predict aggressive behaviors and self-injurious behaviors (SIB) [24]. Participants wore 6 tri-axial accelerometers on wrists, ankles and pockets on their pants. Movement data along with annotated behaviors were collected to build a machine learning model to predict episodes of SIB. The model built achieved an average classification of SIB episodes at an accuracy of 94.6%.

### 1.2.3 VR and AR Based Intelligent Intervention Systems for Children with ASD

In recent years, computer assisted ASD interventions have shown great potential due to their low-cost, appeal to children with ASD, and relatively broader accessibility [60][61]. Many children with ASD have a natural affinity for computer technologies [32] and VR technologies offer them an immersive simulated environment, where they are more engaged with the help of controlled stimuli and real-time feedback [34]. Several VR-based systems have been developed to investigate and teach important life skills to children with ASD and results suggest that children were able to appropriately understand, use and react to virtual environments. For instance, a novel VR-based driving stimulator was developed to teach driving skills to teenagers with ASD [35]. The participants drove a virtual vehicle in a virtual city to complete driving tasks such as passing traffic lights, pulling over, and entering the highway. The simulator detects participant errors and eye gaze and provides appropriate instructions. A virtual haptic training system was designed to assess and improve fine motor skills on children with ASD [37]. The system allows participants to grip and move virtual objects in games and thus provides opportunities for them to improve finger and hand motor control. A haptic device was integrated with a gripper so that the participants needed to move and grip an end effector to fulfill the virtual tasks including picking toys from a crane machine, writing alphabets and collaborating to move objects through mazes. Additionally, a VR-based social cognition training system was developed where participants could practice social tasks including social introductions, conversation initiations, meeting friends, and other social interaction scenarios [36]. The system provides real-time guidance on social initiation to adolescents with ASD in a virtual hall environment so that they can learn how to appropriately interact with strangers. A collaborative virtual environment was built which was a computer-based, distributed, virtual space for multiple users to interact with one another and with virtual items [62]. The system measured the collaborative interactions and verbal communications of children with ASD when they played collaborative puzzle games with their TD peers in remote locations. The system had the ability to promote important collaborative behaviors including information sharing, sequential interactions and simultaneous interactions and to provide real-time feedback based on their game performances. The system increased the social interaction of children with ASD in a feasibility study with 14 pairs of children – 7 ASD and TD pairs and 7 TD and TD pairs.

The success of VR ushered in AR into intervention systems for children with ASD. AR is a computer-based technology that superimposes real-world actions and images on computer-generated displays of virtual characters, scenarios, and interactions, providing a composite view of the situation [63][64]. In recent years, AR-based interventions for children with ASD have been reported in literature. A mirror-like AR system allowed participants to see themselves and their constructed skeletons on the screen [65]. The system captured the movement of the participants without wearing invasive devices and animated their virtual avatars. The system taught children with ASD about their own bodies and allowed them to imitate body gestures. The system was tested with 5 children with ASD and 15 TD children and most children successfully interacted with the system. In another study, a virtual agent was developed to let young adults practice job interviews [66]. Children wore a Magic Leap AR goggle so that they can see a virtual interviewer in the real world and interact with it. The system also measured the performances of the children including eye contact, blinking rate and head orientations. Early tests with young adults with ASD showed that they were drawn to this novel technology and use of virtual humans, praised the immersion of characters, and appreciated the ability to practice valuable skills within a real-world setting. An AR smart glasses system was developed [67] to coach social communication where gamified AR applications provided children with ASD coaching for emotion recognition, face directed gaze and eye contact. The coaching intervention was found to be well tolerated, engaging, and fun at the same time. Although an emerging field, existing research to date has shown the potential of AR-based intervention systems to help children with ASD learn life skills by immersing them in real-life situations [68].

## 1.3 Opportunities for Research Contributions

### 1.3.1 Multimodal Sensory Response Analysis for Early Detection of ASD

The existing works on quantitative classification of high risk ASD in infants have shown great potential. However, most of them focused on audio-visual stimulus with only single channel behavioral response measurements. One sensory processing channel that could advance existing works is related to tactile perception, or the sense of touch. The sense of touch is widely known for the role it plays in discriminating and identifying external stimuli. For example, by investigating the neural basis of somatosensory remapping in human infancy, researchers found that the cortical networks underlying the ability to dynamically update the location of a perceived touch across limb movements become functional during the first year of life [69]. However, there is growing evidence that the sense of touch has another dimension, also known as “affective touch”, which conveys social information just like what someone sees and hears [14]. Affective touch, which can be described as a comforting, caress-like soft touch has been found to impact the social function regions of human brains [70]. Previous research has demonstrated that infants are sensitive to affective touch [71] and that compared to other forms of touch, stroking an infant can not only induce positive emotions but also modulate negative ones [72]. Atypical perception and processing of affective touch were also found in adults with ASD []. Furthermore, the authors hypothesized that the affective touch functionality, which is based on C tactile fiber activation [73], is impaired to some extent in individuals with ASD [69]. More importantly, another study showed that in the presence of affective touch stimulus, individuals with ASD exhibit reduced brain activity in social-emotional-related brain regions compared to TD individuals [15]. Consequently, affective touch represents an identified area of atypical sensory processing related to ASD that can also influence infant responses, making it an optimal target for early detection of neurodevelopmental risk. Analyzing the perception on affective touch on infants with high risk of ASD could augment the current stage of early detection of ASD.

However, there is not yet an appropriate tactile stimulator to deliver controlled affective touch to infants. The affective touch needs to be a gentle stroke with precisely controlled speed and pressure for safety and repeatability. It is also important that the tactile stimuli are delivered without human presence so that there is less interferences to infants. Therefore, an automatic tactile stimulator with good speed and pressure control is needed. The existing analytical models for early detection also relied mostly on single channels of data while the combined multimodal response data streams have the potential to better identify the data patterns between the low-risk and high-risk groups. Thus, there is a need to combine current response measurement methods by fusing the data from multiple channels for better precision and also to compare the importance of each channel.

### 1.3.2 Multimodal Predictive Model of Precursors of Problem Behaviors on Children with ASD

Current research has shown the feasibility of predicting aggression, self-injurious behaviors in children with ASD but it has several limitations. Most works in this field mainly focused on the prediction of problem behaviors themselves instead of the precursors of problem behaviors, which will likely provide some time to intervene and deescalate the situations. A better protocol to integrate the PFA can efficiently generate precursors and gather their behavioral responses. Furthermore, existing studies that apply unimodal systems to predict problem behaviors may omit useful important information that could increase the accuracy of the machine learning models and the chance to compare the relative importance of different data modalities in predicting the precursors. Therefore, integrating multimodal sensors and fusing different data streams can enhance the behavioral response measurements. There is also room for advancement for sensors to better consider the ergonomics of this population to be less invasive and more precise. Also, the current practice of behavioral observation limits the precision and convenience of such observation. The traditional paper-and-pen annotation method often has notable timing errors within the data assessed, which may impair the accuracy of the machine learning model. With an intelligent annotation tool, observers can have significantly less tasks and therefore record events more precisely and focus on observing the response. Thus, with the help of such a tool the timing deviation can be greatly improved and lead to higher prediction accuracies of the machine learning models.

### 1.3.3 AR-based Intervention System for Daily Activity Skills of Children with ASD

VR and AR based intelligent intervention systems have shown great potential in various skill training for children with ASD and there has not been literature to apply these technologies onto daily habits, which plays an important role in the qualities of life of children with ASD. Toothbrushing is one of the most important daily habits and children with ASD experience difficulties with it [74]. There have been a few technology-based systems that aim to help children with ASD to improve their toothbrushing skills. A toothbrushing training program on tablets showed the steps of toothbrushing to children with ASD, which resulted in some improvement. Another study used marker-based video triggering software to show children with ASD a clip of a peer brushing her teeth [75]. All participants learnt how to brush their teeth and maintained the skill in their daily life for 9 weeks after the study. A picture exchange communication system (PECS) based toothbrushing program was used on gingival health in children with ASD [76]. The parents of 37 children rated the program as useful in improving gingival health for children with ASD. A cartoon game called "Brush Up" was utilized on children with ASD and it resulted in significant reduction of visible plaque on post-intervention. Although promising, the existing studies are limited to showing photos and video clips in an open-loop manner and did not teach toothbrushing skills in an interactive way to provide real-time feedback and instructions based on the performance of the children. There is great potential to build an AR environment for daily living activity skill interventions because it demonstrates each part of the process step by step and provides closed-loop feedback and well controlled encouragement and guidance in real-time. These immersive experiences may better engage children with ASD and make the interventions more effective and beneficial, making up a key missing part of skill training intelligent system research.

## 1.4 Goals of the Dissertation

The overall goal for this dissertation is to design, develop and apply multimodal data capture systems, intelligent intervention systems and machine learning to study analytical models in children with ASD. These models aim to enhance the current research on three specific aspects, which are early detection of ASD, problem behavior prediction among children with ASD, and daily living skill coaching intelligent system for children with ASD. The goals of this dissertation will be accomplished through the following specific aims.

### 1.4.1 Specific Aim 1: Multimodal Stimulation and Data Capture System for Early Detection of ASD

To investigate the multisensory responses of infants at high risk of ASD, we designed a novel multisensory stimulation and data capture system (MADCAP). The system delivers multisensory stimuli to infants and collect various dimensions of behavioral responses. In addition to the audio-visual stimuli in the form of videos of mom-characters talking to infants, the tactile stimuli is very crucial in this aim. Therefore, we designed an automated mechanism that actuates a hair brush to replicate soft brushing, which is comparable to an affective touch manually produced by hand [80]. We designed a novel soft brushing mechanism to deliver affective touch on the forearms of children with ASD with precisely controlled speed and pressure using wearable design. For data capture, MADCAP measures EEG, peripheral physiological and eye gaze data.

To the best of our knowledge, this work is the first system to demonstrate a multimodal system incorporating affective touch that has the potential to meaningfully chart differences in coordinating visual, auditory, and tactile processing in infancy. The system has gone through a feasibility study of 17 toddlers aged between 2-5 years with ASD diagnosis and 12 age-matched TD controls. Ten out of 12 participants in TD group and 10 out of 17 participants in ASD group tolerated the entire experiment well and the rest nine could not tolerate the EEG cap or could not sit still. The overall compliance rate of toddlers was 72%. The machine learning models combining MADCAP features and the behavioral scores reached an average of 83% to classify the ASD group. The most informative channels are eye gaze fixation on eyes, the 12-30Hz section of EEG data and the eye gaze fixation on the mouth. This study can pave the way for a more precise and robust machine learning model for risk of ASD during infancy with multisensory response data.

### 1.4.2 Specific Aim 2: Predictive Model of Precursors of Problem Behaviors for Children with ASD

(a) Offline machine learning analysis

To locate behavioral patterns of problem behaviors and precursors, training data with reliable behavioral events are needed. Therefore, we created a novel multimodal data capture platform for precursors of problem behaviors, M2P3, for children with ASD. M2P3 combines an off-the-shelf wearable sensor, E4, a Kinect sensor, and a customized wearable intelligent non-invasive gesture sensor (WINGS). The presented multimodal platform is seamlessly integrated with a newly developed tablet-based software application, behavior data collection integrator (BDCI), to collect data and provide assistance to the assessment team. With these technological assistance, we performed a novel PFA embedded experimental framework to collect training data for a machine learning model that seeks to capture expert Board Certified Behavior Analyst (BCBA)'s ratings as the ground truth for behavioral events. With multimodal data captured and behavioral events captured from the experiments, we developed a predictive model to alert caregivers of problem behaviors for children with ASD (PreMAC) that predicts imminent precursors of problem behaviors.

We recruited 7 children aging from 7-12 with diagnosis of ASD to the feasibility study. All the children tolerated the system very well except for one kid who did not want to sit on the designated chair. Data was successfully collected and the machine learning models on individualized profiles achieved an average prediction accuracy of 98.51%.

(b) Real-time problem behavior prediction

Building on PreMAC, which has proven to have a high offline prediction accuracy, we sought to demonstrate the feasibility of predicting imminent precursors with multimodal real-time signals, which is, to our knowledge, the first in the field. We integrated the learnt models in the data collection modules to make real-time classifications. We made modifications to the PreMAC system to collect real-time signals, extract features, fuse the data, and then make online classifications. We planned to recruit 6 children with ASD and their families for an experimental study. There were two visits including the initial visit for data collection and building PreMAC; second visit to have real time alerts on behaviors of children and compare to the observations of BCBAs. In this dissertation, we present the feasibility of the real-time prediction system.

### 1.4.3 Specific Aim 3: Interactive AR coaching System for Toothbrushing Skills in Children with ASD

To build a more engaging and effective intervention system of daily living activity skills, we designed a novel interactive AR coaching system for toothbrushing skills, CheerBrush, for children with ASD. The system aims to help children with ASD learn independent toothbrushing skills. CheerBrush has both a coaching mode and an assessment mode and it is an AR-based system that teaches children with ASD how to brush their teeth in a fun, playful way. The system captures the movement of the children and let them interact with the virtual objects such as toothbrush with their hands. It decomposes the toothbrushing task in to small steps and guide children through them, with real-time feedback corresponding to their actions. The feedback includes audio, visual cues and virtual avatars. We also designed a mechatronic toothbrush to assess the toothbrushing skills pre- and post-test. The motion of toothbrushing is recorded by the integrated IMU and we compare their performances. We also monitor their physiological signals with an E4 wristband.

We recruited 8 children (4 with ASD, 4 TD) aging from 3-6 for the experiment. All of them played the game but finished different levels of difficulties. Children with ASD spent significantly longer time on different levels of difficulties and they also had less engaged time than the TD group. Both groups showed decreased HR and increased GSR, meaning they stayed calm during the experiments and did not experience increased stress. Both groups also showed better brushing motion in post-test, getting improvement on the skills. The study will continue to expand the experimental group size, adding versatilities and functions of the virtual tasks. The extended results will be demonstrated in the final defense.

## 1.5 Structure of this Dissertation

My research focuses on the multimodal response analysis in children with ASD, including 1) Detection of ASD developmental risk through multisensory stimuli, multimodal data collection, and machine learning; 2) Precursor of problem behavior prediction through wearable sensing and machine learning; 3) Real-time prediction of problem behaviors with multimodal online data; and 4) Intelligent AR based coaching system for toothbrushing skills in children with ASD. The following chapters present these projects in more details and are structured as below.

In Chapter 2, I describe my work on the design and implementation of the multisensory stimulation and data capture system, MADCAP. We designed the novel tactile stimulator that was able to deliver precisely controlled tactile stimuli that can trigger affective touch to toddlers. We integrated an eye tracker for gaze detection, an EEG systemfor brain wave monitoring, and an E4 wristband for physiological data collections. We recruited 27 toddlers in our pilot study to validate the feasibility and tolerability of our system and collected data for analysis and machine learning. The results showed that the multisensory delivery system had a high tolerability rate for toddlers with and without neurodevelopmental disorders. The machine learning results demonstrated that the affective touch features were informative for ASD developmental risk detection and it had the potential to assist screening in a much earlier age.

In Chapter 3, I present my work on the wearable multimodal framework for problem behavior prediction, PreMAC. We designed and developed a customized Wearable Intelligent Non-invasive Gesture Sensor (WINGS) that detects the roll, pitch and yaw angles of the upper body. We also integrated a Kinect and an E4 wristband for capturing facial expressions, head rotations and physiological data. To have more convenient and precise behavioral annotations, we built a tablet application, called the behavioral data collection integrator (BDCI). Through our pilot data collection, we trained both individualized and group machine learning models. The individualized models and the group model had average prediction accuracies over 98% and 82%, respectively.

In Chapter 4, I describe the AR-based coaching system for toothbrushing skills in children with ASD, CheerBrush. We designed and built an AR platform to project virtual objects, such as a toothbrush, in the real-world surroundings of the user on a screen and decomposed steps of toothbrushing from start to finish, with closed-loop multisensory guidance and feedback. We also designed a mechatronic toothbrush to monitor their brushing movements pre- and post-to the coaching sessions. We recruited eight children with and without ASD to participate in our pilot study. The results indicated that children with ASD spent more time on coaching sessions but less time engaged on eye gaze. There were some improvement in the brushing movements and it reached statistical significance for the brushing distance for the ASD group. Finally, from our physiological monitoring during the coaching sessions, we did not find induced stress during coaching.

In Chapter 5, I present my work of a smart application to establish brevity behavioral control in assessments. Inspired by the BDCI in chapter 3, we aimed to develop a smart application to augment visual analysis in functional analysis. We designed an app for behavioral analysts to input behavioral observations and then automatically compute whether we have behavioral control and its level on the children. The graphical user interface also provides graphs and numbers to update the user. We then validated the app with behavioral analysts on the multilevel criterion, inter-rater agreement and input recording reliability. The results suggested the app could run without any glitches, record all inputs quickly and precisely, and the multilevel criterion was reasonable.

In Chapter 6, I describe the real-time problem behavior prediction system. The results in Chapter 3 were promising and we wanted to demonstrate the feasibility of applying these models in real-time. We made modifications to PreMAC platform so that it could collect real-time signals, extract features, and make online classifications. We brought the participants in for two visits, where the first visit was for data collection and the second visit was for validation. We compared the machine learning predictions with the human expert observations as ground truth. The results suggested that the real-time machine learning system had short latencies for online classifications, and the prediction accuracies for individualized models were slightly lower than the offline machine learning, but still very promising for automatic, consistent, and real-time prediction of problem behaviors.

Finally, Chapter 7 summarizes the primary contributions of the dissertation research, in terms of technical contributions and contributions to the science of autism intervention, as well as future directions.

## References

[1] A. S. Weitlauf et al., Therapies for Children With Autism Spectrum Disorder: Behavioral Interventions Update, vol. 318, no. 6. 2017.

[2] M. J. Maenner et al., “Prevalence of autism spectrum disorder among children aged 8 Years-Autism and developmental disabilities monitoring network, 11 Sites, United States, 2016,” MMWR Surveill. Summ., vol. 69, no. 4, pp. 1–12, 2020, doi: 10.15585/MMWR.SS6904A1.

[3] P. Howlin, I. Magiati, and T. Charman, “Systematic Review of Early Intensive Behavioral Interventions for Children With Autism,” AJIDD, vol. 114, no. 1, pp. 23–41, 2009, doi: 10.1352/2009.114.

[4] R. Iovannone, G. Dunlap, H. Huber, and D. Kincaid, “Effective Educational Practices for Students With Autism Spectrum Disorders,” Focus Autism Other Dev. Disabl., vol. 18, no. 3, pp. 150–165, 2003, doi: 10.1177/10883576030180030301.

[5] Z. Warren, A. Vehorn, E. Dohrmann, C. Newsom, and J. L. Taylor, “Brief report: Service implementation and maternal distress surrounding evaluation recommendations for young children diagnosed with autism,” Autism, vol. 17, no. 6, pp. 693–700, 2013, doi: 10.1177/1362361312453881.

[6] C. Williams, B. Wright, G. Callaghan, and B. Coughlan, “Do children with autism learn to read more readily by computer assisted instruction or traditional book methods?” pp. 71–91, 2002.

[7] A. V. S. Buescher, Z. Cidav, M. Knapp, and D. S. Mandell, “Costs of autism spectrum disorders in the United Kingdom and the United States,” JAMA Pediatr., vol. 168, no. 8, pp. 721–728, 2014, doi: 10.1001/jamapediatrics.2014.210.

[8] G. L. C. Chia, A. Anderson, and L. A. McLean, “Use of Technology to Support Self-Management in Individuals with Autism: Systematic Review,” Rev. J. Autism Dev. Disord., vol. 5, no. 2, pp. 142–155, 2018, doi: 10.1007/s40489-018-0129-5.

[9] C. Lord, S. Risi, P. S. DiLavore, C. Shulman, A. Thurm, and A. Pickles, “Autism from 2 to 9 years of age,” Arch. Gen. Psychiatry, vol. 63, no. 6, pp. 694–701, 2006, doi: 10.1001/archpsyc.63.6.694.

[10] CDC, “Screening and diagnosis of autism spectrum disorders (ASD).,” Avaliable https//www. cdc. gov/ncbddd/autism/screening. html, 2018.

[11] J. Veenstra-VanderWeele and Z. Warren, “Intervention in the context of development: Pathways toward new treatments,” Neuropsychopharmacology, vol. 40, no. 1, pp. 225–237, 2015, doi: 10.1038/npp.2014.232.
[12] T. Germani et al., “Brief Report : Assessment of Early Sensory Processing in Infants at High-Risk of Autism Spectrum Disorder,” pp. 3264–3270, 2014, doi: 10.1007/s10803-014-2175-x.
[13] W. Jones and A. Klin, “Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism,” Nature, 2013, doi: 10.1038/nature12715.
[14] M. India, S. L. Line, and H. Olausson, “The skin as a social organ,” pp. 305–314, 2010, doi: 10.1007/s00221-009-2007-y.
[15] I. Gordon, A. C. Voos, R. H. Bennett, D. Z. Bolling, K. A. Pelphrey, and M. D. Kaiser, “Brain Mechanisms for Processing Affective Touch,” vol. 922, pp. 914–922, 2013, doi: 10.1002/hbm.21480.
[16] J. L. Matson and M. Nebel-Schwalm, “Assessing challenging behaviors in children with autism spectrum disorders: A review,” Res. Dev. Disabil., vol. 28, no. 6, pp. 567–579, 2007, doi: 10.1016/j.ridd.2006.08.001.
[17] S. Jeff, M. Arthur, and M. O’Reilly, Challenging behavior and developmental disability.pdf. 2003.
[18] S. M. Myers et al., “Management of children with autism spectrum disorders,” Pediatrics, vol. 120, no. 5, pp. 1162–1182, 2007, doi: 10.1542/peds.2007-2362.
[19] O. Chadwick, N. Piroth, J. Walker, S. Bernard, and E. Taylor, “Factors affecting the risk of behaviour problems in children with severe intellectual disability,” J. Intellect. Disabil. Res., vol. 44, no. 2, pp. 108–123, 2000, doi: 10.1046/j.1365-2788.2000.00255.x.
[20] E. Emerson, “Challenging behaviour: Analysis and intervention in people with severe intellectual disabilities,” 2001.
[21] S. L. Parish, K. C. Thomas, R. Rose, M. Kilany, and P. T. Shattuck, “State medicaid spending and financial burden of families raising children with autism,” Intellect. Dev. Disabil., vol. 50, no. 6, pp. 441–451, 2012, doi: 10.1352/1934-9556-50.06.441.
[22] J. Jessel, G. P. Hanley, and M. Ghaemmaghami, “On the Standardization of the Functional Analysis,” Behav. Anal. Pract., vol. 13, no. 1, pp. 205–216, 2020, doi: 10.1007/s40617-019-00366-1.
[23] M. S. Goodwin et al., “Predicting imminent aggression onset in minimally-verbal youth with autism spectrum disorder using preceding physiological signals,” ACM Int. Conf. Proceeding Ser., pp. 201–207, 2018, doi: 10.1145/3240925.3240980.
[24] K. D. Cantin-Garside, Z. Kong, S. W. White, L. Antezana, S. Kim, and M. A. Nussbaum, “Detecting and Classifying Self-injurious Behavior in Autism Spectrum Disorder Using Machine Learning Techniques,” J. Autism Dev. Disord., 2020, doi: 10.1007/s10803-020-04463-x.
[25] E. Z. Murshid, “Parents’ Dental Knowledge and Oral Hygiene Habits in Saudi,” vol. 14, no. 2, 2014.
[26] American Academy of Pediatric Dentistry, “Guideline on caries-risk assessment and management for infants, children, and adolescents,” no. 35, 5:E157. Pediatric dentistry, 2013.
[27] R. Y. Du, C. K. Y. Yiu, and N. M. King, “Oral Health Behaviours of Preschool Children with Autism Spectrum Disorders and Their Barriers to Dental Care,” J. Autism Dev. Disord., vol. 49, no. 2, pp. 453–459, 2019, doi: 10.1007/s10803-018-3708-5.
[28] A. J. Cuvo, A. Godard, R. Huckfeldt, and R. Demattei, “Training children with autism spectrum disorders to be compliant with an oral assessment,” Res. Autism Spectr. Disord., vol. 4, no. 4, pp. 681–696, 2010, doi: 10.1016/j.rasd.2010.01.007.
[29] C. Lauren and B. Enrique, “Preferences of Parents of Children with Autism Spectrum Disorders Concerning Oral Health and Dental Treatment,” vol. 34, no. December, pp. 480–484, 2012.
[30] B. Lai, M. Milano, M. W. Roberts, and S. R. Hooper, “Unmet dental needs and barriers to dental care among children with autism spectrum disorders,” J. Autism Dev. Disord., vol. 42, no. 7, pp. 1294–1303, 2012, doi: 10.1007/s10803-011-1362-2.
[31] L. Mckeown, “Social relations and breath odour,” 2003.
[32] H. C. Shane and P. D. Albert, “Electronic screen media for persons with autism spectrum disorders: Results of a survey,” J. Autism Dev. Disord., vol. 38, no. 8, pp. 1499–1508, 2008, doi: 10.1007/s10803-007-0527-5.
[33] R. C. Pennington, “Computer-assisted instruction for teaching academic skills to students with autism spectrum disorders: A review of literature,” Focus Autism Other Dev. Disabl., vol. 25, no. 4, pp. 239–248, 2010, doi: 10.1177/1088357610378291.

[34] S. Parsons and S. Cobb, "State-of-the-art of virtual reality technologies for children on the autism spectrum," Eur. J. Spec. Needs Educ., vol. 26, no. 3, pp. 355–366, 2011, doi: 10.1080/08856257.2011.593831.
[35] J. Fan, J. W. Wade, A. P. Key, Z. E. Warren, and N. Sarkar, "EEG-based affect and workload recognition in a virtual driving environment for ASD intervention," IEEE Trans. Biomed. Eng., vol. 65, no. 1, pp. 43–51, 2018, doi: 10.1109/TBME.2017.2693157.
[36] E. Bekele et al., "Multimodal adaptive social interaction in virtual environment (MASI-VR) for children with Autism spectrum disorders (ASD)," Proc. - IEEE Virtual Real., vol. 2016-July, pp. 121–130, 2016, doi: 10.1109/VR.2016.7504695.
[37] H. Zhao, Z. Zheng, A. Swanson, A. Weitlauf, Z. Warren, and N. Sarkar, "Design of a haptic-gripper virtual reality system (HG) for analyzing fine motor behaviors in children with autism," ACM Trans. Access. Comput., vol. 11, no. 4, 2018, doi: 10.1145/3231938.
[38] T. Carr and C. Lord, "A pilot study promoting participation of families with limited resources in early autism intervention," Res. Autism Spectr. Disord., vol. 25, pp. 87–96, 2016, doi: 10.1016/j.rasd.2016.02.003.
[39] E. Walsh, J. Holloway, A. McCoy, and H. Lydon, "Technology-Aided Interventions for Employment Skills in Adults with Autism Spectrum Disorder: A Systematic Review," Rev. J. Autism Dev. Disord., vol. 4, no. 1, pp. 12–25, 2017, doi: 10.1007/s40489-016-0093-x.
[40] M. M. Murray, D. J. Lewkowicz, A. Amedi, and M. T. Wallace, "Multisensory Processes: A Balancing Act across the Lifespan," Trends Neurosci., vol. 39, no. 8, pp. 567–579, 2016, doi: 10.1016/j.tins.2016.05.003.
[41] L. R. Sacrey, S. E. Bryson, and L. Zwaigenbaum, "Prospective examination of visual attention during play in infants at high-risk for autism spectrum disorder : A longitudinal study from 6 to 36 months of age," vol. 256, pp. 441–450, 2013.
[42] F. Shic, S. Macari, and K. Chawarska, "Speech Disturbs Face Scanning in 6-Month-Old," Biol. Psychiatry, vol. 75, no. 3, pp. 231–237, 2013, doi: 10.1016/j.biopsych.2013.07.009.
[43] W. Bosl, A. Tierney, H. Tager-Flusberg, and C. Nelson, "EEG complexity as a biomarker for autism spectrum disorder risk," BMC Med., vol. 9, 2011, doi: 10.1186/1741-7015-9-18.
[44] E. Grossi, C. Olivieri, and M. Buscema, "Diagnosis of autism through EEG processed by advanced computational algorithms: A pilot study," Comput. Methods Programs Biomed., vol. 142, pp. 73–79, 2017, doi: 10.1016/j.cmpb.2017.02.002.
[45] M. Ahmadlou, H. Adeli, and A. Adeli, "Fractality and a wavelet-chaos-neural network methodology for EEG-based diagnosis of autistic spectrum disorder," J. Clin. Neurophysiol., vol. 27, no. 5, pp. 328–333, 2010, doi: 10.1097/WNP.0b013e3181f40dc8.
[46] D. Alie, M. H. Mahoor, W. I. Mattson, D. R. Anderson, and D. S. Messinger, "Analysis of eye gaze pattern of infants at risk of autism spectrum disorder using Markov Models," 2011 IEEE Work. Appl. Comput. Vision, WACV 2011, pp. 282–287, 2011, doi: 10.1109/WACV.2011.5711515.
[47] S. M. Mavadati, H. Feng, A. Gutierrez, and M. H. Mahoor, "Comparing the gaze responses of children with autism and typically developed individuals in human-robot interaction," IEEE-RAS Int. Conf. Humanoid Robot., vol. 2015-Febru, pp. 1128–1133, 2015, doi: 10.1109/HUMANOIDS.2014.7041510.
[48] S. Canavan et al., "Combining Gaze and Demographic Feature Descriptors for Autism Classification," 2017.
[49] V. A. Fusaro et al., "The Potential of Accelerating Early Detection of Autism through Content Analysis of YouTube Videos," vol. 9, no. 4, pp. 1–6, 2014, doi: 10.1371/journal.pone.0093533.
[50] Q. Tariq et al., "Mobile detection of autism through machine learning on home video : A development and prospective validation study," pp. 1–20, 2018.
[51] G. P. Hanley, C. S. Jin, N. R. Vanselow, and L. A. Hanratty, "Producing meaningful improvements in problem behavior of children with autism via synthesized analyses and treatments," J. Appl. Behav. Anal., vol. 47, no. 1, pp. 16–36, 2014, doi: 10.1002/jaba.106.
[52] S. Gehlert and T. S. Ward, "Theories of health behavior," in Handbook of health social work, 2019, pp. 143–163.
[53] C. E. Bullock, W. W. Fisher, and L. P. Hagopian, "Description and Validation of a Computerized Behavioral Data Program: 'bDataPro,'" Behav. Anal., vol. 40, no. 1, pp. 275–285, 2017, doi: 10.1007/s40614-016-0079-0.
[54] Catalyst, "DataFinch." https://datafinch.com/aba-data-collection/.
[55] A. Sano, J. Hernandez, J. Deprey, M. Eckhardt, M. S. Goodwin, and R. W. Picard, "Multimodal annotation tool for challenging behaviors in people with autism spectrum disorders," UbiComp'12 - Proc. 2012 ACM Conf. Ubiquitous Comput., pp. 737–740, 2012, doi: 10.1145/2370216.2370378.

[56] J. Jessel, E. T. Ingvarsson, R. Metras, H. Kirk, and R. Whipple, “Achieving socially significant reductions in problem behavior following the interview-informed synthesized contingency analysis: A summary of 25 outpatient applications,” J. Appl. Behav. Anal., vol. 51, no. 1, pp. 130–157, 2018, doi: 10.1002/jaba.436.
[57] S. Sarabadani, L. C. Schudlo, A. A. Samadani, and A. Kushki, “Physiological Detection of Affective States in Children with Autism Spectrum Disorder,” IEEE Trans. Affect. Comput., pp. 1–14, 2018, doi: 10.1109/TAFFC.2018.2820049.
[58] F. Albinali, M. S. Goodwin, and S. Intille, “Detecting stereotypical motor movements in the classroom using accelerometry and pattern recognition algorithms,” Pervasive Mob. Comput., vol. 8, no. 1, pp. 103–114, 2012, doi: 10.1016/j.pmcj.2011.04.006.
[59] N. M. Rad and C. Furlanello, “Applying Deep Learning to Stereotypical Motor Movement Detection in Autism Spectrum Disorders,” IEEE Int. Conf. Data Min. Work. ICDMW, vol. 0, pp. 1235–1242, 2016, doi: 10.1109/ICDMW.2016.0178.
[60] B. O. Ploog, A. Scharf, D. Nelson, and P. J. Brooks, “Use of computer-assisted technologies (CAT) to enhance social, communicative, and language development in children with autism spectrum disorders,” J. Autism Dev. Disord., vol. 43, no. 2, pp. 301–322, 2013, doi: 10.1007/s10803-012-1571-3.
[61] A. L. Wainer and B. R. Ingersoll, “The use of innovative computer technology for teaching social communication to individuals with autism spectrum disorders,” Res. Autism Spectr. Disord., vol. 5, no. 1, pp. 96–107, 2011, doi: 10.1016/j.rasd.2010.08.002.
[62] L. Zhang, Q. Fu, A. Swanson, A. Weitlauf, Z. Warren, and N. Sarkar, “Design and Evaluation of a Collaborative Virtual Environment (CoMove) for Autism Spectrum Disorder Intervention,” ACM Trans. Access. Comput., vol. 11, no. 2, pp. 1–22, 2018, doi: 10.1145/3209687.
[63] J. Carmigniani, B. Furht, M. Anisetti, P. Ceravolo, E. Damiani, and M. Ivkovic, “Augmented reality technologies, systems and applications,” Multimed. Tools Appl., vol. 51, no. 1, pp. 341–377, 2011, doi: 10.1007/s11042-010-0660-6.
[64] S. C.-Y. Yuen, G. Yaoyuneyong, and E. Johnson, “Augmented Reality: An Overview and Five Directions for AR in Education,” J. Educ. Technol. Dev. Exch., vol. 4, no. 1, 2011, doi: 10.18785/jetde.0401.10.
[65] X. Casas, G. Herrera, I. Coma, and M. Fernández, “A Kinect-based Augmented Reality system for individuals with autism spectrum disorders,” GRAPP 2012 IVAPP 2012 - Proc. Int. Conf. Comput. Graph. Theory Appl. Int. Conf. Inf. Vis. Theory Appl., pp. 440–446, 2012, doi: 10.5220/0003844204400446.
[66] A. Hartholt et al., “Virtual humans in augmented reality: A first step towards real-world embedded virtual roleplayers,” HAI 2019 - Proc. 7th Int. Conf. Human-Agent Interact., pp. 205–207, 2019, doi: 10.1145/3349537.3352766.
[67] R. Liu, J. P. Salisbury, A. Vahabzadeh, and N. T. Sahin, “Feasibility of an autism-focused augmented reality smartglasses system for social communication and behavioral coaching,” Front. Pediatr., vol. 5, no. June, pp. 1–8, 2017, doi: 10.3389/fped.2017.00145.
[68] K. Khowaja et al., “Augmented reality for learning of children and adolescents with autism spectrum disorder (ASD): A systematic review,” IEEE Access, vol. 8, pp. 78779–78807, 2020, doi: 10.1109/ACCESS.2020.2986608.
[69] V. Velzen, S. Rigato, J. B. Ali, and A. J. Bremner, “Report The Neural Basis of Somatosensory Remapping Develops in Human Infancy,” pp. 1222–1226, 2014, doi: 10.1016/j.cub.2014.04.004.
[70] J. Brauer, Y. Xiao, T. Poulain, A. D. Friederici, and A. Schirmer, “Frequency of Maternal Touch Predicts Resting Activity and Connectivity of the Developing Social Brain,” no. August, pp. 3544–3552, 2016, doi: 10.1093/cercor/bhw137.
[71] M. T. Fairhurst, L. Löken, and T. Grossmann, “Physiological and Behavioral Responses Reveal 9-Month-Old Infants ’ Sensitivity to Pleasant Touch,” 2014, doi: 10.1177/0956797614527114.
[72] M. Pelaez-nogueras et al., “The Effects of Systematic Stroking versus Tickling and Poking on Infant Behavior,” 1997.
[73] F. Mcglone, J. Wessberg, and O. Hakan, “Perspective Discriminative and Affective Touch : Sensing and Feeling,” vol. 4, no. Table 1, 2014, doi: 10.1016/j.neuron.2014.05.001.
[74] S. Lopez Cazaux, G. Lefer, A. Rouches, and P. Bourdon, “Toothbrushing training programme using an iPad® for children and adolescents with autism,” Eur. Arch. Paediatr. Dent., vol. 20, no. 3, pp. 277–284, 2019, doi: 10.1007/s40368-018-0396-y.

[75] D. F. Cihak, E. J. Moore, R. E. Wright, D. D. McMahon, M. M. Gibbons, and C. Smith, “Evaluating Augmented Reality to Complete a Chain Task for Elementary Students With Autism,” J. Spec. Educ. Technol., vol. 31, no. 2, pp. 99–108, 2016, doi: 10.1177/0162643416651724.
[76] O. B. Al-Batayneh, T. S. Nazer, Y. S. Khader, and A. I. Owais, “Effectiveness of a tooth-brushing programme using the picture exchange communication system (PECS) on gingival health of children with autism spectrum disorders,” Eur. Arch. Paediatr. Dent., vol. 21, no. 2, pp. 277–283, 2020, doi: 10.1007/s40368-019-00485-x.
[77] D. Bian et al., “A Novel Multisensory Stimulation and Data Capture System (MADCAP) for Investigating Sensory Trajectories in Infancy,” IEEE Trans. Neural Syst. Rehabil. Eng., vol. 26, no. 8, pp. 1526–1534, 2018, doi: 10.1109/TNSRE.2018.2854672.
[78] D. Bian, Z. Zheng, A. Swanson, and A. Weitlauf, “Design of a Multisensory Stimulus Delivery System for Investigating Response Trajectories in Infancy,” vol. 1, pp. 471–480, doi: 10.1007/978-3-319-58706-6.
[79] Z. K. Zheng, D. Bian, A. Swanson, and A. Weitlauf, “Soft-Brush : A Novel Tendon Driven Tactile Stimulator for Affective Touch in Children with Autism,” vol. 1, pp. 15–22, doi: 10.1007/978-3-319-94373-2.
[80] C. Triscoli, H. Olausson, U. Sailer, H. Ignell, and I. Croy, “CT-optimized skin stroking delivered by hand or robot is comparable,” vol. 7, no. December, pp. 1–5, 2013, doi: 10.3389/fnbeh.2013.00208.
[81] R. L. Gabriels, M. L. Cuccaro, D. E. Hill, B. J. Ivers, and E. Goldson, “Repetitive behaviors in autism: Relationships with associated clinical features,” Res. Dev. Disabil., 2005, doi: 10.1016/j.ridd.2004.05.003.

# Chapter 2: Novel Multisensory Stimulation and Data Capture System (MADCAP) for Investigating Sensory Trajectories in Infancy


## 2.1 Abstract

Sensory processing differences, including auditory, visual and tactile, are ideal targets for early detection of neurodevelopmental risks, such as autism spectrum disorder. However, most existing studies focus on the audiovisual paradigm and ignore the sense of touch. In this project, we investigated the multimodal sensory trajectories of infants at high risk of autism. We present a multisensory delivery system that can deliver audio, visual, and tactile stimuli in a controlled manner and capture peripheral physiological, eye gaze and electroencephalogram response data. The novelty of the system is the ability to provide affective touch. In particular, we have developed a tactile stimulation device for infants to deliver tactile stimulus with precisely controlled brush stroking speed and pressure on the skin. A usability study of ten 3-20 month-old infants was conducted to investigate the tolerability and feasibility of the system. Results have shown that the system is well tolerated by infants and all the data were collected robustly.

Building on the promising system feasibility results, we further improved the system with a new generation of tactile stimulator, Soft-Brush, which has a soft-sleeve shape design that conveniently wraps around forearms of infants. Soft-Brush is much lighter and more compact and comfortable. We ran experimental validation on Soft-Brush and it significantly improved the tolerance rate and had noticeable less distractions during data collection. Thus, we ran a pilot study on 29 toddlers aging from 2-4 years, with 12 in the typically developed group and 17 in the autistic group. With machine learning approach, we found out that adding the new channel of stimulation, affective touch, provided additional predictive power. The MADCAP extracted features combined with the existing behavioral scores showed improved predicting accuracies. This work paves the way for future studies charting the meaning of sensory response trajectories in infancy.


## 2.2 Introduction

Autism spectrum disorders (ASD) is a common neurodevelopmental disability characterized by social and communication impairments and is associated with costly human experience and financial impact [1]. Despite the fact that a reliable diagnosis of ASD can be made by the age of 2 years, with many symptoms evident much earlier, most children are not accurately identified with ASD until after age four due to multiple factors, including difficulties in accessing care and a lack of trained providers [2]. Consequently, these children do not receive early intervention in the first couple years of life, a time period recognized as optimal for enhancing developmental outcomes due to neural plasticity [3]. Although the neural basis of complex social and communicative behaviors develops over the course of childhood, brain response to more basic sensory stimuli, such as touch, sight, smell, are present much earlier, even during the first few months of life, which is long before the observable behavioral and communication symptoms of ASD become apparent [4]. Given that hypo- and hyper-responsiveness to sensory input is a core diagnostic feature of ASD that can cause significant impairment over time [5][6], it logically follows that children at risk of ASD or other neurodevelopmental disorders may show subtle sensory differences within the first year of life, earlier than ASD can reliably be diagnosed at present, creating an opportunity to identify those children at an earlier age. Thus, they may benefit from closer developmental monitoring.

Existing prospective studies of high-risk infant siblings of children with ASD (Sibs-ASD) suggest that sensory differences related to visual processing clearly emerge in the first two years of life [5]. A growing number of studies have investigated visual attention to faces and other social stimuli in Sibs-ASD [7][8][9]. For both high- and low-risk infants, most of these studies have described early point-in-time group level similarities on simple performance measurements of visual scanning and preferential looking to core facial features. These studies have also suggested that high-risk infants may show subtle processing differences in the brain-based mechanisms for responding to these stimuli. These subtle processing differences may contribute to neurodevelopmental impairment over time. For example, for those infants who are eventually diagnosed with ASD, attention directed to eyes during infant-directed audiovisual speech initially appears intact but declines from 2 to 6 months of age, a pattern not observed in infants who do not develop ASD [7].

These important findings identify a potentially critical developmental trajectory of decreased visual attention in high-risk infants. However, existing social attention paradigms are limited by their focus on solely audio-visual modalities of early sensory learning. We know that neural mechanisms for processing various sensory inputs, such as tactile, vestibular, and auditory inputs, start to come online prenatally, playing an immediate role in the postnatal social-sensory experiences that lay a foundation for multisensory processing and social learning over time [10]. Paradigms using additional sensory processing channels to augment existing visual attention findings may provide more robust methods for detecting actionable neurodevelopmental risk at earlier time points.

One sensory processing channel that could augment existing visual attention work is related to tactile perception, or sense of touch. The sense of touch is widely known for the role it plays in discriminating and identifying external stimuli. For example, by investigating the neural basis of somatosensory remapping in human infancy, Silvia et al. found that the cortical networks underlying the ability to dynamically update the location of a perceived touch across limb movements become functional during the first year of life [11]. However, there is growing evidence that the sense of touch has another dimension, also known as "affective touch," which conveys social information just like what someone sees and hears [12]. Affective touch, which can be described as a comforting, caress-like soft touch (e.g., maternal touch), has been found to impact the "social brain" [13]. Previous research has demonstrated that infants are sensitive to affective touch [14] and that compared to other forms of touch, stroking an infant can not only induce positive emotions but also modulate negative ones [15]. In working with adults with ASD, Croy et al. found that they show atypical perception and processing of affective touch. Additionally, the authors hypothesized that the affective touch functionality, which is based on C tactile fiber activation [16], is impaired to some extent in individuals with ASD [17]. Furthermore, Kaiser et al. demonstrated that in the presence of affective touch stimulus, individuals with ASD exhibit reduced brain activity in social-emotional-related brain regions compared to typically developing (TD) individuals [18]. As such, affective touch represents an identified area of atypical sensory processing related to ASD that can also influence infant response, making it an optimal target for early detection of neurodevelopmental risk.

One of the primary contributions of the current work is to design an automated mechanism for affective touch and then, integrate it with an audio-visual stimuli presentation system to study multisensory processing of infants at high- and low-risk of ASD. It is difficult to produce skin-to-skin affective touch in laboratory settings. However, an analogous tactile stimulation which is produced by a mechanical source (e.g., soft brushing) is comparable to an affective touch that is manually produced by hand [19]. Previous tactile stimulation work in infants has utilized trained human confederates to administer pleasant social touch via dorsal forearm stroking with Hake brushes at predetermined velocities and pressures [14]. Although adequate for documenting generalized physiological responses to the stimulus, this manual control has several limitations: speed and pressure are not precisely controlled or measured, stroking is hard to coordinate with other stimuli/measurements, and the human presence may confound certain experimental paradigms.

In this chapter, we present the design and development of a novel multisensory stimulation and data capture system (MADCAP) for infants that delivers multiple sensory stimuli and simultaneously captures multi-dimensional data. In particular, we have developed two generations of tactile stimulation devices for infants to deliver tactile stimulus with precisely controlled brush stroking speed and pressure on the skin. A pilot study was conducted to demonstrate the system feasibility and collect multimodal data. With that data, machine learning models were built to and classify ASD diagnosis. To the best of our knowledge, this is the first work to demonstrate a multimodal technological system including affective touch that has the potential to meaningfully chart differences in coordinating visual, auditory, and tactile processing in infancy.

### 2.3 Multisensory Stimulation and Data Capture (MADCAP) System Design

MADCAP includes three main components: a multisensory stimulation delivery module consisting of auditory, visual, and tactile stimuli, a multi-dimensional data capture module capturing a participant's response, and a supervisory controller module to synchronize the interconnection between the two modules. Fig. 1 shows the schematic of the overall system.

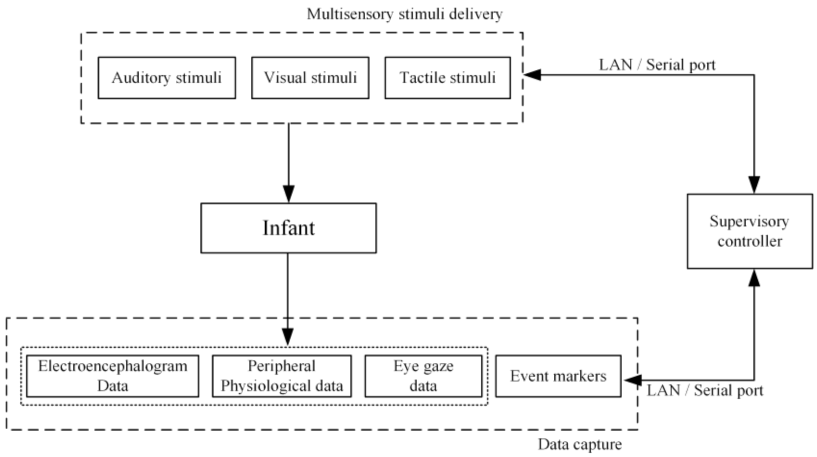



**Fig. 1.** MADCAP system schematic

*2.3.1. Multisensory stimulation delivery module*

The presented multisensory stimulation delivery module is capable of delivering both synchronized and asynchronized tactile, auditory and visual stimuli to the subjects. While the audio-visual stimuli delivery subsystem is not entirely novel, the tactile stimulation device for affective touch is the first prototype in present and is discussed in detail.

*1) Tactile stimulation device*

In this work, we have developed a tactile stimulation device (TSD) specifically for infants to deliver affective touch at precisely controlled brushing speed and pressure. Since this device is meant for infants, special consideration was given to the size, speed, comfort, and safety of the device, which are described below.

TSD is designed to impart affective touch on the dorsal side of the forearm of infants under 24 months of age. The housing of the system needs to provide ample, comfortable space where infants can place their arm in it to be stroked by a soft brush. In order to determine the overall size of the device, we considered both the arm circumference and arm length for infants under the age of 24 months. The arm circumference median with *z*-scores of 3 is 18.45cm and the average arm length is 13cm [20]. The compartment that holds the arm should have a variable radius to accommodate different subjects. The brush should have a stroke length longer than the maximum possible forearm length of infants within the age range.

Brush speed is one of the key variables in creating affective touch. In previous studies, brush stroking speed in the range of 1 cm/s to 10 cm/s was found to be most pleasant and effective in producing affective sensation [21]. We designed our system to cover this speed range with more precision, although if necessary for future work, the speed could be increased.

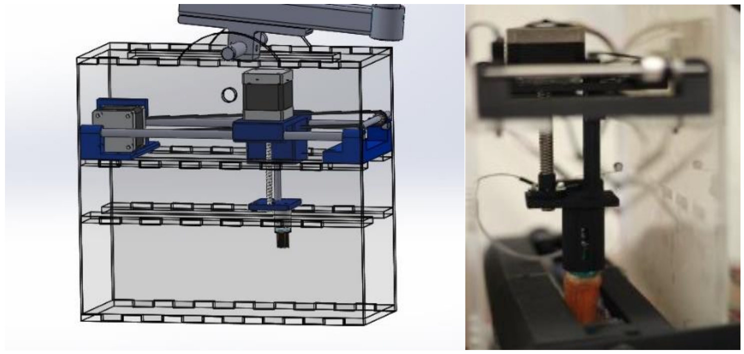


**Fig. 2.** a) A CAD drawing of the mechanism; and b) the fabricated system

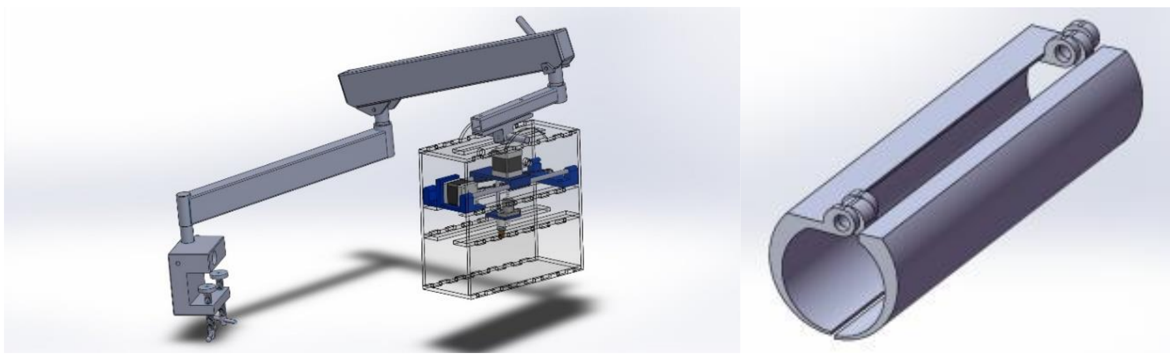


**Fig. 3.** a) The gravity compensated manipulator supporting the weight of device; and b) arm holder

Comfort is also important for a device used by infants because they may feel frustrated or afraid when their range of motion is restricted (e.g., arms restrained within a mechanical system). As a result, we designed the tactile stimulation device such that an infant can move her arm, along with the device, freely in 3-D space while remaining attached. The design also specifically eliminated sharp edges and rough surfaces. Moving mechanical elements were placed in a space that was out of infants' reach as well as sight.

Based on the above considerations, the TSD was designed and constructed as shown in Fig. 2. The device has three compartments from top to bottom. The top compartment contains actuators that control the brush. The middle compartment includes a replaceable soft brush. A force sensing resistor is attached to the bottom of the brush so that the pressure applied on the arm from the brush can be measured and modulated. The infant's arm rests in the bottom compartment. The infant's forearm is placed into a soft strap to hold it in place and guarantee that the brush will contact the forearm. The device is attached to a gravity compensating articulating arm which gives the infant a certain degree of freedom to move his/her arm while the relative position between the arm and the device remains the same (Fig. 3a).

For the mechanical design, the brushing mechanism has two degrees of freedom, a horizontal motion to move the brush along the length of the forearm and a vertical motion to move the brush up and down to control contact pressure as shown in Fig. 2a. Two stepper motors (NEMA 17) were selected to actuate the brush in both directions. The precise positioning of the stepper motor allows the brushing movement to be precisely controlled and repeated. The stepper motor that controls the horizontal movement, was chosen in such a way that the brush could reach a maximum horizontal speed of 13cm/s, which covered the most effective stroking speed range to produce affective touch. A lead screw stepper motor was chosen to control the vertical movement of the brushing mechanism. Lead screw converts a turning motion into a linear motion and provides precise positioning, which is suitable for the vertical movement of the brushing mechanism. In the middle compartment, the brush is attached to the device through a Lego brick-type design, so that brushes with various size and material can be attached and detached easily. The brushing pressure control mechanism lies in the bottom part of the brush, which is discussed later in more detail.

The infant's arm is held inside an arm holder in the bottom compartment. The infant is not able to reach or see any sharp parts like the lead screw or the motors in the upper two compartments. The arm holder, as shown in Fig. 3b, was designed and 3D-printed to keep the relative position between the arm and the assembly constant while the assembly moves in 3D space. The holder could open and close to adapt to the sizes of different subjects. The holder had a gradient radius from 35mm to 40mm to better fit the shape of an arm. Soft cushions were attached to the inner surface of the holder for better comfort.

As shown in Fig. 3a, the assembly described above is hung on a mechanical manipulator (Dectron, USA). By adjusting the spring inside the rod, the manipulator could compensate for the weight of the assembly as well as for the baby's arm. Thus, the infant could freely move her arm with the attached device.

For the electronic design and control, an Arduino-based microcontroller circuit board (chipKIT uC32) was used to control the two stepper motors because it has a faster processor than the traditional Arduino board, thereby guaranteeing smooth movement of the two stepper motors. We also chose two motor driver boards (EasyDriver) to drive the stepper motors and two limit switches (D2VW-5L2-1HS, Omron) to set up their initial positions. The back and forth motion was controlled in an open-loop manner since the stepper motor was fairly accurate in maintaining speed. The pressure on the skin was controlled in a closed-loop manner.

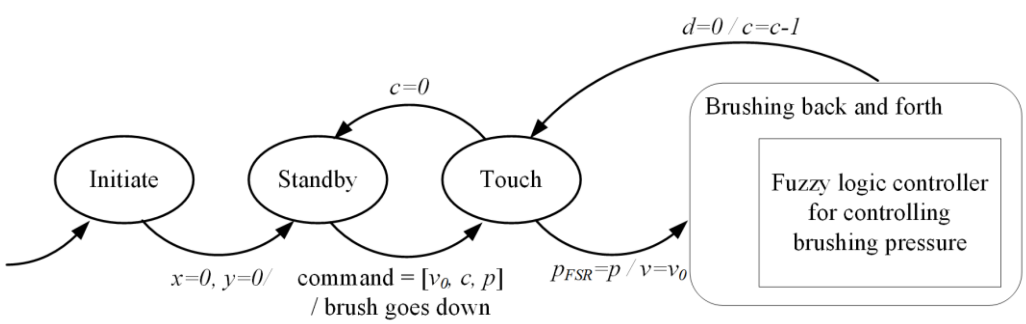



**Fig. 4.** FSM for controlling the horizontal movement of the brush.

The TSD is designed to simulate affective touch by stroking an infant's arm with soft brush back and forth at a given speed and pressure. As such, we developed the TSD to receive the control command, including translational speed (*v0*), brushing counts (*c*), and brushing pressure (*f*), to perform back and forth stroking based on the Finite State Machine (FSM) shown in Fig. 4. The *x*, *y* are the horizontal and vertical positions of the brush, *c* is the brushing count, *v0* is the brushing speed, *f* is the brushing pressure, *fFSR* is the pressure computed from force sensor, *d* is the horizontal distance of the brush away from the initial position.

The FSM starts with the "Initiate" state, where the lead screw goes up and the translational stepper motors goes back, until both limit switches are triggered to set the initial positions. The FSM then moves into the "Standby" state, waiting for a command. When a valid command comes in, the "Touch" state is triggered and the brush goes down until the force on the brush reaches the desired value. In the brushing state, the brush strokes the baby's arm back and forth and the counter subtracts one stroke after each round of stroking. The brush goes back to the "Standby" state once the desired stroke counts is completed.

To achieve consistent brushing pressure and a smooth actuation of the lead screw during the "Brushing back and forth" state, a fuzzy logic controller was implemented. Fuzzy logic control is a natural choice for our application since instead of applying pressure that varies continuously, we are interested in applying pressures to the forearm that can be categorized as "soft", "medium", and "high." Compared to other control methods, fuzzy logic control maintains the variable in a desired range rather than a specified value. This approach combines regulation algorithms with logical reasoning and reduces mechanical chattering so that the system has a smoother motion and less noise.

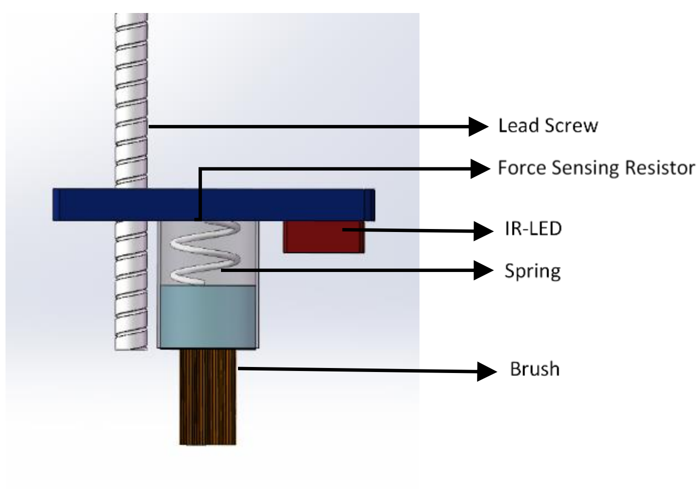



**Fig. 5.** Illustration of the brush and sensors for maintaining consistent brushing pressure. Force Sensing Resistors (FSR 402, Interlink Electronics) were integrated at the base of the brush to measure the pressure at the end of the brush. Infrared LED (GP2Y0A51SK0F, Sharp) was used to measure the relative position between brush and the arm.

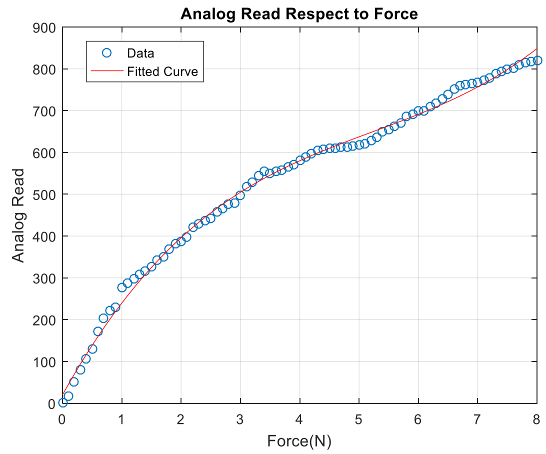



**Fig. 6.** The relationship between readings of the force sensor and actual force upon the sensor

In order to apply controlled brushing pressure on the forearm, allowing the system to control the brush such that it touches the forearm gently and without abrupt pressure, we implemented a mechanism to measure the contact pressure as the input of the fuzzy logic controller. The mechanism is illustrated in Fig. 5. It contains a brush, a platform connected to the lead screw, an IR-LED, and a force sensing resistor (FSR). The lead screw actuates the brush along the vertical direction based on the output of the fuzzy logic controller. Before the brush touches the forearm, the IR-LED measures the distance between the arm surface and the end of the brush to control the vertical speed towards the arm. The IR-LED has a very accurate linear measurement in the distance range from 2-15cm and it is suitable for this purpose. The end of the brush is connected to the FSR through a spring which is used to avoid the disturbance from tilting during brushing, thereby assuring an exact reading of the normal pressure. In order to map the sensor analog reading to the normal pressure, we calibrated the FSR to get the mappings as shown in Fig. 6. Gaussian membership functions were used to map the contact pressure to the five states because it best fits the data with the actual tactual feeling. Thus the fuzzy logic controller has five input and output states. The input states are: None, Soft, Medium, Hard, and Heavy; the output states are: High Speed Down, Low Speed Down, Stop, Low Speed Up and High Speed Up.

A real-time input is mapped to the state which has the largest membership function value. The membership functions are derived based on real-life experiments, as shown in Fig. 7. The membership function values, which varies from 0 to 1, stand for the likelihood of an input or output belonging to one state.

Fuzzy logic rules generate appropriate output with respect to the input. To maintain different pressure of touch, rules are defined and tested. For example, the rules to produce medium touch are defined as follows:

1. If Pressure is None, then Brush speed is High Speed Down.

2. If Pressure is Soft, then Brush speed is Low Speed Down.

3. If Pressure is Medium, then Brush speed is Stop.

4. If Pressure is Hard, then Brush speed is Low Speed Up.

5. If Pressure is Heavy, then Brush speed is High Speed Up.

These rules generate an appropriate output for real-time inputs to keep the brushing pressure in a desired range.

*2) Audiovisual stimulation delivery module*

We designed a custom audiovisual stimulation delivery module to present clips of an adult female narrating a short story to the infant in our usability study (Section IV-B). This module has the ability to communicate with various other stimulation delivery module and the data capture modules.

Unity (https://unity3d.com) software was used to implement the module. A finite state machine (Fig. 8) was used to control the logic of the stimuli presentation. All parameters of the stimuli delivery such as the content of the audiovisual stimuli, the duration of rest between each stimulus, and the dose of the stimuli, could be adjusted easily according to the experimental protocol. In addition, the user defined event markers, such as start/end experiment and start/end stimulus, were logged into files using the JSON (www.json.org) format.

This audiovisual stimulation delivery module also communicated with the tactile stimulation device via the supervisory controller to make sure multisensory stimuli were properly synchronized. Furthermore, this module sent the user-defined event markers to the data capture module for later data analysis. We will discuss the details of the inter-module communications in Section II-C.

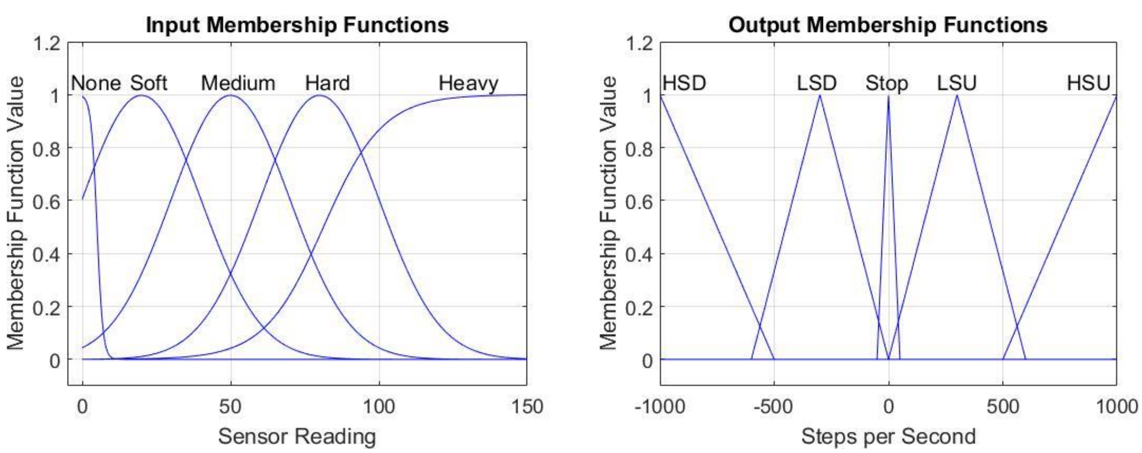



**Fig. 7.** The membership functions of the input and output.

HSD: High Speed Down; LSD: Low Speed Down; LSU: Low Speed Up; HSU: High Speed Up

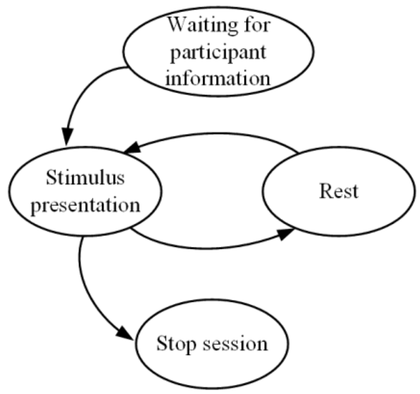



**Fig. 8.** Audiovisual delivery module FSM. The program starts in the “waiting for participant information” state. After the experimenter inputs the participant information, the program start the “Stimulus presentation” state. The participant has a certain amount of time to rest between each stimulus. After several rounds of stimulus presentation according to the protocol, the program will come to the “Stop session” state and record the data.

*2.3.2 Data capture module*

MADCAP has the ability to track eye gaze of infants when they looked at the audiovisual stimuli. The Tobii X120 eye tracker (Tobii Pro AB), which has a sampling rate of 120 Hz, was used to measure gaze position across defined regions within the audiovisual presentation. The (X, Y) coordinates of the gaze position—(0, 0) for upper left corner—as well as time stamps and event markers were recorded.

The infant’s peripheral physiological data, including blood volume pulse (BVP) and electrodermal activity (EDA), were recorded by using the E4 wristband (Empatica Inc.), which is an unobtrusive device suitable for infant study. The E4 wristband is worn on the infant’s ankle during recording. The sampling rate for BVP and EDA is 64Hz and 4Hz respectively. By using the hardware API provided by the E4 wristband, we developed a custom program in C# for the E4 wristband to record time-stamped physiological data within the protocol. The physiological data measured by the E4 wristband were streamed to the custom program wirelessly via Bluetooth.

In addition, electroencephalogram (EEG) data was recorded using a dense-array EEG system (EGI Inc.). EEG was recorded from a 128 channel Geodesic sensor net and the sampling rate was 1000Hz for each channel. All 128 channels were recorded continuously and event markers were registered via serial port for off-line segmentation of the data. The vertex electrode was used as a reference and the data were later re-referenced to an average reference.

*2.3.3. Supervisory controller module*

As can be seen from the system diagram (Fig. ), the supervisory controller module served as a bridge between the multisensory stimulation delivery and the data capture modules. It played a crucial role in the MADCAP system, making sure all the stimulations were presented in a time-synchronized manner and all the data capture were properly controlled based on specific events and time stamp.

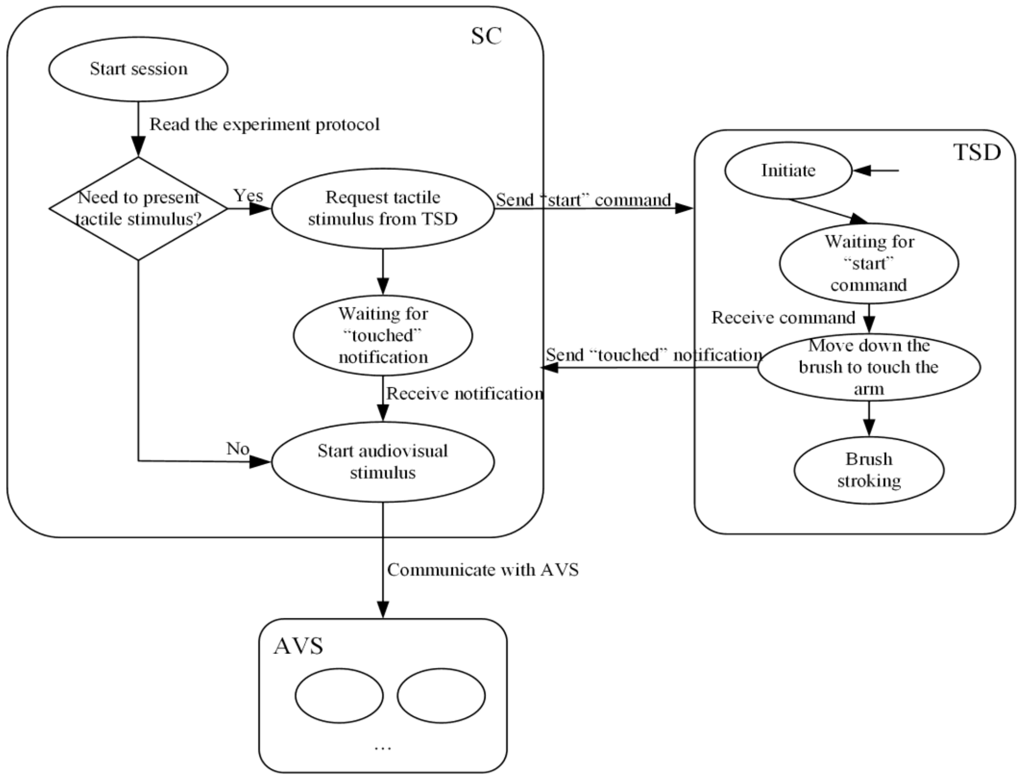



**Fig.9.** Supervisory controller FSM to synchronize stimuli deliveries.

SC: Supervisory Controller module; TSD: Tactile Stimulation Device; AVS: AudioVisual Stimuli presentation module

The supervisory controller program communicates with the tactile stimulation device through the serial port at a baud rate of 9600 bits/second. As depicted in Fig. 9, a command is sent to the tactile stimulation device to initiate the brush stroking when a tactile stimulus is needed. The brush stroking does not start immediately because it must first be moved vertically to come into contact with the infant's arm, as described in Section II-A-1. As soon as the brush touches the infant's arm and the pressure sensor detects the touch, the tactile stimulation device sends a notification message to the supervisory controller. Then the supervisory controller initiates the audiovisual stimulus. In this way, we guarantee the audiovisual and tactile stimuli fire at the same time.

The supervisory controller communicates with both physiological data recording program and eye gaze data recording program over a TCP/IP based socket in a Local Area Network (LAN). For EEG data recording, event markers can be added by inputting Transistor-Transistor Logic (TTL) pulse to the amplifier, thus we designed a circuit board which can be controlled by supervisory controller program to generate different kinds of TTL pulse to represent different event markers.

Throughout the experiments, the supervisory controller monitored the stimulation delivery modules. When a user defined event occurred (e.g., start brush stroking/audiovisual stimulus), the supervisory controller sent an event marker in JSON format to data capture modules over socket or serial port.

## 2.4 Tactile Stimulation Device Validation

TSD was validated in two steps. First TSD was validated separately in MATLAB Simulink. Then we tested the real-world performance of TSD on adults.

### *2.4.1. Simulation Results*

MATLAB Simulink was used to simulate and validate the fuzzy logic controller that maitains the brushing pressure of TSD. The basic concept of this simulation is to stroke an uneven surface, much more uneven than a human arm, and to see whether the brushing pressure still stays consistent. An additional 10 Hz Gaussian noise was introduced in the simulation to create realistic situations.

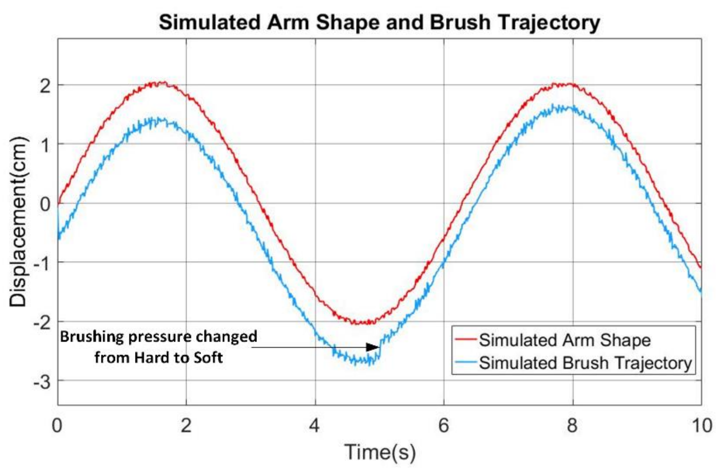



**Fig. 10.** Simulink results

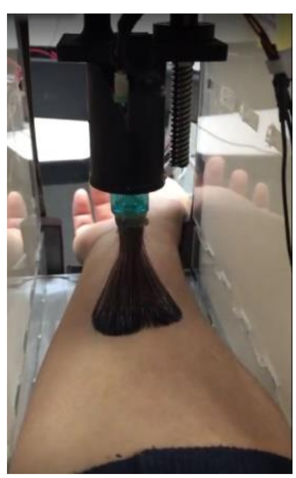


**Fig 11.** Human experiment

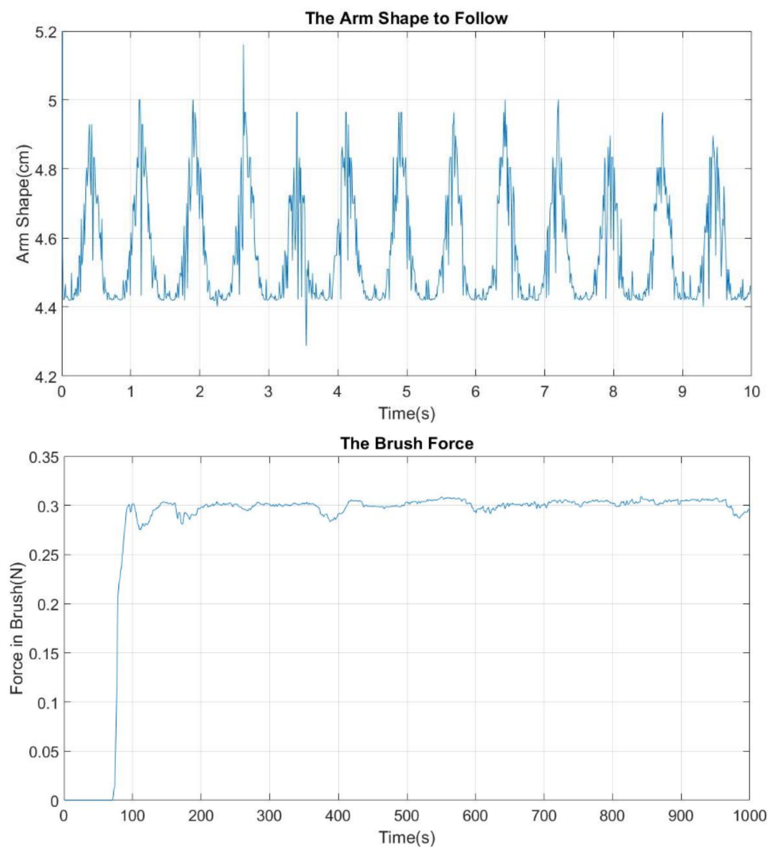



**Fig. 12.** Real system experiment and results

Simulations of the brush stroking under different brushing conditions were run and the results are shown in Fig. 10. The simulation used an extremely wavy arm shape with a peak to peak value of 4 cm which is much greater than the possible variance of arm radius. The figures demonstrates the simulation results of one brush stroke (10cm). In the middle point of the stroke (5cm), the brushing pressure changed from Hard to Soft. The top figure shows the brush trajectory. The red curve simulates the shape of the arm and the blue curve is the moving trajectory of the brush. We can see that the blue curve follows the shape of the red curve while keeping a vertical distance lower than the red curve meaning that the brush bristles are bent. After time 5s, the brushing pressure is adjusted to soft and the vertical distance between the two curves is shorter meaning that the brush bristles are bent less. The relative position between the red curve and blue curve before and after 5cm point stays relatively constant. The bottom figure shows the brush speed in the vertical direction. At the 5cm point, there is an obvious speed up in the negative direction, which means the brush moves backward to release some pressure on the surface. In addition, the controller had a rather short adjustment time and the relative position became stable in a very short time after starting the simulation.

*2.4.2 Real System Results*

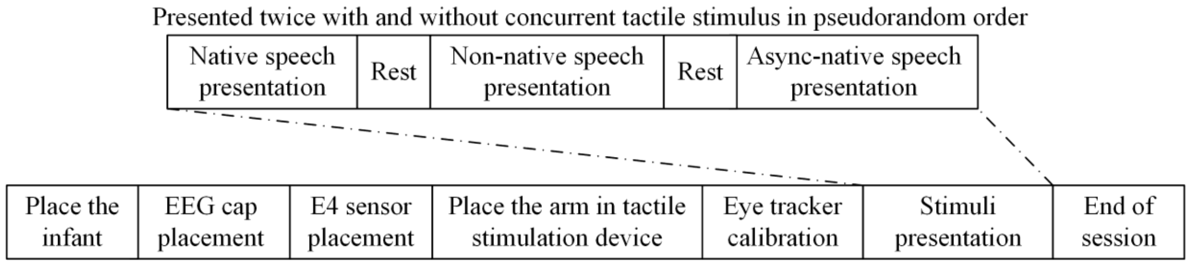



**Fig. 13.** Procedure of the usability study

The real system validation was done on the arms of adults without the arm holder as shown in Fig. 11. The arm of the tested adult should be comparable to an infant's arm inside the arm holder. The brush shook slightly and the reason might be the noise in the sensor readings. The arm shape to follow which is measured by the IR-LED during horizontal reciprocation and the force in the brush during the motion is shown in Fig. 12. The test subjects reported consistent brushing pressure during the stroking procedure.

## 2.5 MADCAP Usability Study

In order to demonstrate the tolerability and feasibility of MADCAP, we tested our integrated audio, visual, and tactile protocol across a sample of 6 infants from 3-8 months of age (3 girls, 3 boys; mean age = 5.45 months, SD = 1.60). The protocol was reviewed and approved by the Institutional Review Board (IRB) at Vanderbilt University. After receiving a thorough explanation of the experiment, parents gave written informed consent for their children's participation.

### *2.5.1 Experimental Setup*

Within a sound-attenuated room infants were seated in an infant/toddler seat (appropriate to age) and positioned 50 cm from the LCD monitor, video recording device, and eye tracker. If parents requested to have the infant on the lap or if the infant refused to sit in an infant seat, the parent would be permitted to hold their infant in their lap but was asked to minimally distract the infant. The Geodesic EEG cap was placed on the infant's head. Then the infant's left forearm was positioned within the bottom compartment of the tactile stimulation device and velcro strapped. The E4 wristband was attached to the right ankle (Fig. 14.) Subsequently, the eye tracker calibration was accomplished by using presenting audio-enabled, animated cartoon pictures. We used 2 or 5-point calibration procedures depending on the infant's age [20]. The infants then participated in a single session lasting approximately 10 minutes where they were exposed to three distinct presentations of audiovisual speech stimuli (each 50s in length).

### *2.5.2 Procedure*

The audiovisual speech presentations were the same stimuli utilized by Lewkowicz [21] to demonstrate the multisensory coherence of fluent audiovisual speech. Specifically, the stimuli presented included clips of an adult female narrating a short story in English (native tongue), Spanish (non-native tongue), and audio-asynchronous native (English, with timing of audio presentation delayed 500ms relative to video). Each audiovisual speech recording was presented with and without a concurrent tactile stimulation. In the presence of tactile stimulation, the infant's dorsal forearm was continuously stroked (i.e., back and forth) at a speed of 3cm/s, the pressure around 0.2N, with a soft makeup brush. The stroking speed and pressure were chosen based on existing studies which have shown that gentle touch at medium speed produces the most pleasant effect. The task stimuli were presented in pseudorandom order and the infant received 10-second of rest between each stimulus to reduce sensory habituation. The procedure of the usability study is shown in Fig. 13.

*2.5.3 Usability Study Results*

The target population for MADCAP is infants under the age of two years. Because of the vulnerability of this population as well as their inherent difficulty understanding experimental expectations, establishing tolerability and feasibility is crucial in determining if additional data can be collected robustly.

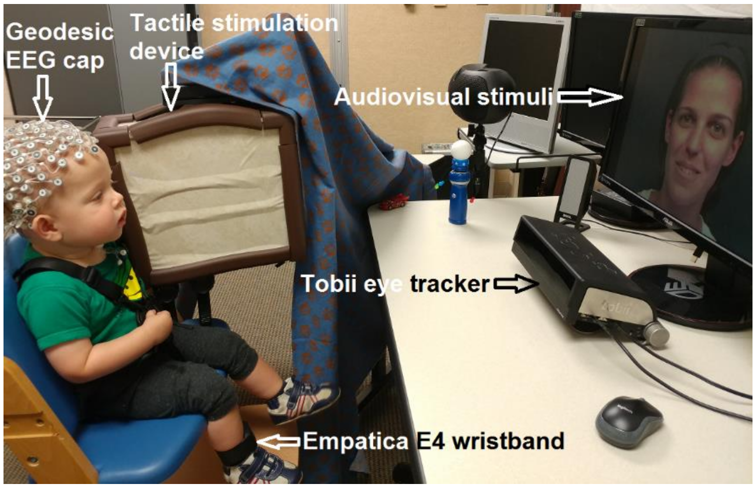



**Fig. 14.** MADCAP experimental setup.

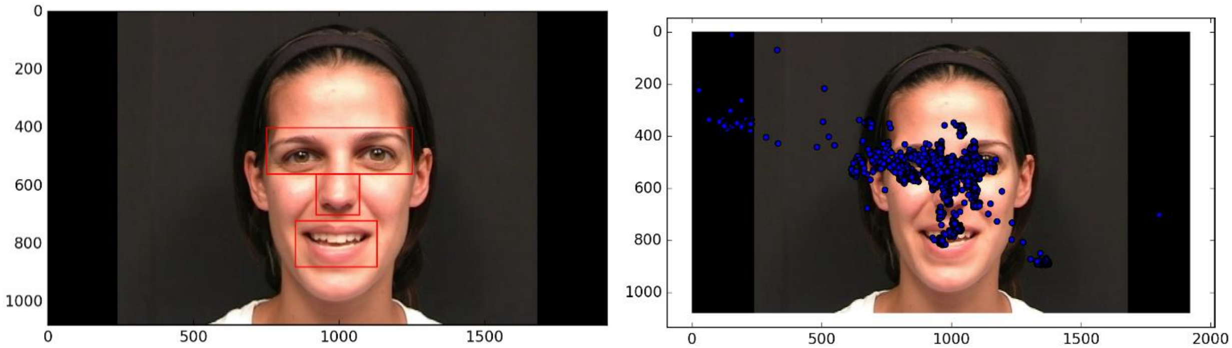



**Fig. 15.** a) Defined ROI in speech stimulus; and b) an example of gaze position for one session

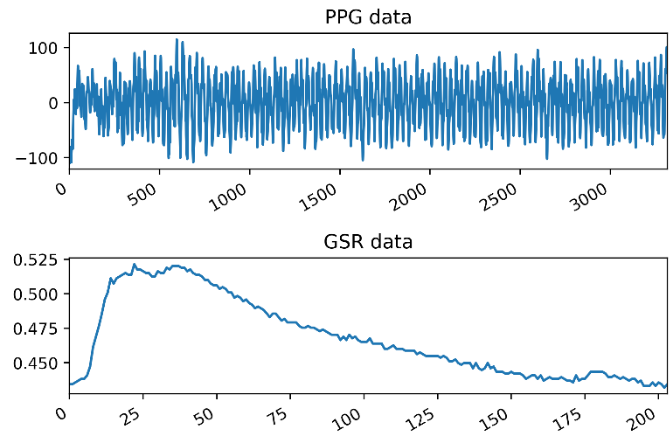



**Fig 16.** An example of physiological data from one trial (50s)

*1) Tolerability and feasibility of the system*

Results of this study demonstrated that 7 out of 10 infants completed the whole procedure. Of the three participants who did not complete the session, one could not tolerate the EEG cap. The other two participants, who had smaller arms, pulled their arms out of the tactile device in the middle of the experiments. This indicated that the current design will require modifications in future work in order to comfortably yet effectively restrict the movement of children's arms.

Eye gaze data, physiological data, and EEG data were otherwise robustly collected with event markers.

Of note, we found that the rotation of the motor generated a small noise. It started after the computer program sent the command 'start tactile stimulus' and before the tactile stimulus onset. It was caused by the rotation of the lead screw stepper motor. The noise was not loud enough to distract the infants and was greatly reduced after the brush started to stroke the infant's forearm.

*2) Data collection*

Eye gaze data were collected at a sampling rate of 120Hz. Examining how infants scan social stimuli provides us with valuable information about the distribution of interest and attention. In this study, we defined three Regions of Interest (ROI) surrounding eyes, nose, and mouth [9] (Fig. 15a). We focused on how much time infants spent fixating on these ROIs. Since infants often varied in the amount of time they spent looking at the stimuli, the fixation data were normalized (looking time to ROIs divided by total looking time to the stimuli) for each stimulus presentation.

Results indicated that, across all the sessions, participants looked at the stimulus screen 27% of the time with gaze toward demarcated ROI for 57% of this time (Fig. 15b.) In terms of coherent and delayed audiovisual stimuli, the participants spent similar percentages of time looking at the ROIs (56% for coherent audiovisual stimuli and 57% for delayed audiovisual stimuli.)

Peripheral physiological data including PPG and EDA were recorded throughout the whole session. The sampling rates for PPG and EDA are 64Hz and 4Hz, respectively. The data were examined and multiple features were successfully extracted. Heart rate (HR) was computed by detecting peaks in the PPG signal. Tonic and phasic components of EDA were decomposed separately from the original signal. The tonic component is the baseline level of EDA and is generally referred to as skin conductance level (SCL). The phasic component is the part of the signal that changes when stimuli are presented and is known as skin conductance response (SCR). We documented the average HR and SCR for the participants with the presence of affective touch (HR: 127.02, SCR: 3.5) and without the presence of affective touch (HR: 133.74, SCR: 3.74). Fig. 16 shows an example of the recorded physiological data from one trial, which is about 50 seconds. EEG data (Fig. 16) were recorded at the sampling rate of 1000 Hz with event markers. Several common EEG features, such as delta band, theta band, alpha band, and beta band, were extracted for future data analysis.

## 2.6 Soft-Brush: A Novel Tendon Driven Tactile Stimulator for Affective Touch Delivery

From the usability study of MADCAP, the first generation prototype of tactile stimulator has a few limitations. The TSD is overwhelmingly big and rigid for infants, which has a box-shaped design with dimensions of 28cm x 12cm x 30cm. The box hangs on from a gravity-compensated supporting manipulator that guarantees the box could move with the infant’s arm freely in 3D space. However, the inertia of the box and damping of the manipulator needs a certain force for motion and sometimes it is too much for the infants. Study with infants has shown that limitation of motion may lead infants to worse mental conditions [22] such as car seat crying. In general, the researchers need to touch the infants’ arms during experimental setup with MADCAP but study has shown that almost all babies generate wariness and even fear when touched by strangers [23]. Even though there are no sharp edges in the TSD and there is cloth cover on surfaces, the contacting surfaces are still not soft and comfortable enough.

The target population for MADCAP study is infants under the age of two. Tolerability due to comfort and safety is especially important to the nature of the infant participants. To ensure a higher tolerance rate, a more soft and comfortable affective touch stimulator is needed. Meanwhile, the tactile stimulator should be more compact and light-weight so that the infants do not feel constrained to move their arms. In this section, we present Soft-Brush, a novel infant-friendly silicone tactile stimulator with a tendon-actuated mechanism that can be worn like a sleeve.

### *2.6.1. Soft-Brush System Design*

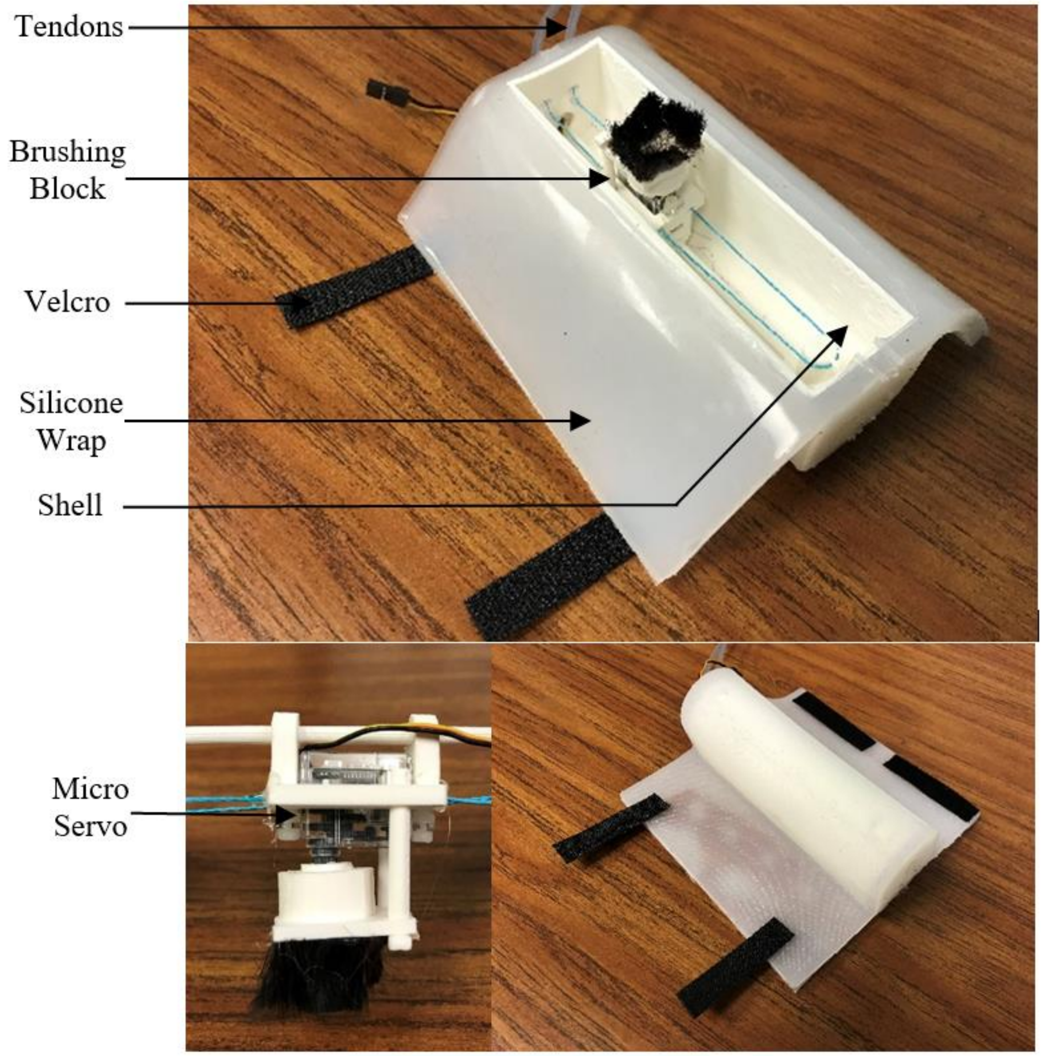



**Fig 17.** Soft-Brush and its brushing block

Soft-Brush consists of three main components, which are a silicone wrap, a brushing block and an actuating part, as shown in Fig. 17. The combined weight of the silicone wrap and the brushing block is only 161 grams, which is far less than the TSD. The lightweight design allows the babies to move their arms much more freely.

The silicone wrap has a sleeve-type design and wraps around an infant's forearm. The soft plastic rubber (Ecoflex 00-30), which is skin-safe and comfortable to touch, was used to make the wrap. A mold shown in Fig. 18 was first designed and 3D printed to get the exact shape of the wrap. After curing, we sewed four Velcro strips on both sides of the wrap to stick them together. This also allows us to use the wrap on infants with a wide range of arm sizes, from 3-months to 2-years-olds. Then we sewed a 3D-printed shell inside the chamber to keep its shape in case that the infant tries to squeeze the silicone wrap. The wrap can be soaked into warm water and it is very easy to apply Lysol and baby wipes for disinfection.

**Fig. 18.** Silicone wrap curing in 3D-printed mold

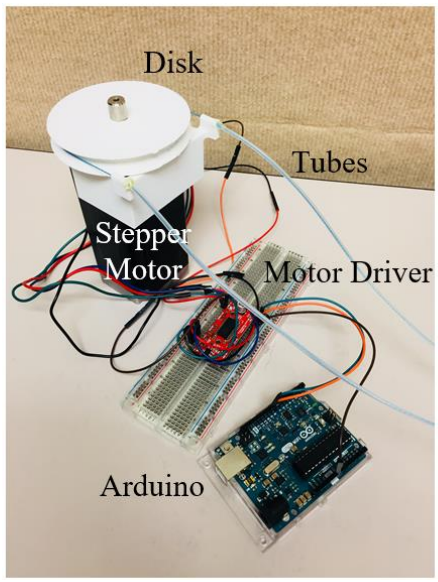



**Fig. 19.** Actuating part

The brushing block moves back and forth on a support axle and soft brush hair was glued on the bottom to produce affective touch [17]. The pressure and velocity are the two key factors for such affective touch. A micro servo (Hitech, HS24BB) shown in Fig. 17 with a lead screw is integrated to drive the brush up and down. By reading the encoder information of the servo, we compute the relative distance between the brush and the skin, which correlates to the brushing pressure. The brush can go 11.5cm and 1cm in horizontal and vertical directions, respectively. The brushing block is able to cover the 3-12 cm/s speed range that is optimal for creating affective touch [12].

The actuating part is positioned 50cm away from the wrap. The tendons (Bravefisherman, 0.48mm) stick on a disk housed on a Nema 23 stepper motor (Stepper-Online, 23HS45) and then go through plastic tubes and into the chamber, actuating the brushing block. To reduce the nonlinear friction between the tubes and the tendons, we try to keep the tubes as straight as possible during the experiments in order to achieve the smoothest brushing. Nevertheless, sufficient smooth motion can be achieved even if the tubes are bent to a circle curve. A controller (Arduino, Uno) controls both the stepper and servo motors. The Arduino connects to a workstation via serial port to deliver the tactile stimulus synchronously with audio-visual stimulus. The tendons are non-extendable so we can control the brushing velocity through the open-loop speed control of the stepper motor that is precise enough for our application.

### *2.6.2. Soft-Brush Experimental Validation*

To test whether Soft-Brush can outperform TSD in terms of tolerability, we ran experiments following the same procedure as MADCAP usability study. Before the experiment, the Soft-Brush was soaked into warm water with baby shampoo so that it is warm and fragrant. The TSD typically needs 30-45 seconds to put on whereas the Soft-Brush only needs 5-10 seconds to do so. Moreover, researchers no longer need to touch the infants' arms during the setup to avoid stranger touching effect. The experimental setup is shown in Fig. 20. The subject sits in a high chair wearing Soft-Brush on the left forearm as well as E4 sensor (Empatica, E4 wristband) on the ankle and an EEG cap (EGI, Geodesic Sensor Net) on the head. After the setup, the subject watches six sessions of videos. Tactile stimulus is applied in three of the six sessions in random orders. Throughout the experiment, EEG, physiological and eye gaze data were recorded. After the experiments, we can temporarily take the supporting axle along with the brushing block off so we can wash and disinfect the silicone wrap conveniently.

### *2.6.3. Soft-Brush Validation Results*

We have conducted 12 experiments to data. Ten infants tolerated the entire process, which amounts to a tolerance rate of 83.33%. Note that the tolerance rate of the TSD was 46.15%. The reasons for the two infants who did not went through the whole process are as follows: One infant got upset during the experiment so we stopped the experiment; the other infant did not tolerate the EEG cap and we went through the rest of the process without EEG recording. Therefore, if we take out cases due to other causes (e.g., EEG tolerability) to compute the acceptance rate of tactile stimulators only, the acceptance rate of the Soft-Brush is 91.67% (11 out of 12). If we apply the same criteria for the TSD, the tolerance rate is 50.00%.

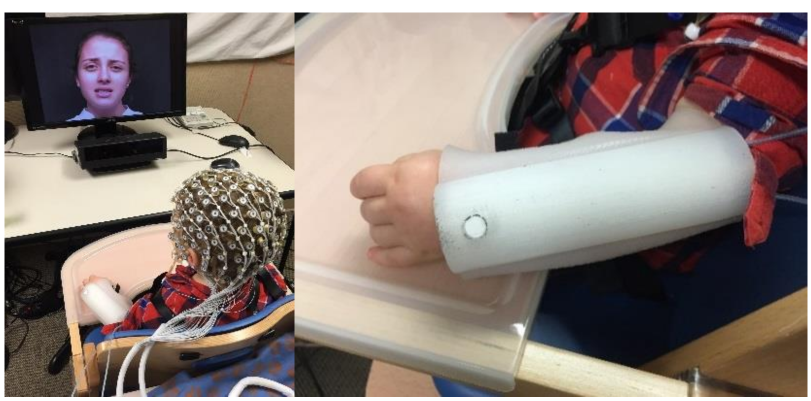


**Fig. 20.** Experimental setup and Soft-Brush during experiment

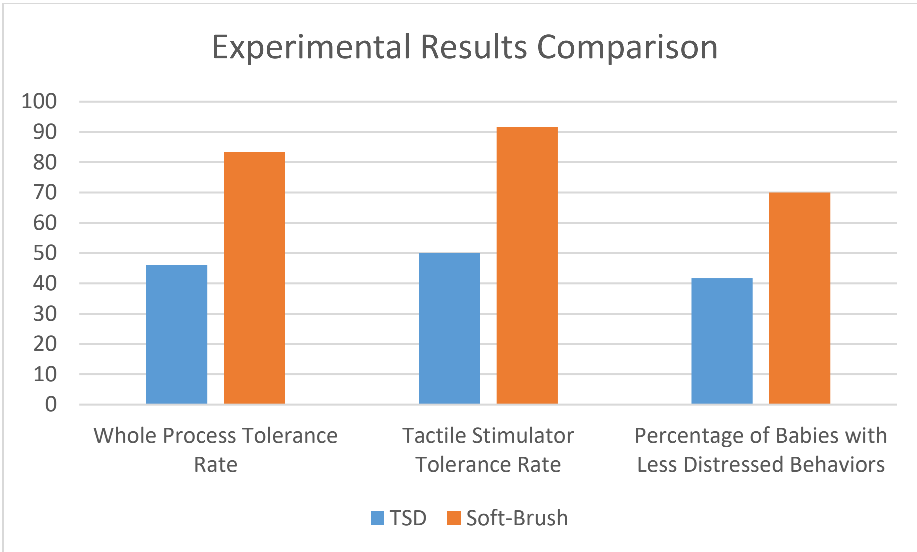



**Fig. 21.** Experimental validation results

Moreover, the Soft-Brush created less distraction and constraint to the infants during the experiments based on our observations. They were able to move their arms freely. The shorter setup time, less stranger touching and more comfortable feeling led to a pleasant mental state in the infants. Of those infants who completed the entire procedure, 7 out of 10 infants (70%) did not display any distressed behaviors (e.g., crying) as compared to 41.67% for the TSD. The calmer and happier state of babies during the experiments leads to a more reliable data collection.

The results indicate that the Soft-Brush is capable of delivering the tactile stimulus with much higher tolerance rate than the TSD. In addition, the infants wearing the Soft-Brush displayed less distress during the experiments. The observed behaviors of the infants during experiments suggest that infants feel much more comfortable wearing the Soft-Brush. In addition, feedback from the parents agrees with this observation.

## 2.7 MADCAP Pilot Study

With the improved tactile stimulator Soft-Brush, we were able to collect data with MADCAP more efficiently and analyze the multimodal sensory trajectories of infants at high risk of ASD. Therefore, we conducted a pilot study on toddlers aging from 2-5 years, in ASD group that has reliable diagnosis of ASD and TD group, respectively. We delivered auditory, visual, and tactile stimuli to the toddlers in both synchronous and asynchronous manner according to the protocol described later. At the same time, we collected physiological responses, including peripheral physiological, EEG, and eye gaze data, and behavioral data via questionnaire from the parent.

*2.7.1. Experiment*

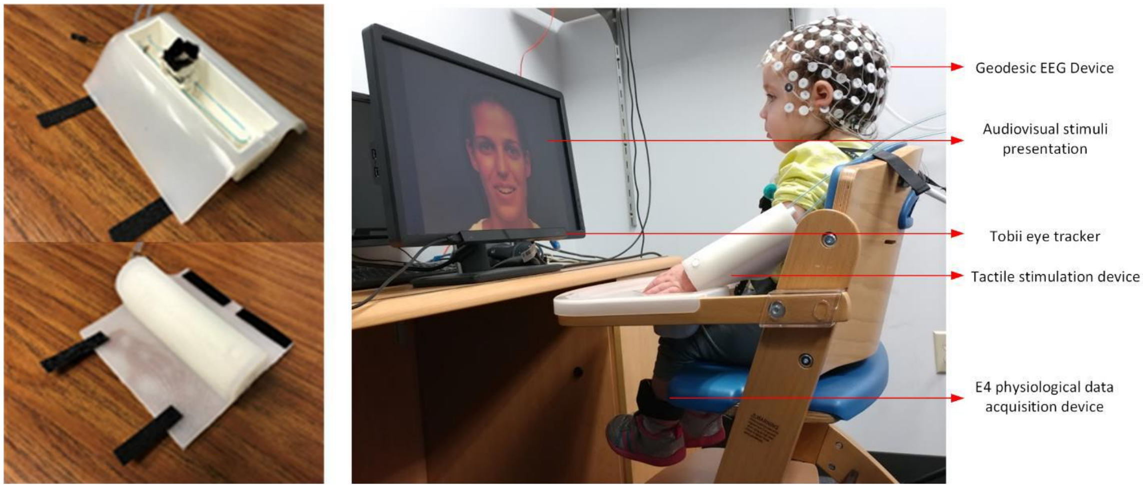



**Fig. 22.** Pilot study experimental setup

We recruited 17 toddlers aged between 2-5 years with ASD diagnosis and 12 age-matched typically developing controls. Each participant had one visit for this study. The final analyses contained data from 10 participants with ASD and 10 TD controls. The details about the system's tolerability is presented later. The protocol was reviewed and approved by the Institutional Review Board (IRB) at Vanderbilt University. After receiving a thorough explanation of the experiment, parents gave written informed consent for their children's participation.

Within a sound-attenuated room, the infant was seated in a high chair or a toddler seat depending on their sizes. The chair was positioned 50cm from the monitor, a video recording device, and an eye tracker. If the toddler could not sit along on the chair, they could sit on the lap of their parent but we asked the parent to have minimal distract on the toddler. The parent also needed to wear an infrared lenses to avoid being tracked by the eye tracker. The Geodesic EEG cap was then placed on the head of the toddler. Then we wore the Soft-Brush on the left forearm of the toddler and secured the Velcro straps. The E4 wristband was attached to the ankle of the participant. Subsequently, the eye tracker calibration was accomplished by using audio-enabled, animated cartoon pictures. Before the stimuli were presented to the participants, a three-minute baseline data was collected for peripheral physiological and EEG while the participant was watching a non-social neutral-content video. The participants then participated in a multi-stimuli presentation session, lasting approximately 8 minutes, wherein they were exposed to three distinct presentations of audiovisual stimuli, each lasted 50 seconds, along with affective touch. In the end, an event-related potential (ERP) session was conducted to explore the ERP when the participant is exposed to tactile stimulus alone. The experimental procedure is shown in Fig. 23.

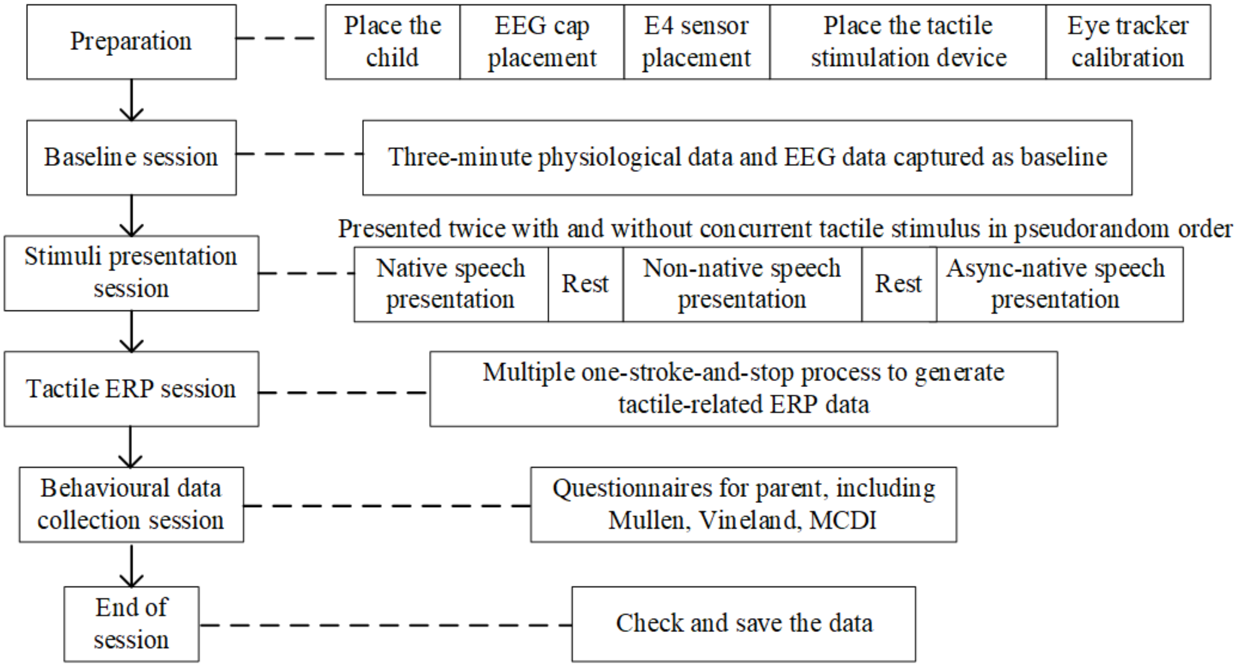



Fig. 23. Pilot study procedure

Throughout the whole sessions, multimodal data were collected with time stamps. Eye gaze, PPG, EDA, 32 EEG signals were collected at sampling rates of 120Hz, 64Hz, 4Hz, and 1000Hz, respectively. Besides the participants' physiological profile data, we also collected behavioral data in the form of questionnaire. They are the Mullen Scales of Early Learning (MSEL) and MacArthur-Bates Communicative Development Inventories (MCDI). MSEL is a standardized test that assesses development in several domains, including expressive and receptive language, for children birth to 68 months. The MCDI is a parent report that assesses early vocabulary and broader spoken language ability.

*2.7.2. Results*

Ten out of 12 participants in the TD group and 10 out of 17 participants in ASD group were used in the final data analysis. Nine participant's data were excluded from the study due to the following reasons: 1) 5 participants did not tolerate the EEG cap; 2) 3 participants could not sit still through the experimental session; and 3) 1 participant's data in TD group was removed because of noticeable developmental delay. The overall compliance rate of toddlers in this study reached 72%.

Different features were extracted from MADCAP data and machine learning approach was applied to classify the diagnostic results with the MADCAP features. The behavioral scores of the participants were also used for comparison. Four sets of machine learning experiments were run to detect the diagnostic results. The details of the experiments and the overall results are shown in the table below.

Table 1: Machine Learning Experiments

| Experiments | Data used for developing model | Accuracies (M, SD) |
| --- | --- | --- |
| 1 | MADCAP features from sessions without affective touch | 0.65 (0.07) |
| 2 | MADCAP features from sessions with affective touch | 0.68 (0.07) |
| 3 | Behavioral scores | 0.80 (0.05) |
| 4 | Behavioral scores and features from second run | 0.83 (0.05) |

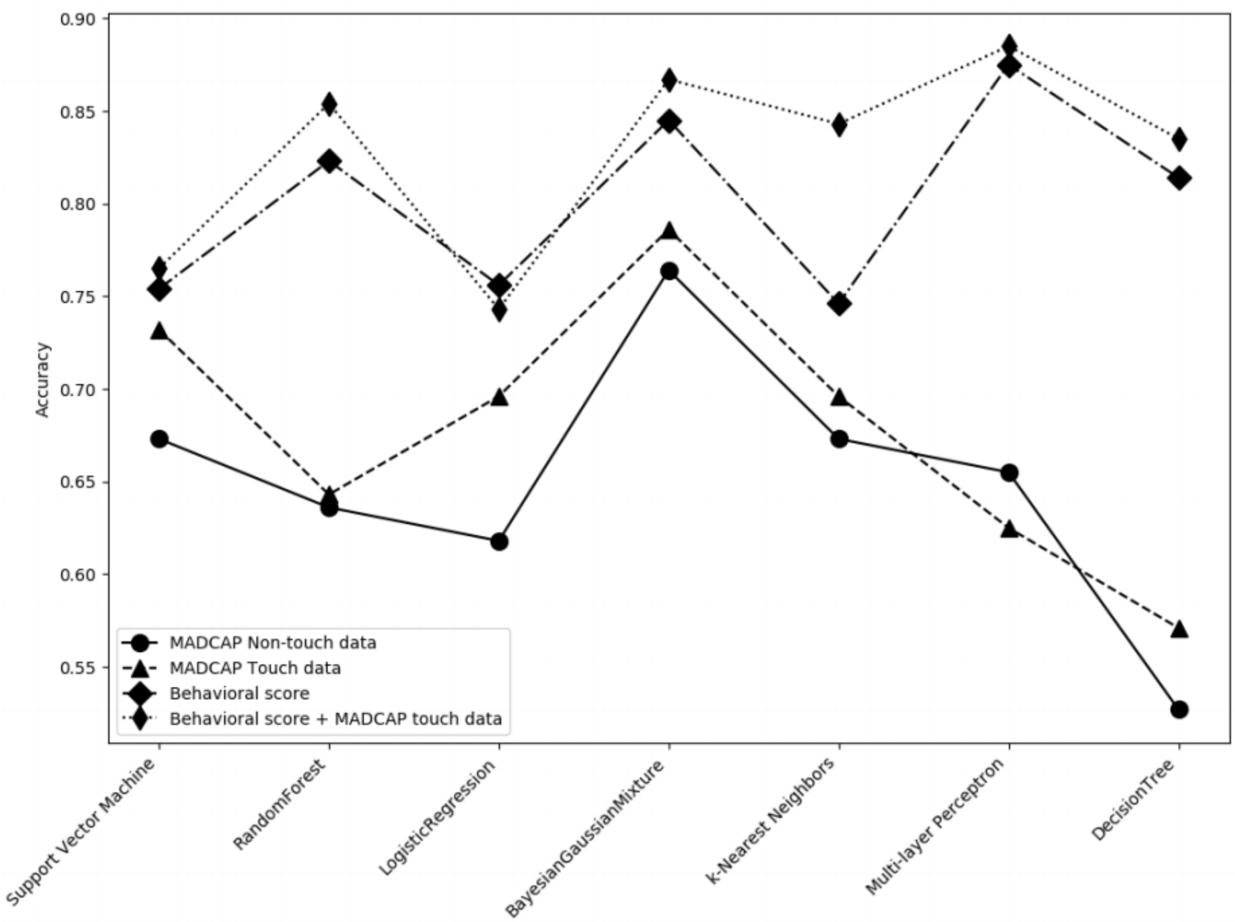



**Fig. 24.** Results of machine learning models

To avoid bias in the training and testing samples, also known as overfitting, all experiments were performed using a leave-one-subject-out approach. In this approach, each time one subject's data was used to test the model, the remaining data were used to train the model. Seven most commonly used machine learning algorithms were developed for each set of experiments. For each algorithm, randomized grid search was used to find the optimal hyper-parameters. Multiple algorithms were used to ensure that potential differences in prediction results between groups resulted from the data rather than the algorithms. The accuracies for all the algorithms in different experiments are shown in Fig. 24.

## 2.8 Conclusion

We investigated the multimodal sensory trajectories of infants at high risk of ASD. We designed a novel multisensory stimulation and data capture system that delivers audio-visual and tactile stimuli while capturing physiological responses of infants. To the best of our knowledge, this is the first multisensory stimulation delivery system capable of providing auditory, visual, and precisely-controlled tactile stimuli in a synchronized and asynchronized manner. In particular, we developed a tactile stimulation device to deliver affective touch with well-controlled speed and pressure. The tactile stimulation device was validated by both a MATLAB Simulink and in real world conditions. The simulation results show that the TSD could produce smooth stroking behavior with consistent brushing pressures, meaning that the affective touch could be effectively triggered. We conducted a usability study for MADCAP and found promising results. The tactile stimulation device could be precisely controlled in terms of speed and pressure, also synchronization with other stimuli. Infants between 3-20 months old fairly tolerated the system and the 10-minute protocol and all the physiological data were robustly collected.

To further improve the tolerance rate in the initial study, we designed a novel tendon-driven soft silicone tactile stimulator, Soft-Brush to improve comfort and portability of MADCAP. This novel mechanism has a soft-sleeve design that wraps around forearms around infants, avoiding stranger contact and significantly decreasing the size and weight of the tactile stimulation device. It is also much more comfortable from our observations and the feedback of parents and can therefore have less distraction on the response data collection. The new design increased the whole session tolerance rate from 46.15% to 83.33%.

With the more efficient generation of tactile stimulator, we were able to conduct a pilot study on 29 children (17 ASD, 12 TD) to collect physiological responses under multisensory stimulus. Machine learning approach was applied on the data collected and it shows fine classifying results with behavioral scores. Four sets of machine learning experiments were run to demonstrate the potential predictive power of these features. First, when affective touch was present, the models resulted in higher predictive accuracies, indicating that adding the new channel of stimulation, affective touch, provided additional information and thus predictive power. This new channel was often overlooked by existing literature. Second, with MADCAP features, models of behavioral scores showed improved prediction accuracies, reaching an average of 83%. This result demonstrated that predictive power of MADCAP data is a potential step forward diagnosing ASD in an earlier age, bringing the current age limit forward. It is worth mentioning that the results presented here do not have an application value yet because toddlers at this age already have reliable diagnosis from clinicians. But our process showed the collection of these biomarkers and demonstrated the feasibility of locating these response patterns. This study paves the road for early detection of ASD with sensory response patterns in infancy, which will better plan intervention for the individuals and their families.

The limitation of this study is the small group size. With more time and support, it would be ideal to include more toddlers to demonstrate more generalized results.

## References


[1] A. V. S. Buescher, Z. Cidav, M. Knapp, and D. S. Mandell, "Costs of autism spectrum disorders in the United Kingdom and the United States," *JAMA Pediatr.*, vol. 168, no. 8, pp. 721–728, 2014, doi: 10.1001/jamapediatrics.2014.210.

[2] J. Baio, "Prevalence of autism spectrum disorders: Autism and developmental disabilities monitoring network, 14 sites, United States, 2008," in *Autism Spectrum Disorders: Guidance, Research and Federal Activity*, 2012.

[3] J. Veenstra-VanderWeele and Z. Warren, "Intervention in the context of development: Pathways toward new treatments," *Neuropsychopharmacology*, vol. 40, no. 1, pp. 225–237, 2015, doi: 10.1038/npp.2014.232.

[4] M. M. Murray, D. J. Lewkowicz, A. Amedi, and M. T. Wallace, "Multisensory Processes: A Balancing Act across the Lifespan," *Trends Neurosci.*, vol. 39, no. 8, pp. 567–579, 2016, doi: 10.1016/j.tins.2016.05.003.

[5] T. Germani *et al.*, "Brief Report : Assessment of Early Sensory Processing in Infants at High-Risk of Autism Spectrum Disorder," pp. 3264–3270, 2014, doi: 10.1007/s10803-014-2175-x.

[6] A. S. Weitlauf *et al.*, *Therapies for Children With Autism Spectrum Disorder: Behavioral Interventions Update*, vol. 318, no. 6. 2017.

[7] W. Jones and A. Klin, "Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism," *Nature*, 2013, doi: 10.1038/nature12715.

[8] L. R. Sacrey, S. E. Bryson, and L. Zwaigenbaum, "Prospective examination of visual attention during play in infants at high-risk for autism spectrum disorder : A longitudinal study from 6 to 36 months of age," *Behav. Brain Res.*, vol. 256, pp. 441–450, 2013, doi: 10.1016/j.bbr.2013.08.028.

[9] F. Shic, S. Macari, and K. Chawarska, "Speech Disturbs Face Scanning in 6-Month-Old," *Biol. Psychiatry*, vol. 75, no. 3, pp. 231–237, 2013, doi: 10.1016/j.biopsych.2013.07.009.

[10] D. J. Lewkowicz, I. Leo, and F. Simion, "Intersensory Perception at Birth : Newborns Match Nonhuman Primate Faces and Voices," vol. 15, no. 1, pp. 46–60, 2010, doi: 10.1111/j.1532-7078.2009.00005.x.

[11] V. Velzen, S. Rigato, J. B. Ali, and A. J. Bremner, "Report The Neural Basis of Somatosensory Remapping Develops in Human Infancy," pp. 1222–1226, 2014, doi: 10.1016/j.cub.2014.04.004.

[12] M. India, S. L. Line, and H. Olausson, "The skin as a social organ," pp. 305–314, 2010, doi: 10.1007/s00221-009-2007-y.

[13] J. Brauer, Y. Xiao, T. Poulain, A. D. Friederici, and A. Schirmer, "Frequency of Maternal Touch Predicts Resting Activity and Connectivity of the Developing Social Brain," no. August, pp. 3544–3552, 2016, doi: 10.1093/cercor/bhw137.

[14] M. T. Fairhurst, L. Löken, and T. Grossmann, "Physiological and Behavioral Responses Reveal 9-Month-Old Infants ' Sensitivity to Pleasant Touch," 2014, doi: 10.1177/0956797614527114.

[15] M. Pelaez-nogueras *et al.*, "The Effects of Systematic Stroking versus Tickling and Poking on Infant Behavior," 1997.

[16] F. Mcglone, J. Wessberg, and O. Hakan, "Perspective Discriminative and Affective Touch : Sensing and Feeling," vol. 4, no. Table 1, 2014, doi: 10.1016/j.neuron.2014.05.001.

[17] I. Croy, H. Geide, M. Paulus, K. Weidner, and H. Olausson, "Affective touch awareness in mental health and disease relates to autistic traits – An explorative neurophysiological investigation," *Psychiatry Res.*, vol. 245, pp. 491–496, 2016, doi: 10.1016/j.psychres.2016.09.011.

[18] I. Gordon, A. C. Voos, R. H. Bennett, D. Z. Bolling, K. A. Pelphrey, and M. D. Kaiser, "Brain mechanisms for processing affective touch," *Hum. Brain Mapp.*, 2013, doi: 10.1002/hbm.21480.

[19] C. Triscoli, H. Olausson, U. Sailer, H. Ignell, and I. Croy, "CT-optimized skin stroking delivered by hand or robot is comparable," vol. 7, no. December, pp. 1–5, 2013, doi: 10.3389/fnbeh.2013.00208.

[20] M. K. Das, N. Bhattacharyya, and A. K. Bhattacharyya, "WHO child growth standards," *European Journal of*

*Pediatrics*. 2010, doi: 10.1007/s00431-009-1041-x.
[21] L. S. Löken, M. Evert, and J. Wessberg, “Pleasantness of touch in human glabrous and hairy skin: Order effects on affective ratings,” *Brain Res.*, 2011, doi: 10.1016/j.brainres.2011.08.011.
[22] D. E. Elder, L. Russell, D. Sheppard, G. L. Purdie, and A. J. Campbell, “Car seat test for preterm infants: Comparison with polysomnography,” *Arch. Dis. Child. Fetal Neonatal Ed.*, 2007, doi: 10.1136/adc.2006.109488.
[23] S. S. Moy *et al.*, “Sociability and preference for social novelty in five inbred strains: An approach to assess autistic-like behavior in mice,” *Genes, Brain Behav.*, 2004, doi: 10.1111/j.1601-1848.2004.00076.x.

# Chapter 3: A Predictive Multimodal Framework to Alert Caregivers of Problem Behaviors for Children with ASD (PreMAC)


## 3.1 Abstract

Autism Spectrum Disorder (ASD) impacts 1 in 54 children in the US. Two thirds of children with ASD display problem behaviors. If a caregiver can predict that a child is likely to engage in problem behavior, they may be able to take action to minimize that risk. Although experts in Applied Behavior Analysis can offer caregivers recognition and remediation strategies, there are limitations to the extent to which human prediction of problem behavior is possible without the assistance of technology. In this chapter, we propose a machine learning-based predictive framework, PreMAC, that uses multimodal signals from precursors of problem behaviors to alert caregivers of impending problem behavior for children with ASD. A multimodal data capture platform, M2P3, was designed to collect multimodal training data for PreMAC. The development of PreMAC integrated a rapid functional analysis, the interview-informed synthesized contingency analysis (IISCA), for collection of training data. A feasibility study with seven 10- to 15-year-old children with ASD was conducted to investigate the tolerability and feasibility of the M2P3 platform and the accuracy of PreMAC. Results indicate that the M2P3 platform was well tolerated by the children and PreMAC could predict precursors of problem behaviors with improved prediction accuracy than the existing studies. The importance of different sensor modalities were also compared. This study utilizes the behavioral precursors, integrates multimodal data and extends the current predictive capacity of problem behaviors in terms of predicting accuracy and time window.


## 3.2 Introduction

Autism spectrum disorder (ASD) is a common neurodevelopmental disability characterized by social communication difficulties and repetitive patterns of interest and behaviors [1]. Current prevalence estimates indicate that one in 54 children in the US are diagnosed with ASD [2], two-thirds of whom display problem behaviors [3]. Although there are multiple terms that could be used for the very diverse behaviors targeted by our system, we utilize the term “problem behavior”, consistent with the Applied Behavior Analytic literature and the developers of the IISCA tool on which our system is based [4][5][6]. Common problem behaviors that co-occur with ASD include self-injury, aggression and elopement [7]. These behaviors severely impede involvement of children in community and educational activities [8] and can put children and their caregivers at risk of potential physical harm [9]. Persistent problem behaviors can prevent children from learning new skills [10], excluding them from school services and community opportunities and aggravating financial burden on caregivers [11].

A validated practice for treating chronic problem behaviors is the Functional Analysis (FA), in which a Board Certified Behavior Analyst (BCBA) systematically manipulates environmental variables suspected to evoke and reinforce problem behaviors and directly observes the behaviors of concern under these controlled conditions in clinical settings in order to individualize treatment protocols that may benefit the child [12]. Although FA can provide an empirical understanding of the variables that impact behavior [13] and has been extensively researched, it is usually resource-intensive, requiring full engagement with a BCBA and other team members. Additionally, while the significant resources invested in a FA may result in the identification of certain antecedent conditions (for instance diverted caregiver attention or reduced access to preferred items) as likely to occasion problem behavior for a child, the FA stops short of building a model for truly predicting problem behavior outside of the clinical context. Disruptive, dangerous and chronic problem behaviors that occur outside of clinical settings and their corollary impact can lead to considerable stress for families, educators and children themselves, on top of the financial burdens of procuring best practice behavioral assessment and intervention services [14], [15].

To address some of these limitations of the FA as most frequently described in the published literature, researcher-clinicians have developed a novel process for FA called the Practical Functional Assessment (PFA). The PFA, which analyzes the occurrence of precursors to problem behavior within the experimental design of the FA instead of the problem behaviors themselves, has been developed in recent years to increase safety, speed and acceptability [16], [17]. The PFA has demonstrated clinical utility when identifying and measuring precursor behaviors, which are observable behaviors - such as changes in body movement, affect, or vocalizations - that reliably precede the onset of problem behaviors. In fact, it has been shown that precursors are functionally directly related to dangerous problem behaviors [18]. Because of this, assessors can use precursors as safe proxies for problem behaviors within the assessment context to reduce the potential for unsafe behavioral escalation. However, despite the advantages of PFA over prior FA procedures, it too stops short of creating or verifying a prediction model for problem behavior, and it relies upon measures of human behavior that can be reliably observed through the human eye [19]. Thus, although PFA does not require that actual problem behaviors occur and as a result, improves safety, it still requires extensive expertise and specific resources. It also continues to require the occurrence and observation of the precursor behaviors themselves rather than other types of signals, such as physiological signals, which may serve as reliable precursors to problem behavior. Furthermore, it is not the specific aim of the PFA to build an experimental model that can provide real-time prediction of problem behavior.

The goal of the current work is to capitalize on the strengths of the PFA to develop a clinically-grounded multi-modal data-driven machine learning (ML)-based problem behavior prediction model, PreMAC, which can be utilized within the community to potentially reduce the need for intensive human data collection. We hypothesize that with the advancement of wearable sensors and affective computing, it is possible to create a ML-based prediction model to accurately predict problem behavior (as well as observable precursors to problem behavior) using real-time sensor data that can provide minute changes in one's internal and external states within a given context.

Affective computing is an emerging field that aims to enable intelligent systems to recognize, infer and interpret human emotions and mental states [20]. There are many successful applications of affective computing to analyze and infer emotions and sentiments using facial expression, body gestures and physiological signals [21]. Primarily, facial expressions are direct cues of human affects and researchers have explored automatic recognition of facial expressions. For example, the Computer Expression Recognition Toolbox (CERT) is a software tool for fully automatic real-time facial expression recognition [22] that can automatically code the intensity of 19 different facial actions from the Facial Action Unit Coding System (FACS) [23] and 6 different prototypical facial expressions. In [24], salient facial patches were extracted from facial landmarks, which were then used to classify emotions ranging from anger, fear, disgust, happiness, sadness and surprise. Body gestures also provide a significant source of features for emotion. In [25], skeletal movements captured by video-based sensor technology were used to identify different human emotions. In [26], a set of postural, kinematic, and geometric features that were extracted from sequences of 3D skeletons were used in a multi-class support vector machine (SVM) classifier to recognize emotions from body gestures. Physiological signals have also been used for affective computing. For example, stress differences among computer users were detected by Galvanic Skin Response (GSR), Blood Volume Pulse (BVP), Pupil Diameter (PD) and Skin Temperature (ST) [27]. Arousal states during virtual tasks were classified by signals including GSR, respiration, BVP and Electroencephalography (EEG) [28]. Although less explored, several prosodic and acoustic features have been used in the literature to detect emotions. Differences between portrayed speech and natural emotional speech were classified by features including pitch, energy, pauses and speaking rate [29].

Affective computing has been successfully applied to inferring emotional and behavioral states of children with ASD based on various sensory data. Peripheral physiological responses such as heart rate (HR) and GSR have been used to predict imminent aggression [30]. The results demonstrated that the individualized and group models were able to predict the onset of aggression one minute before occurrence with good accuracy. With the same dataset, a more recent study[6] utilized support vector machine and it resulted a significantly better prediction accuracies over different prediction window lengths. In [31], skin conductance and respiration were used to build an ensemble of classifiers to differentiate the arousal level and valence in children with ASD. The results suggest the feasibility of objectively discerning affective states in children with ASD using physiological signals. With regard to behavior recognition from body motion, accelerometer data was used in [32] to recognize stereotypical hand flapping and body rocking behaviors, which may occur in some children with ASD. Stereotypical motor movements in ASD were detected using deep learning and resulted in a significant increase in classification performance relative to traditional classification methods [33].

In addition to the work on unimodal systems described above, several studies have shown promise regarding detection of affective and behavioral states of children with ASD using data from multimodal sources. For example, a multi-modal stimulation and data capture system with a soft wearable tactile stimulator was developed to investigate the sensory trajectories of infants at high risk of ASD [34], [35]. Wearable multimodal bio-sensing systems have been developed to capture eye gaze, EEG, GSR and photoplethysmogram (PPG) data [36]. Communication and coordination skills of children with ASD were assessed with multimodal signals, including speech, gestures, and synchronized motion [37].

These and other existing studies demonstrate the potential of affective computing for children with ASD. With the advancement of low-cost robust sensors and computational frameworks it has become possible to create data-driven inference systems that are both accessible and affordable [38]. In general, multimodal systems that integrate several modalities, capture more information and hence increase the accuracy and robustness of machine learning models [39]. With regard to predicting precursors to problem behaviors, it is possible that including multiple modalities involving movements, physiology, social orientations and facial expressions could improve prediction accuracies and robustness. These modalities may directly capture the measurable indicators of emotional states of a child that may lead to problem behaviors such as fidgeting, arm crossing, cursing and grimacing [40].

Given the success of existing data-driven frameworks for affective computing, a recent study attempted to characterize different aggressive behaviors and self-injurious behaviors (SIB) [41]. Movement data along with annotated behaviors were collected to build a machine learning model to predict episodes of SIB. The learnt model could classify different types of SIB at high accuracies. However, note that this research explored prediction of problem behaviors and not their precursors. In contrast to the existing work in the literature, we focus on the prediction of precursors of problem behaviors instead of predicting the problem behaviors themselves, which will provide even more time to intervene and deescalate the situation prior to the occurrence of dangerous outbursts. Furthermore, existing studies that apply unimodal systems to predict problem behaviors may omit useful important information that could increase the accuracy of the machine learning model and the chance to compare the relative importance of different data modalities in predicting the precursors.

This chapter introduces the work of development of PreMAC that aims to predict imminent precursors of problem behaviors using multimodal data and behavioral states. Offering caregivers more time in advance could limit behavioral escalation and prevent dangerous problem behaviors. We present a novel PF- embedded experimental framework to collect training data for this model that seeks to capture expert BCBA’s direct behavior observations as the ground truth. In order to develop the PreMAC, we first created a novel **M**ulti-**M**odal data capture **P**latform for **P**recursors of **P**roblem behaviors, M2P3, for children with ASD. M2P3 combines an off-the-shelf wearable sensor, E4 [42], a Kinect sensor [43], and a customized **W**earable **I**ntelligent **N**on-invasive **G**esture **S**ensor (WINGS). The presented multi-modal platform is seamlessly integrated with a newly developed tablet-based software application, Behavior Data Collection

Integrator (BDCI), to collect data and provide assistance to the assessment team completing a modified PFA. Note that the traditional behavioral assessment modalities rely primarily upon paper-and-pencil recording methods for data entry although there have been a few attempts recently to automate the process [44][45][46]. The customized BDCI help experts record ground truth for PreMAC in a convenient and precise manner that can be easily integrated with the M2P3- and WINGS-generated data.

The rest of this chapter is organized as follows. Section 3.3 presents the overall framework to build PreMAC. Section 3.4 presents the details of the M2P3 platform design including sensor integration, software development and customized sensor design. Section 3.5 introduces the protocol of our feasibility study and pilot data collection. Section 3.6 presents the PreMAC training and prediction results. Finally we conclude the work with a discussion of results and potential future work in Section 3.7.

## 3.3 PreMAC Development Framework

The framework for the development of PreMAC is shown in Fig. 1. We have embedded the Interview Informed Synthesized Contingency Analysis (IISCA) into our data collection for training the prediction model. IISCA is a commonly used type of PFA [47]. In an IISCA assessment, a BCBA methodically manipulates the environment to test caregiver hypotheses around what environmental stimuli serve as antecedents or establishing operations (EOs) to a problem behavior. During the IISCA, the child interacts with the BCBA who systematically evokes problem behavior (or more often the reported precursors to problem behavior) by presenting the EOs identified by the caregiver. The BCBA then provides contingent reinforcement for the problem behavior or precursor to halt behavioral escalation and verify that the child's behavior is functioning to receive the reinforcers indicated by their caregivers. At the same time, M2P3 is deployed for use with the children with ASD for multimodal data collection. Another BCBA observer watches the sessions and uses the BDCI to report their ratings on the behaviors of the child. The multimodal data are then de-noised, synchronized and processed and are used to extract features. The ratings of the observer for the behavioral states are used as the ground truth. The features of the multimodal data are then mapped against the ground truth to train PreMAC to predict the precursors of problem behaviors. Cross validation is then run on the PreMAC to address its accuracy and analysis results.

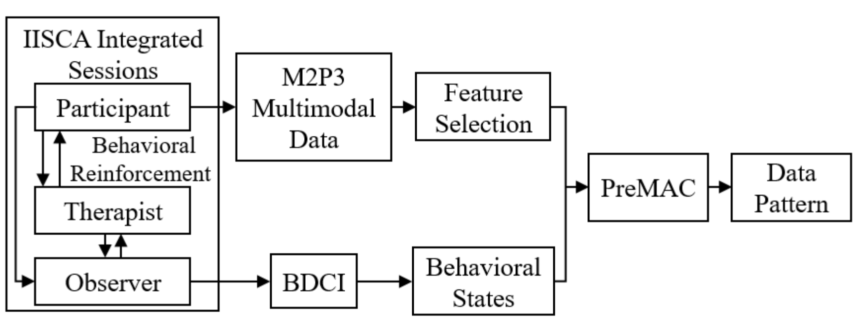



**Fig. 1.** Study Flow Chart

## 3.4 Multimodal Data Collection Platform Design

In order to collect adequate multimodal signals for PreMAC, we developed the M2P3. It integrates and synchronizes multiple data modalities of different time scales. The platform architecture is shown in Fig. 2. The data modalities of M2P3 include facial expressions and head rotations from the Kinect, peripheral physiological and acceleration signals from the E4, and body movements from WINGS. We also developed a tablet application, BDCI, to collect direct behavior observation data.

### *3.4.1. Kinect and E4 Sensors*

M2P3 consists of several platform components. A Microsoft Kinect V2 was used to detect the facial expressions and head rotations of the children. Microsoft Kinect API computes positions of eyes, nose and mouth among different points on the face from its color camera and depth sensor to recognize facial expressions and compute head rotations. We integrated the API to read these measurements in C# scripts. M2P3 is designed to track the first child that enters the camera view of the Kinect. The facial expressions that can be recognized by the API are: happy, eyes closed, mouth open, looking away and engaged. These measures are classified with facial features in real-time and vary on a discrete numerical scale that ranges from 0, 0.5 and 1, meaning no, probably, and yes, respectively. Facial expressions such as happy and engaged are not determinant measures of arousal but have strong indicators of such states [48]. Whether the child is engaged is decided by whether the child opens both eyes and look towards the Kinect. The head rotations are measured in terms of roll, pitch and yaw angles of the head. The sampling rates of the head rotations and facial expressions are both 10Hz and the signals are recorded with time stamps with millisecond precision. The Kinect is placed on the wall by a 3D printed structure which can adjust the pan and tilt angles of the Kinect so that it directly faces the child as shown in Fig. 3.

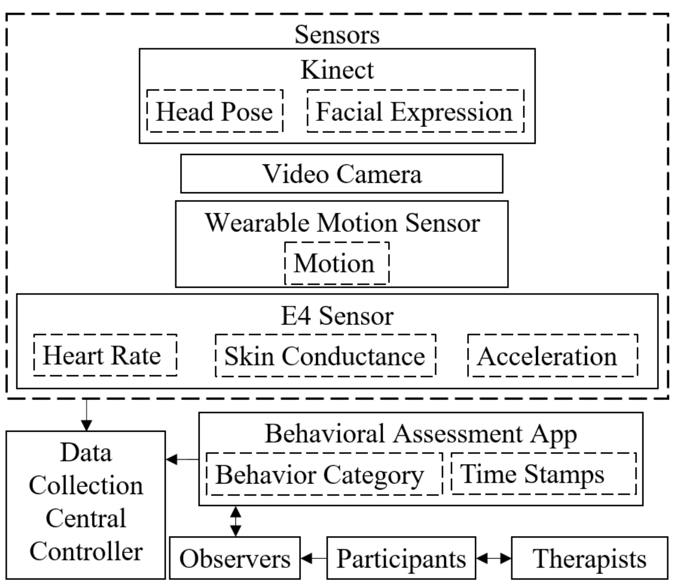



**Fig. 2.** Platform Architecture

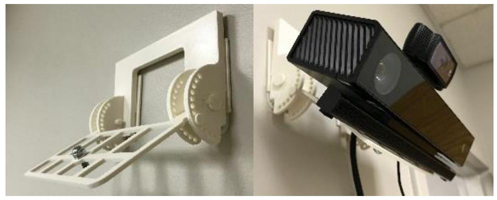
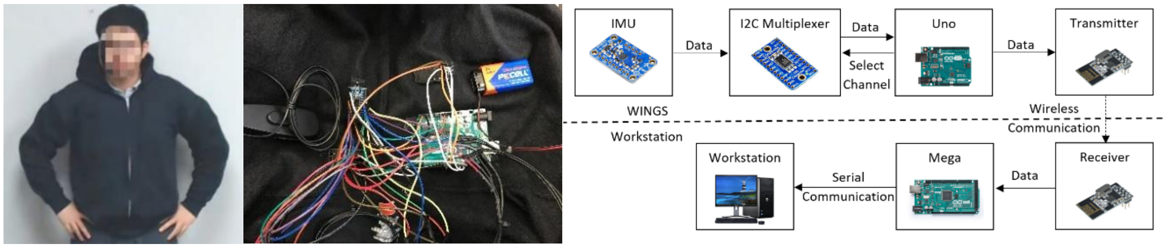


**Fig. 3.** Kinect Setup




(a) (b) (c)

**Fig. 4.** (a) WINGS on a user; (b) Electronics hidden inside WINGS; and (c) WINGS electronic design

Four physiological signals - blood volume pulse (BVP), electrodermal activity (EDA), body temperature and three axis acceleration from an accelerometer - are collected through the E4 wristband. The wristband itself is non-invasive and resembles a smart watch. The sampling rates for BVP and EDA are 64Hz and 4Hz, respectively. We used the API provided for the E4 to record the data with precise time stamps. The real-time physiological data stream is transferred to a central controller by wireless Bluetooth communication.

A central controller is created in Unity, a widely used game engine [49], in C#, to integrate all the data collection modalities. The data collection can be started or stopped by the click of a button. The user interface also displays data being captured by the console and a point cloud showing the field of view of the Kinect.

*3.4.2. WINGS*

The **W**earable **I**ntelligent **N**on-invasive **G**esture **S**ensor or WINGS is a body movement tracking sensor designed for children with ASD. It is a portable, non-invasive tool for measuring upper body motion as shown in Fig. 4(a) and 4(b). There are, in general, two popular ways to track motion and gestures: one is based on computer vision (CV), and the other is based on inertial measurement units (IMU) [50][51]. Despite CV being less-invasive, it has limitations with regard to field of view, occlusion, portability and computational demands [52]. On the other hand, the IMU-based gesture sensor although body worn, could be a better solution in unstructured environment such as in homes and schools where the children will move around. WINGS integrates IMUs to measure the acceleration and orientation of the torso and limbs using a combination of accelerometers and gyroscopes. To increase the likelihood that the platform will be tolerated by children with varying levels of activity, sensory sensitivity, and cognitive functioning, we created WINGS within an off-the-shelf cotton hoodie where the IMUs [53] are sewn within an enclosed space between inner and outer cloth layers. The remaining electronic components including controllers, battery, transmitters and the circuit are sewn within the hood.

Children cannot see or touch any of the electrical and mechanical elements. The total weight of WINGS is 232 grams. When worn, it feels like a normal hoodie. WINGS presents the advantage of allowing children to have an unrestricted workspace. However, we note that some children with ASD will not tolerate wearable sensors and in such cases, WINGS will not be the solution. The total cost of one WINGS is about 170 dollars, although the unit cost will reduce as the production increases. A variety of sizes of WINGS were made to fit children of different sizes.

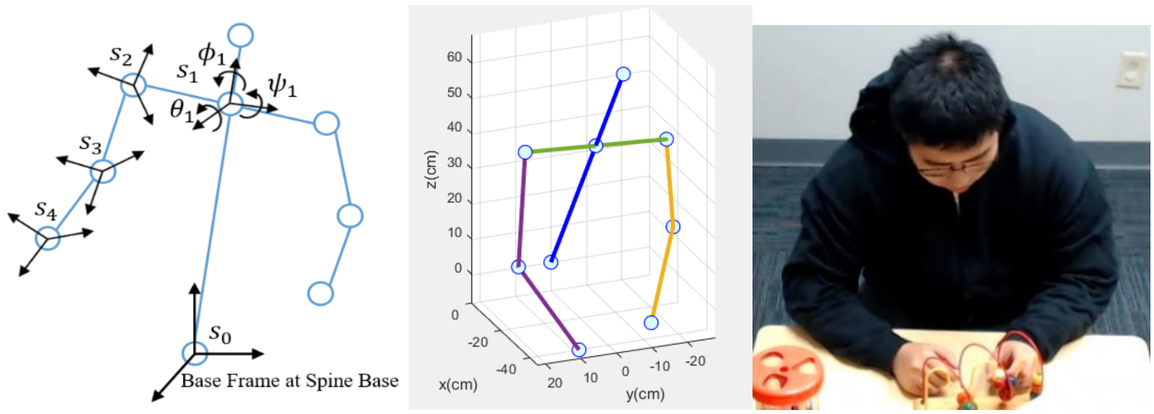
The electronic components of WINGS include an Arduino Uno microcontroller, an I2C multiplexer, a 9V battery, a wireless transmitter and 7 IMUs. Fig. 4(c) shows the data flow scheme of the system. In order to fully construct the upper-body gestures of a child wearing WINGS, we need 7 IMUs to measure joint angles of each forearm, upper arm and the three locations on the back for optimal sensor locations for self-stimulatory behaviors detection [54]. Four cables from each IMU connect to the Uno controller hidden in the hood. Each IMU uses an I2C communication with the Uno microcontroller while the I2C multiplexer [55] searches and loops through the IMUs. The Uno sends the data via a wireless transmitter to a 2.4GHz receiver and the receiver then sends the data further to an Arduino Mega microcontroller. The wireless transmitter and receiver have a SPI communication with the Arduinos. The Mega controller sends the data to a workstation for data storage through a serial communication. When tested, the battery life for WINGS was more than 25 hours, which is adequate for sessions in clinic, school, and other outpatient settings.

From the 3 components of the accelerometer readings, $\mathrm{accl}_x$ , $\mathrm{accl}_y$ and $\mathrm{accl}_z$ , and 3 components of the magnetometer readings, $\mathrm{mag}_x$ , $\mathrm{mag}_y$ and $\mathrm{mag}_z$ , we can compute the roll, pitch and yaw angles ( $\theta$ , $\psi$ , $\phi$ ) of the torso and limbs using Equations (1), (2) and (3) as shown below. The roll and pitch angles are computed by the IMU orientations with respect to the gravitational direction. The yaw angle is computed by the relative IMU orientations with respect to the earth's magnetic field.

$$\theta=\tan^{-1}\left(\frac{\mathrm{accl}_y}{\sqrt{\mathrm{accl}_y^2+\mathrm{accl}_z^2}}\right) \tag{1}$$

$$\psi=\tan^{-1}\left(\frac{\mathrm{accl}_x}{\sqrt{\mathrm{accl}_y^2+\mathrm{accl}_z^2}}\right) \tag{2}$$

$$\phi=\tan^{-1}\left(\frac{\mathrm{mag}_z s\psi-\mathrm{mag}_y c\theta}{\mathrm{mag}_x c\theta+\mathrm{mag}_y s\theta s\psi+\mathrm{mag}_z c\psi s\theta}\right) \tag{3}$$




(a) (b) (c)

**Fig. 5.** (a) Forward Kinematics of WINGS; (b) WINGS Skeleton Visualization; and (c) User Gesture

Knowing the roll, pitch and yaw angles of different joints, we are able to compute the 3D positions and orientations of each joint using forward kinematics [56]. As shown in Fig. 5(a), the base frame is set at the spine base of the child. The base frame's positive directions along the x, y and z axes are front, left of the child and up, respectively. Then the coordinate frame $s_n$ is attached to each body joint. Homogeneous transformation matrices $H^{Jointn-1}_{Jointn}$ between the *n*th joint and the *(n-1)*th joint consist of two parts: a 3-by-3 rotation matrix $R^{n-1}_n$ and a 1-by-3 translation vector $d^{n-1}_0$ . The rotation and translation matrices can align and move the previous coordinate frame to the current coordinate frame, respectively. The rotation matrix is computed by roll, pitch and yaw angles while the translation vector is computed by the body link lengths which are manually measured for different sizes of WINGS. Each homogeneous transformation matrix is computed using Equation (4). The overall homogeneous transformation matrix $H^{Origin}_{Jointn}$ between the base frame and the *n*th frame can be computed by multiplying all the homogeneous transformation matrices as in Equation (5). From this matrix, $d^0_n$ provides the 3D position of the *n*th joint position with respect to the base frame.

$$H^{Jointn-1}_{Jointn}=\begin{bmatrix} R^{n-1}_n & d^{n-1}_0 \\ \vec{0} & 1 \end{bmatrix}=\begin{bmatrix} R_{x,\psi}R_{y,\theta}R_{z,f} & d^{n-1}_0 \\ \vec{0} & 1 \end{bmatrix}=$$
$$\begin{bmatrix} cf_n c\theta_n & cf_n s\theta_n s\psi_n\text{-}sf_n c\psi_n & sf_n s\psi_n+cf_n s\theta_n c\psi_n & x^{n-1}_n \\ sf_n c\theta_n & sf_n s\theta_n s\psi_n+cf_n c\psi_n & sf_n s\theta_n c\psi_n\text{-}cf_n s\psi_n & y^{n-1}_n \\ \text{-}s\theta_n & c\theta_n s\psi_n & c\theta_n c\psi_n & z^{n-1}_n \\ 0 & 0 & 0 & 1 \end{bmatrix} \tag{4}$$

$$H^{Origin}_{Jointn}=H^{Origin}_{Joint1}gH^{Joint1}_{Joint2}LH^{Jointn-1}_{Jointn}$$
$$=\begin{bmatrix} R^0_1 & d^0_1 \\ \vec{0} & 1 \end{bmatrix}L\begin{bmatrix} R^{n-1}_n & d^{n-1}_n \\ \vec{0} & 1 \end{bmatrix}=\begin{bmatrix} R^0_n & d^0_n \\ \vec{0} & 1 \end{bmatrix} \tag{5}$$

Thus, we have the 3D positions of each body joint and we can construct the body gestures made using these joints. A MATLAB program was written to visualize the upper body gestures in real time. Fig. 5(b) shows a visualized gesture and Fig. 5(c) shows its corresponding photo. The lines represent the limbs and the blue dots represent the joints.

The precision of the IMU measured roll, pitch and yaw angles is approximately 1 degree. To quantitatively validate the overall precision of WINGS, we conducted a test where a user wore WINGS and sat in a chair at a designated point. Then the user reached nearby designated 3D points using his shoulder, elbow and wrist. Thus, the relative 3D positions between that joint and the spine base could be measured manually and we compared it to the results computed by WINGS. The user used each joint to reach the designated point for 10 times and the average errors of the shoulder, elbow and wrist were 5.7mm, 9.6mm and 11.7mm, respectively. These precisions are adequate for human gesture measurements for our purpose.

### *3.4.3. Behavioral Data Collection Integrator*

To record under which conditions target behaviors were observed, the IISCA requires observers to record the occurrence of precursors to problem behaviors or problem behaviors themselves, typically using paper and pen while timestamping events via a stopwatch [57]. There have been some attempts recently to automate this process. A computerized behavioral data program "BDataPro" allows real-time data collection of multiple frequency and duration-based behaviors [44]. Catalyst, another software for behavioral assessment, allows collection and management of a wide variety of data for behavioral intervention, including skill acquisition and behavior reduction [45]. An annotation tool for problem behaviors for people with ASD was also developed to log data more conveniently [46]. These existing annotation tools cannot efficiently and precisely record and integrate direct behavioral observation with multimodal data collection. To increase the portability, convenience, and precision of behavioral data collection, we designed a tablet application to assist human therapists with recording data during IISCA procedures, the BDCI. BDCI was written in Unity and implemented on an Android tablet [58].

The application has three pages: Initialization, Session, and Summary. In the Initialization page, there are fields for the observer to input child information, therapist information, as well as session number and type. Once initialized, the observer clicks the start button to begin the session. In the meantime, the application generates a text file to store information and the interface moves to the second page, the Session page. By clicking each button, the application writes a data entry containing the category of the event and its time stamp precise in milliseconds. As shown in Fig. 6, there are several buttons on the Session page related to observer actions and child behaviors. Two buttons are available for the observer to switch between two therapist-imposed conditions within this assessment protocol: establishing operations (EO) and Reinforcing stimulus ($S^R$). Establishing operations represents those antecedent conditions reported to evoke behavioral escalation by caregivers. Reinforcing stimulus ($S^R$) represents those intervals in which antecedent conditions are arranged to prevent, de-escalate, and restore a state of happy and relaxed engagement.

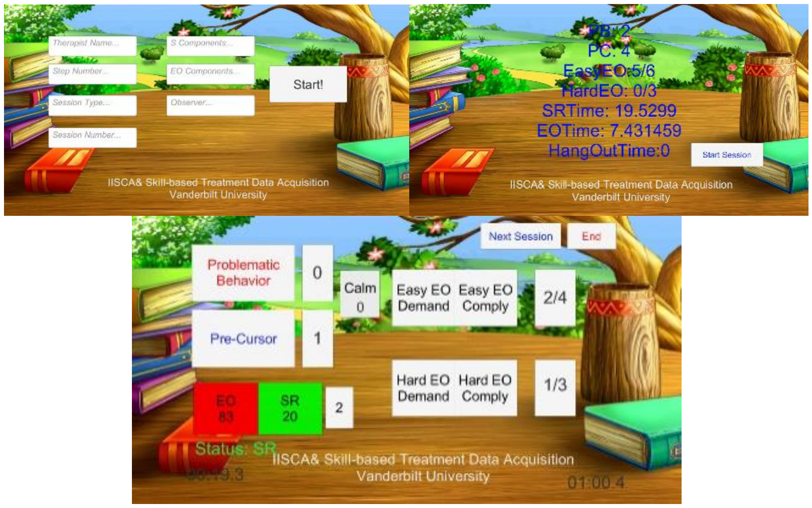



**Fig. 6.** Screenshots of the BDCI Application

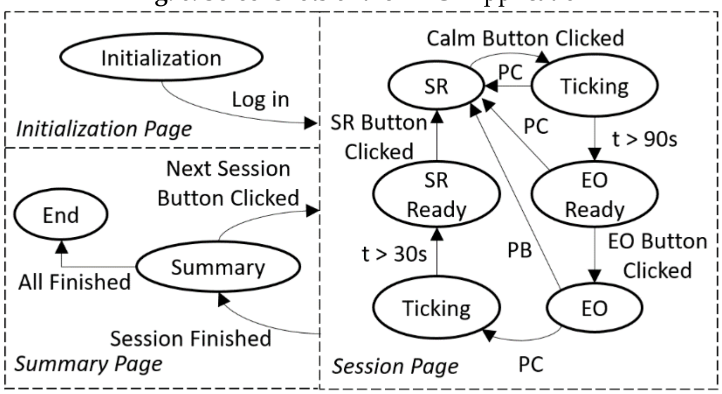



**Fig. 7.** FSM of BDCI

For this app, the current antecedent condition is highlighted in green. The observer can toggle between the two conditions by clicking the relevant button. Event recording of problem and precursor behaviors are recorded by the app. The elapsed time for the current assessment session and condition within the session are shown in the lower part of the screen. According to the IISCA protocol, a specified minimum duration of 90 seconds of the child demonstrating a happy, relaxed, and engaged affect within the $S^R$ condition is needed prior to re-instituting the EO condition. This procedure is used in order to 1) prevent the child from escalating to higher intensities of problem behavior or becoming emotionally dysregulated to such a degree that there is a reduction to their awareness of their environment and 2) ensure that the child contacts reinforcement as a consequence of their problem behavior so that they learn to emit the behavior with higher frequency. This second point may sound counter-intuitive as therapists are teaching the child to engage in high rates of undesirable behavior, but bear in mind that it is the earliest and least disruptive forms of the child's escalation cycle that is being strengthened through this assessment, and it is when high rates of precursor behavior are evoked within the experiment that a robust individualized prediction model for problem behavior can be built. The app includes stopwatches for time management that can cue the data collector and BCBA when a change of condition is appropriate. When an antecedent condition is not ready to be implemented, the button turns red and includes a countdown for the time remaining until the next condition can be implemented.

The app was designed using a finite state machine (FSM) that integrated with our modified IISCA protocol. The app starts with the initialization state. After logging in the session information, the app goes to the $S^R$ state. It was important to the assessment process and our subsequent analyses to include time stamps for when the child was demonstrating a calm affect. At the outset of each experiment, therapists discussed the importance of collecting this information with caregivers and sought their assistance in using their expert knowledge of their child's affective states to ensure that the data collector was accurate in recording periods of observable calm in the participating child. Caregivers observed every minute of every experiment through a one-way mirror, and provided real-time feedback to the data collector as to when the child became or ceased to be calm. The data collector in turn pressed the calm button within the BDCI app and a timer provided feedback to the data collector as to the duration of the current interval of calm. A continuous happy, relaxed, and engaged state lasting at least 90 seconds was sought (by keeping reinforcement in place) to prevent behavioral escalation and give the child's body time to provide "calm" data to the M2P3 and WINGS that could be compared with the data generated when they were escalating behaviorally . If any precursor or problem behaviors happen during this time, the $S^R$ condition must continue. If the child is observed to remain continuously calm, the app indicates a readiness for the EO conditions at the end of 90 calm seconds and the observer will click the EO button as the therapist begins to present the evocative EO conditions. In the EO state, if the precursor button is clicked, the event is recorded, and the app will provide a 90 second count down after which the app indicates readiness for the $S^R$ condition. If a single assessment session is finished, the app proceeds to the summary state; if all the sessions are already finished in the summary state, the app will move to the end state. The FSM is shown in Fig. 7.

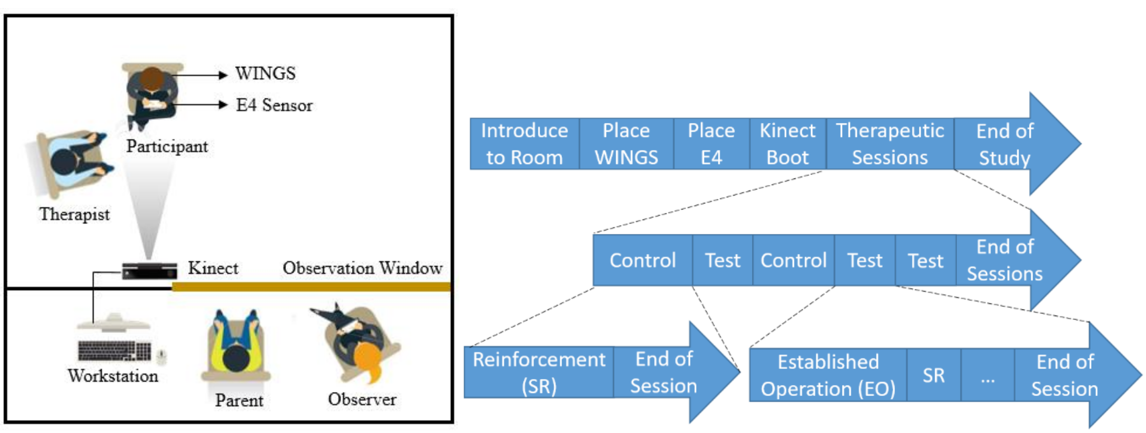



**Fig. 8.** (a) Experimental Setup; and (b) Experimental Procedure

## 3.5 Data Collection Experiment

In order to collect training data for PreMAC and also to demonstrate the feasibility and tolerability of M2P3, we conducted a feasibility study with 7 participants with ASD from 10-15 years old (6 male, 1 female; mean age = 12.20 years, SD = 1.37). These children all had diagnoses of ASD from licensed clinical psychologists. Participants' caregivers reported that the participants presented with frequent episodes of problem behavior which are predictable and significant enough to be provoked by a novel therapist within a novel clinical setting as part of the study protocol. The protocol was reviewed and approved by the Institutional Review Board (IRB) at Vanderbilt University. The research team members explained the purpose, protocols, and any potential risks of the experiment to both the parents and the participants and answered all their questions before seeking informed consent from the parents and informed assents from the participants. Because the purpose of the study was to evoke and respond to precursors to problem behaviors and prevent escalation to dangerous problem behaviors, parents and two dedicated BCBA data collector observed the assessment sessions to ensure that all precursors and problem behaviors as well as the target emotions of happy, relaxed and engaged were correctly recorded. Behavioral states were coded accordingly to clearly-defined written criteria across two observers, as described above. The precursors and problem behavior episodes and the calm states were noted by the observers with the help of observing caregivers and then recorded by the observers using BDCI.

### *3.5.1. Experimental Setup*

As shown in Fig. 8(a), the child-proof room has two compartments, the experimental space and the observation space. The participant sits in the experimental space with a BCBA therapist. The seat for participants is 2 meters away from the Kinect and a video camera. The participant wears an E4 sensor on the non-dominant wrist and WINGS on the upper body. Four observers including an engineer, one of the participants' primary caregivers, a BCBA data collector and a BCBA assessment manager are seated in the observation space, which has a one-way mirror towards the experimental space. The observers and the parent can see the therapist and the participant through a one-way mirror. The therapist had a Bluetooth headphone to relay information from the manager and the manager ensured that the time components of the experimental protocol were correctly executed.

The participant was first invited to the experimental space by the therapist. Then the door was closed to separate the experimental space from the observation room. The therapist then put the E4 sensor on the wrist of the participant and helped him or her wear WINGS. Meanwhile, the parent and the other observers entered the observation room. The Kinect can track up to 16 people at the same time by assigning a specific body ID for each user. In this experiment, the Kinect calibration was performed with only the participant in the Kinect camera view. In this way, the body ID of the participant was recognized so the program only recorded the data of the participant, and not the therapist. Each experiment lasted for approximately one hour.

### *3.5.2. Experimental Procedure*

The experiment followed a modified IISCA protocol [59]. We conducted multiple therapeutic sessions in a single experimental visit to capture data on different behavioral states. These sessions are labeled as control (C) and test (T). The sessions are structured as CTCTT, which represents a A-B-A-B-B Reversal design for single subject research [60]. The control sessions contain only $S^R$ conditions and the test sessions alternate between EO and $S^R$ presentations. EO is followed by $S^R$ and EO is applied once again after at least 90 seconds have elapsed during which the participants stay calm. During EO presentations, the therapist simulated the antecedent conditions that were most likely to evoke precursors and problem behaviors. These tasks were reported by the parents in an open-ended interview days before the actual experimental session. The most commonly reported tasks that induced problem behaviors include asking them to complete homework assignments, removing preferred toys or electronics from them and withdrawing preferred social attention from them. During $S^R$ condition presentations, the therapist offers free access to their favorite toys and electronics, stops asking them to work, removes all the work-related materials, and provides them with the reported preferred attention such as making eye contact, smiling and showing interest. The primary caregiver of the participants observed from behind the one-way mirror, watched the behaviors of the participant, and gave feedback to the data collector and manager who verified the occurrence of precursors or problem behaviors and the presence or absence of a calm state. At times, the caregiver provided advice on how to calm the child or how to provoke problem behavior. The structure of the whole experimental procedure is shown in Fig. 8(b).

### *3.5.3. Feasibility Study Results*

All 7 participants completed their entire experimental sessions. The average length of each experimental session was 54.2 minutes (min = 36.5 minutes, max = 63.1 minutes, SD = 11.5 minutes). The time variation across sessions was largely due to differences in how long it took for each participant to calm down during $S^R$ sessions. WINGS was the most invasive component in the M2P3 platform; 6 out of 7 participants tolerated it without a problem. Some participants even put WINGS on themselves. The only participant who did not tolerate WINGS the entire time put it on at the beginning and then decided to take it off after 15 minutes because he had a high level of caregiver-reported tactile sensitivity.

The other wearable platform component, the E4 wristband, was less invasive and tolerated well by all participants. With regard to staying within the view of the Kinect, one participant was unable to stay seated at the table throughout the entire experiment and instead spent some time on the floor with toys. Thus, the Kinect was not able to track the participant for the entire duration of the experiment.

## 3.6 PreMAC Training

### *3.6.1. Multimodal Data Collection and Signal Processing*

M2P3 collected data using four components: WINGS, E4 wrist band, Kinect, and BDCI. WINGS provided movement data of the upper body; E4 sensor provided peripheral physiological data and the 3-axis acceleration signal of the wrist; Kinect provided facial expressions and head rotations data; and BDCI supplied behavioral states.

Movement data collected by WINGS had a sampling rate of 15 Hz. Less than 0.3% of WINGS data entries were corrupted due to wireless communication or signal noises. A low-pass filter was applied with a cut-off frequency of 10 Hz to raw signals for accelerations that contained sensing noises. The threshold was chosen according to the usual speed of human motions [61] so that the noises were filtered out while keeping information-rich signals for analysis. Peripheral physiological data, BVP and EDA, were collected with the sampling rates of 64 Hz and 4 Hz, respectively. BVP signals were filtered by a 1 Hz to 8 Hz band pass filter. Acceleration signals of the E4 were collected at 32 Hz and a low-pass filter of 10Hz was applied to it.

There was a significant amount of missing data for the head rotations and facial expressions. The Kinect failed to collect these measurements 26.9% of the time when participants were looking down or away. An interpolation algorithm was used for the missing data points, where the numerical mean value of the 20 closest available head rotations and the most frequent class among the 20 closest available facial expressions were chosen respectively, to fill the missing data.

### *3.6.2. Feature Extraction*

From the processed and filtered data, different features were selected and extracted using insight from the BCBA involved in our study as well literature on problem behaviors. Head banging is common amongst children with ASD who exhibit problem behaviors [62]. Measures of head rotations can indicate such physical movements and roll, pitch and yaw angles of the head were thus chosen as features. Children often show observable precursor behaviors that indicate that problem behavior is likely to follow. Examples of these observable precursors include changes in facial expression, such as pouting, eyebrow knitting and intensely staring. From facial expressions, we extracted features of closing of eyes and mouths, engagement, and looking away. Pitch, roll, and yaw angles of the torso, and limbs were also selected as features as they were used to construct the upper body gestures using Equations 1-3. Certain gestures are related to problem behaviors [63] including fist throwing, upper body swinging, laying back and repetitive arm movements. Besides gestures, the average magnitude of accelerations of each joint can also be used as a measure of activity level as shown in Equation (4). The activity level indicates the intensity of physical movements of the children. Fast and repetitive movements such as throwing fist and fidgeting are a common category of problem behaviors [64] and these movements have increased activity level.

$$AL=\frac{\sum_{i=1}^{n}\sqrt{a_{x_i}^2+a_{y_i}^2+a_{z_i}^2}}{7} \tag{4}$$

Peripheral physiological data is a strong indicator of arousal and several features including heart rate (HR) and EDA have been shown to have correlations with problem behaviors [30][31]. Thus, the HR level was computed by inter-beat-intervals (IBI) of the BVP raw signal. From the EDA data, two types of data were separated and characterized, which were tonic skin conductance level (SCL) and phasic skin conductance response (SCR) [65]. These features correlate well with the arousal level of a person [66].

Features from different modalities were combined for training and testing of PreMAC. There were altogether 27 features: 15 roll, pitch, and yaw from forearms, upper arms, and torso from WINGS; 1 heart rate and 2 skin conductance features from E4; 3 accelerations from E4, and 6 facial expressions from Kinect. These features were generated at different instants due to varying processing times of different signal modalities. In this work, in order to combine all these features for each time step, we used WINGS' features at any given instant as the basis and added the other features mentioned above that were closest in time with the WINGS' features.

The BDCI provides the time stamps of precursors of problem behaviors captured by the BCBA observers. With these time stamps, we assigned either absence or presence of imminent precursor classes to each *1-by-27* vector of multimodal data. With the insight from our IISCA practitioners, we chose the most representative data for precursors of problem behaviors within a time window. The time window is between 90 seconds before the episode and the point of precursor generation. Thus, if the multimodal data set was collected within 90 seconds prior to the precursor, it would be assigned label 1. Otherwise the class was assigned label 0. The two classes 0 and 1 had an average ratio of 6:4. This was not a significantly unbalanced dataset so class balancing was not necessary to maintain more information. For our experiment, each child had an average of 27242 samples of the multimodal data.

*3.6.3. Machine Learning*

PreMAC includes both individualized models and group models, predicting whether a precursor to problem behaviors is going to happen in a window of time. The individualized models were built with data from each participant. The group models were built with data from all the participants to explore the general group behavioral patterns. In order to find the most accurate ML algorithm, we explored several standard ML algorithms with our datasets. The library scikit-learn [67] was used on Jupyter Notebook [68]. The samples were randomly divided into training and test sets with a ratio of 80 to 20. Then a 5-fold cross validation was run to compute the accuracies of each algorithm. The prediction accuracies of individualized models of several algorithms are shown in the table below.

Table 1. Comparison of Machine Learning Algorithms

| *Machine Learning Algorithm* | *Individualized Model* | *Group Model* |
|---|---|---|
| *Random Forest* | 98.51% | 97.72% |
| *Support Vector Machine* | 88.71% | 88.00% |
| *k Nearest Neighbors* | 94.94% | 97.06% |
| *Decision Tree* | 94.76% | 96.48% |
| *Discriminant Analysis* | 74.55% | 69.54% |
| *Naïve Bayes* | 68.69% | 75.27% |
| *Neural Network* | 91.72% | 92.07% |

Table 1 shows the accuracies of different algorithms and the individualized model accuracies are the average accuracies among all the participants. For individualized models, Random Forest (RF), k Nearest Neighbors (kNN), Decision Tree (DT) and Neural Network (NN) have high prediction accuracies while Support Vector Machine (SVM), Discriminant Analysis (DA) and Naïve Bayes (NB) have comparatively lower accuracies. For group models, RF, kNN, and DT also show high accuracies while DA has significantly lower accuracy. DT, kNN, NB and NN have higher prediction accuracies in group models as compared to their corresponding individual models but the RF classifier remains the most accurate one. Group model has the combined samples from all participants and the increased sample size could help make more accurate prediction. However, since problem behaviors of each child can vary significantly, group model may try to predict average problem behaviors. This trade-off may be the reason that some models have high accuracies for group models while others do not. The RF individualized model has the best average prediction accuracy. The confusion matrix of an example individualized model is shown in table 2. The confusion matrix of the group model is shown in table 3.

Table 2. Individual Model Confusion Matrix

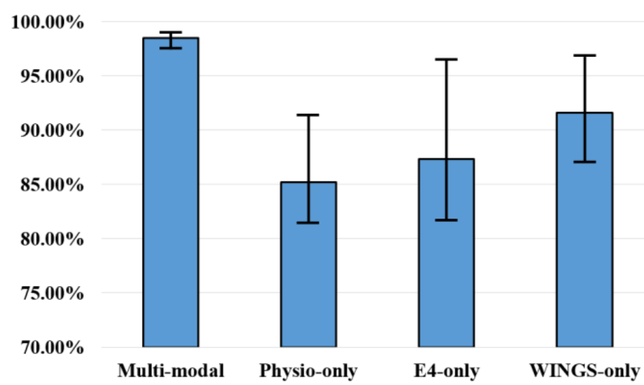
| **n=6004** | **Predicted Yes** | **Predicted No** |
|---|---|---|
| **Actual Yes** | TP = 3548 | FP = 42 |
| **Actual No** | FN = 49 | TN = 2365 |

Table 3. Group Model Confusion Matrix

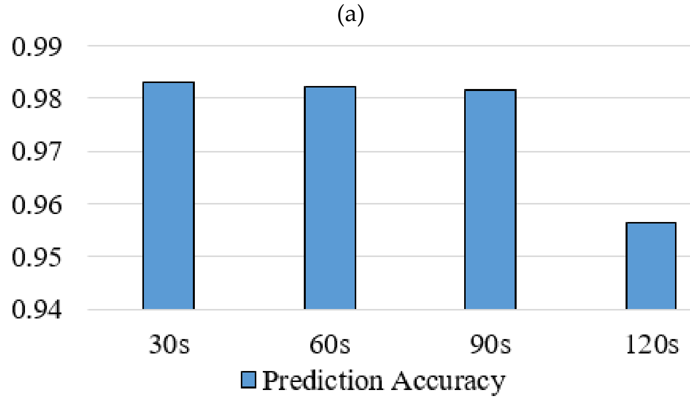
| **n=38139** | **Predicted Yes** | **Predicted No** |
|---|---|---|
| **Actual Yes** | TP = 21850 | FP = 130 |
| **Actual No** | FN = 336 | TN = 15823 |

The RF algorithm also offers estimates of importance of each feature. The feature importance is computed as the total decrease in node impurity weighted by the probability of reaching that node. We also analyzed the relative importance of motion-based, physiology-based, and facial expression-based features. The features included were motion signals of each joint, physiological signals, head rotations and facial expressions. As shown in Fig. 9(c), the most important features were from torso, right shoulder and right wrist, where the total of all features equals to 1. The results are consistent with our experimental observations. The main precursors included participants banging their arms against their torsos and moving their right arms, as the right arm was the dominant arm for the participants. For data modalities from E4 sensor, the physiological features including both HR and EDA had an importance of 0.0566 and the 3-axis accelerations had an importance of 0.0689. Head Rotations had an importance of 0.0399 and facial expression features were the least important with a value of 0.0061. The poor performance of facial expressions may be due to the missing portion of data when children were looking down and away.

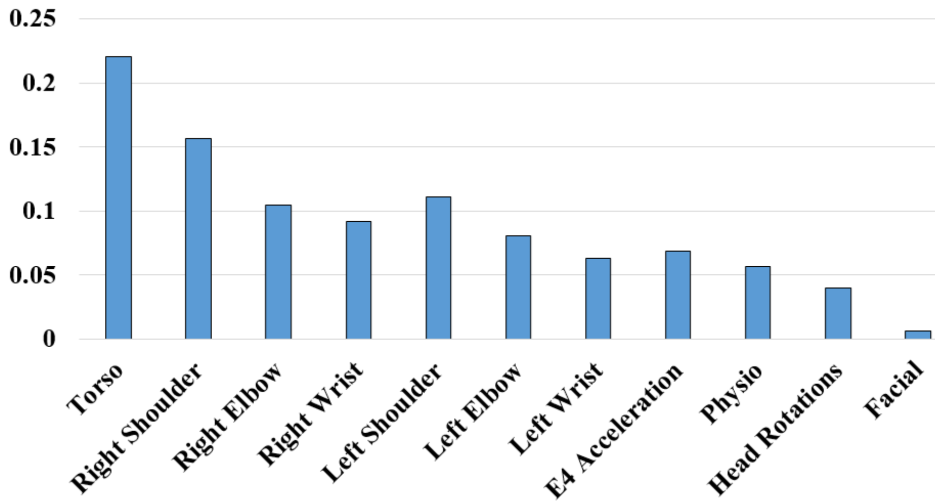
We also analyzed the prediction accuracies with only WINGS data, only physiological data and only E4 data to compare the contributions of different data modalities. We utilized the best performing algorithm RF to learn the data pattern of each child. The average prediction accuracies and the range of each model are shown in Fig. 9(a). The average prediction accuracies for the multimodal model, physiological data only model, WINGS data only model and E4 data only model were 98.51%, 85.22%, 87.31% and 91.63%, respectively. Multimodal individualized models have high accuracies with small variances, meaning the performance is robust among children. The customized WINGS provides data for significantly improved prediction accuracies compared to the commercial E4 sensor.




(a)




(b)




(c)

**Fig. 9.** (a) Prediction Accuracies of Different Data Modality Models; (b) Interval Analysis; and (c) Feature Importance Comparison

As mentioned earlier, label 1 (imminent precursor) was assigned to data that was collected within the last 90 seconds prior to the observation. To analyze the effect of the time window on the prediction of precursors, we varied the time window for the class of imminent precursors from 30 seconds prior to the observed precursor to 120 seconds in steps of 30 seconds. To avoid effects of different ratio of classes, we oversampled the minority class so that the two classes had a 1:1 ratio. As shown in Fig. 9(b), prediction accuracies for 30, 60, 90 seconds do not have significant differences but it significantly decreases for the 120 second time window. It is noteworthy that this reliable 90-second prediction is for the precursor that reportedly precedes actual problem behavior by an unknown number of seconds or minutes. A 90-second warning prior to even precursors happening provides enough time for a caregiver to withdraw a demand, provide desired attention, or redirect the child to a preferred item or activity. This analysis validates that the 90 seconds window seems to be the optimal window for precursor prediction that has good prediction accuracy with ample time for the caregivers to intervene.

### 3.7 Conclusion

Existing modes of capturing the behavioral assessment data, which inform intervention planning and efficacy for children with ASD and severe problem behavior, rely upon paper-and-pencil data taken by human observers, and are unable to provide a valid real-time prediction model. To augment this best-practice clinical care model, which underpins many evidence-based intervention approaches for individuals with ASD and other developmental disabilities [16], we developed a novel machine learning based predictive model, PreMAC. Based on multimodal data input, PreMAC creates individualized and group profiles of imminent behavioral escalation among children with ASD based upon physiological, gestural, and motion-based *precursors* that a problem behavior is about to occur. This multi-modal data capture platform, M2P3, collects training data from two portable wearable devices (including one of our own creation, WINGS) and a newly designed tablet application, BDCI, all of which represent low cost options for future real-world community deployment.

PreMAC integrates important relevant data that cannot be reliably collected by a human observer. Specifically, it collects data regarding only subtly visible (e.g. joint angle) or utterly invisible (e.g. skin conductance) precursors of problem behavior at a high level of accuracy. The emphasis of our system design on precursors rather than problem behaviors themselves holds the potential to increase the safety of participants as well as their caregivers by minimizing the risk of a severe problem behavior actually occurring. In summary, this system rapidly generates a robust prediction model with ample time to be clinically and practically relevant all with little-to-no dangerous behaviors occurring at any time during the assessment. If integrated within a system that could somehow signal an adult, this would give caregivers and potentially people with ASD themselves more lead time in intervening on and mitigating problem behaviors leading to better safety, and could even be leveraged to improve treatment approaches designed to address problem behavior.

Our innovative data collection process is novel in its integration of multimodal data collection with cutting edge functional assessment technology from the field of Applied Behavior Analysis. Each step of this work was informed by stakeholder feedback which was then integrated into the system design. Importantly, particularly when designing a system intended for future real-world clinical use, results of this feasibility study suggest that children with ASD with problem behaviors tolerated both the platform and experimental protocol well. Any instances of data loss, such as the participant that sat out of view of the Kinect or the participant who found WINGS uncomfortable, provided us with important information that will guide future adaptation and deployment of updated systems. All modalities of data were robustly collected and synthesized.

To our knowledge, PreMAC extends existing sensory modalities of problem behavior prediction with upper body motion and social orientations and it is the first machine learning model to integrate an IISCA to evoke precursors of problem behaviors instead of dangerous episodes of problem behavior. Within our controlled laboratory context, PreMAC offered a significant increase in prediction accuracy, an average of 98.51% for individualized profiles, as compared to the existing published results which predicted behaviors themselves rather than precursors [6][32][41]. Potential reasons for higher prediction accuracy of PreMAC include more sensing modalities, more accurate precursor time stamps through BDCI, and large data sample size of each child. It is also worth mentioning that this work is predicting behavioral precursors of problem behaviors that precede problem behavior episodes, demonstrating great potential to offer more time in advance for caregivers to intervene.

Based on our analysis, body motion is the most predictive sensing modality for imminent precursors of problem behaviors and WINGS alone may provide adequate information to predict imminent precursors.

We further investigated the importance of different limb movements, head rotations, physiological data and facial expressions. The torso movement is the most effective feature and movement on the dominant side is more effective than the other side. Physiological data is comparatively much less effective than body movements and facial expressions almost do not contribute to the prediction accuracies. This paves the way for future work to identify the most efficient sensor to integrate into an online platform for home and school settings.

Several limitations exist that warrant attention in future work. First, the Kinect and the video camera are the two non-portable components in the platform that, at present, impede data collection in an out-of-lab setting. In the future, we will continue working towards a totally portable data collection platform for home and school settings, which will better assess behaviors of children with ASD and usual problem behaviors. WINGS have combined upper body motion detection with the softest clothing most typically worn by children in this age group and further testing will include more stakeholder input including questions about possible improvements to increase maximum comfort level across a broad range of sensory profiles. Because this is not an autism-specific system, but rather one designed for any child with problem behaviors, updated phenotypic information was not obtained for the purpose of this small pilot study. In future work, we will obtain measures of autism severity, problem behavior frequency, and cognitive skills using standardized tools to better understand the likely variability that will present across a larger sample of individuals. In spite of these limitations, the proposed platform collects multi-modal data with wearable sensors including customized WINGS, a novel tablet application gathering precise time stamps for function analysis, and an IISCA protocol to generate high-density precursors with very few actual problem behaviors. The platform was validated on 7 children with ASD and the performance of PreMAC was promising.

## References


[1] A. S. Weitlauf *et al.*, *Therapies for Children With Autism Spectrum Disorder: Behavioral Interventions Update*, vol. 318, no. 6. 2017.

[2] M. J. Maenner *et al.*, "Prevalence of autism spectrum disorder among children aged 8 Years-Autism and developmental disabilities monitoring network, 11 Sites, United States, 2016," *MMWR Surveill. Summ.*, vol. 69, no. 4, pp. 1–12, 2020, doi: 10.15585/MMWR.SS6904A1.

[3] J. L. Matson and M. Nebel-Schwalm, "Assessing challenging behaviors in children with autism spectrum disorders: A review," *Res. Dev. Disabil.*, vol. 28, no. 6, pp. 567–579, 2007, doi: 10.1016/j.ridd.2006.08.001.

[4] W. E. Sullivan, V. Saini, N. M. DeRosa, A. R. Craig, J. E. Ringdahl, and H. S. Roane, "Measurement of nontargeted problem behavior during investigations of resurgence," *J. Appl. Behav. Anal.*, 2020, doi: 10.1002/jaba.589.

[5] H. C. Gover, T. A. Fahmie, and C. A. McKeown, "A review of environmental enrichment as treatment for problem behavior maintained by automatic reinforcement," *J. Appl. Behav. Anal.*, 2019, doi: 10.1002/jaba.508.

[6] T. Imbiriba, D. C. Cumpanasoiu, J. Heathers, S. Ioannidis, D. Erdomuş, and M. S. Goodwin, "Biosensor prediction of aggression in youth with autism using kernel-based methods," *ACM Int. Conf. Proceeding Ser.*, pp. 88–93, 2020, doi: 10.1145/3389189.3389199.

[7] S. Jeff, M. Arthur, and M. O'Reilly, *Challenging behavior and developmental disability.pdf*. 2003.

[8] S. M. Myers *et al.*, "Management of children with autism spectrum disorders," *Pediatrics*, vol. 120, no. 5, pp.

1162–1182, 2007, doi: 10.1542/peds.2007-2362.
[9] O. Chadwick, N. Piroth, J. Walker, S. Bernard, and E. Taylor, "Factors affecting the risk of behaviour problems in children with severe intellectual disability," *J. Intellect. Disabil. Res.*, vol. 44, no. 2, pp. 108–123, 2000, doi: 10.1046/j.1365-2788.2000.00255.x.
[10] E. Emerson, "Challenging behaviour: Analysis and intervention in people with severe intellectual disabilities," 2001.
[11] S. L. Parish, K. C. Thomas, R. Rose, M. Kilany, and P. T. Shattuck, "State medicaid spending and financial burden of families raising children with autism," *Intellect. Dev. Disabil.*, vol. 50, no. 6, pp. 441–451, 2012, doi: 10.1352/1934-9556-50.06.441.
[12] M. D. Shriver, C. M. Anderson, and B. Proctor, "Evaluating the validity of functional behavior assessment," *School Psych. Rev.*, vol. 30, no. 2, pp. 180–192, 2001, doi: 10.1080/02796015.2001.12086108.
[13] O. Healy, S. Lydon, J. Holloway, and M. Dwyer, "Behavioral Interventions for Aggressive Behaviors in Autism Spectrum Disorders," *Comprehensive Guide to Autism*. pp. 461–486, 2014, doi: 10.1007/978-1-4614-4788-7_21.
[14] M. Ghaemmaghami, G. P. Hanley, J. Jessel, and R. Landa, "Shaping complex functional communication responses," *J. Appl. Behav. Anal.*, vol. 51, no. 3, pp. 502–520, 2018, doi: 10.1002/jaba.468.
[15] J. Jessel, G. P. Hanley, M. Ghaemmaghami, and R. Metras, "An evaluation of the single-session interview-informed synthesized contingency analysis," *Behav. Interv.*, vol. 34, no. 1, pp. 62–78, 2019, doi: 10.1002/bin.1650.
[16] G. P. Hanley, C. S. Jin, N. R. Vanselow, and L. A. Hanratty, "Producing meaningful improvements in problem behavior of children with autism via synthesized analyses and treatments," *J. Appl. Behav. Anal.*, vol. 47, no. 1, pp. 16–36, 2014, doi: 10.1002/jaba.106.
[17] C. A. Warner *et al.*, "Toward accurate inferences of response class membership," *J. Appl. Behav. Anal.*, vol. 53, no. 1, pp. 331–354, 2020, doi: 10.1002/jaba.598.
[18] S. Gehlert and T. S. Ward, "Theories of health behavior," in *Handbook of health social work*, 2019, pp. 143–163.
[19] J. Jessel, G. P. Hanley, and M. Ghaemmaghami, "On the Standardization of the Functional Analysis," *Behav. Anal. Pract.*, vol. 13, no. 1, pp. 205–216, 2020, doi: 10.1007/s40617-019-00366-1.
[20] S. Poria, E. Cambria, R. Bajpai, and A. Hussain, "A review of affective computing: From unimodal analysis to multimodal fusion," *Inf. Fusion*, vol. 37, pp. 98–125, 2017, doi: 10.1016/j.inffus.2017.02.003.
[21] E. Cambria, "Affective Computing and Sentiment Analysis," *IEEE Intell. Syst.*, vol. 31, no. 2, pp. 102–107, 2016, doi: 10.1109/MIS.2016.31.
[22] G. Littlewort *et al.*, "The computer expression recognition toolbox (CERT)," *2011 IEEE Int. Conf. Autom. Face Gesture Recognit. Work. FG 2011*, pp. 298–305, 2011, doi: 10.1109/FG.2011.5771414.
[23] J. F. Cohn and E. Paul, "Measuring Facial Action.pdf," 2005, [Online]. Available: https://psycnet.apa.org/record/2006-03862-002.
[24] S. L. Happy and A. Routray, "Automatic Facial Expression Recognition Using Features of Salient Facial Patches," *IEEE Trans. Affect. Comput.*, vol. 6, no. 1, pp. 1–12, 2015.
[25] A. Kapur, A. Kapur, N. Virji-babul, and G. Tzanetakis, "Gesture-Based Affective Computing on Motion," *1st Int. Conf. Affect. Comput. Intell. Interact. (ACII 2005)*, pp. 1–7, 2005.
[26] S. Piana, A. Staglianò, F. Odone, A. Verri, and A. Camurri, "Real-time Automatic Emotion Recognition from Body Gestures," 2014, [Online]. Available: http://arxiv.org/abs/1402.5047.
[27] D. Wu *et al.*, "Optimal arousal identification and classification for affective computing using physiological signals: Virtual reality stroop task," *IEEE Trans. Affect. Comput.*, vol. 1, no. 2, pp. 109–118, 2010, doi: 10.1109/T-AFFC.2010.12.
[28] J. Zhai and A. Barreto, "Stress detection in computer users through non-invasive monitoring of physiological signals," *Biomed. Sci. Instrum.*, vol. 42, pp. 495–500, 2006.
[29] D. Morrison, R. Wang, and L. C. De Silva, "Ensemble methods for spoken emotion recognition in call-centres," *Speech Commun.*, vol. 49, no. 2, pp. 98–112, 2007, doi: 10.1016/j.specom.2006.11.004.
[30] M. S. Goodwin *et al.*, "Predicting imminent aggression onset in minimally-verbal youth with autism spectrum disorder using preceding physiological signals," *ACM Int. Conf. Proceeding Ser.*, pp. 201–207, 2018, doi: 10.1145/3240925.3240980.
[31] S. Sarabadani, L. C. Schudlo, A. A. Samadani, and A. Kushki, "Physiological Detection of Affective States in Children with Autism Spectrum Disorder," *IEEE Trans. Affect. Comput.*, pp. 1–14, 2018, doi: 10.1109/TAFFC.2018.2820049.

[32] F. Albinali, M. S. Goodwin, and S. Intille, “Detecting stereotypical motor movements in the classroom using accelerometry and pattern recognition algorithms,” *Pervasive Mob. Comput.*, vol. 8, no. 1, pp. 103–114, 2012, doi: 10.1016/j.pmcj.2011.04.006.
[33] N. M. Rad and C. Furlanello, “Applying Deep Learning to Stereotypical Motor Movement Detection in Autism Spectrum Disorders,” *IEEE Int. Conf. Data Min. Work. ICDMW*, vol. 0, pp. 1235–1242, 2016, doi: 10.1109/ICDMW.2016.0178.
[34] D. Bian *et al.*, “A Novel Multisensory Stimulation and Data Capture System (MADCAP) for Investigating Sensory Trajectories in Infancy,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 8, pp. 1526–1534, 2018, doi: 10.1109/TNSRE.2018.2854672.
[35] Z. K. Zheng, D. Bian, A. Swanson, and A. Weitlauf, “Soft-Brush : A Novel Tendon Driven Tactile Stimulator for Affective Touch in Children with Autism,” vol. 1, pp. 15–22, doi: 10.1007/978-3-319-94373-2.
[36] Siddharth, A. N. Patel, T. P. Jung, and T. J. Sejnowski, “A Wearable Multi-Modal Bio-Sensing System Towards Real-World Applications,” *IEEE Trans. Biomed. Eng.*, vol. 66, no. 4, pp. 1137–1147, 2019, doi: 10.1109/TBME.2018.2868759.
[37] E. Delaherche, M. Chetouani, F. Bigouret, J. Xavier, M. Plaza, and D. Cohen, “Assessment of the communicative and coordination skills of children with Autism Spectrum Disorders and typically developing children using social signal processing,” *Res. Autism Spectr. Disord.*, vol. 7, no. 6, pp. 741–756, 2013, doi: 10.1016/j.rasd.2013.02.003.
[38] D. Johnston, H. Egermann, and G. Kearney, “Measuring the behavioral response to spatial audio within a multi-modal virtual reality environment in children with autism spectrum disorder,” *Appl. Sci.*, vol. 9, no. 15, 2019, doi: 10.3390/app9153152.
[39] S. Poria, E. Cambria, A. Hussain, and G. Bin Huang, “Towards an intelligent framework for multimodal affective data analysis,” *Neural Networks*, vol. 63, pp. 104–116, 2015, doi: 10.1016/j.neunet.2014.10.005.
[40] C. A. Warner, “Progressive Applications of Extinction Procedures to Assess Response Class Membership of Multiple Topographies of Problem Behavior,” 2019.
[41] K. D. Cantin-Garside, Z. Kong, S. W. White, L. Antezana, S. Kim, and M. A. Nussbaum, “Detecting and Classifying Self-injurious Behavior in Autism Spectrum Disorder Using Machine Learning Techniques,” *J. Autism Dev. Disord.*, 2020, doi: 10.1007/s10803-020-04463-x.
[42] Empatica, “E4 wristband.” https://www.empatica.com/research/e4/.
[43] Microsoft, “Kinect V2.” https://developer.microsoft.com/en-us/windows/kinect.
[44] C. E. Bullock, W. W. Fisher, and L. P. Hagopian, “Description and Validation of a Computerized Behavioral Data Program: ‘bDataPro,’” *Behav. Anal.*, vol. 40, no. 1, pp. 275–285, 2017, doi: 10.1007/s40614-016-0079-0.
[45] Catalyst, “DataFinch.” https://datafinch.com/aba-data-collection/.
[46] A. Sano, J. Hernandez, J. Deprey, M. Eckhardt, M. S. Goodwin, and R. W. Picard, “Multimodal annotation tool for challenging behaviors in people with autism spectrum disorders,” *UbiComp’12 - Proc. 2012 ACM Conf. Ubiquitous Comput.*, pp. 737–740, 2012, doi: 10.1145/2370216.2370378.

# Chapter 4: A Novel Interactive Augmented Reality Coaching System for Toothbrushing Skills in Children with ASD


## 4.1 Abstract

Autism Spectrum Disorder (ASD) is a common neurodevelopmental disorder that impacts one in every 54 children in the United States. Some children with ASD have learning and fine motor skill impairments, which lead to difficulties completing independent daily living task such as toothbrushing. Lack of toothbrushing skills may cause increased need for dental care and negative social feedback from peers. Technology based intelligent intervention systems offer the advantages of being accessible, engaging and cost-effective. In this work, we present a novel interactive augmented reality coaching system, CheerBrush, to improve the toothbrushing skills of children with ASD. CheerBrush allows children to manipulate virtual objects like a toothbrush and toothpaste with their actual hand motions to practice the steps of toothbrushing. The virtual tasks of CheerBrush demonstrate these steps and guide the children by audio and visual cues, while also showing the brushing process through a virtual avatar. CheerBrush also assesses toothbrushing skills with a custom designed mechatronic toothbrush to evaluate the system's coaching effectiveness. A feasibility study with 8 children (4 children with ASD and 4 typically developing children) was conducted to evaluate the acceptability and effectiveness of CheerBrush. The data collected during the feasibility study showed improvements in the toothbrushing motions and reduced stress for the children in the post-test. CheerBrush catches real-time movement of children and interacts with them by augmented reality, feedback and multimodal hints. We believe that CheerBrush has the potential to provide a low-cost, engaging and beneficial intelligent intervention system to improve the toothbrushing skills of children with ASD.


## 4.2 Introduction

Autism spectrum disorder (ASD) is a common neurodevelopmental disorder characterized by deficits in communication and social interaction together with restricted, repetitive and stereotyped patterns of behavior [1-3]. One in 54 children is diagnosed with ASD in the US [4, 5], with prevalence rates amongst school-aged children increasing from 1.16% to 2.00% between 2007 and 2012 [6]. Although ASD is a neurodevelopmental disability with life-long impact, intensive educational practices and behavioral interventions can make a positive impact on the lives of children with ASD [7, 8]. However, conventional interventions for ASD are costly and often inaccessible [9, 10] due to limited resources, with some treatments negatively impacted by low child engagement or motivation [11]. It is estimated that the average lifetime cost for ASD care is up to $3.2 million for each individual with ASD and their families [12].

Therefore, the development of novel ASD intervention paradigms that can provide low-cost and efficacious treatment options for a broader ASD population are important and in urgent need.

In recent years, computer assisted ASD interventions have shown potential due to their low-cost, their appeal to children with ASD, and their relatively broader accessibility [13-15]. Many children with ASD exhibit a natural affinity for computer technologies that lead to a higher level of engagement and fewer disruptive behaviors in computer-based interactions [16, 17]. In particular, virtual reality (VR) technologies that allow users to actively participate in the interactive and immersive simulated situations have been used to provide an attractive, cost-effective, replicable, quantitatively measurable, and controlled intervention environment with real-time feedback [18, 19].

Several VR-based systems have been developed to investigate and teach important life skills to children with ASD and results suggest that children were able to appropriately understand and use and react to virtual environments. For instance, a novel VR-based driving stimulator was developed to teach driving skills to teenagers with ASD [20]. The participants drive a virtual vehicle in a virtual city to complete driving tasks such as passing traffic lights, pulling over, and entering the highway. The simulator detects participant errors and eye gaze and provides appropriate instructions. Another VR-based system, a virtual haptic training system, was designed to assess and improve fine motor skills on children with ASD [21]. The system allows participants to grip and move virtual objects in games and thus provides opportunities for them to improve finger and hand motor control. Additionally, a VR-based social cognition training system was developed where participants can practice social tasks including social introductions, conversation initiation, meeting friends, and other social interaction scenarios [22]. These VR-based intervention systems have been shown to be acceptable, beneficial, as well as accessible to children with ASD to help with a variety of living skills.

The success of VR ushered in augmented reality (AR) into intervention systems for children with ASD. AR is a computer-based technology that superimposes real-world actions and images on computer-generated displays of virtual characters, scenarios, and interactions, providing a composite view of the situation [23-24]. In recent years, AR-based interventions for children with ASD have been reported in literature. A mirror-like AR system that allowed participants to see themselves and their constructed skeletons on the screen was developed to teach children with ASD about their own body and allowed them to imitate body gestures [25]. In another study, a virtual agent was developed to let young adults practice job interviews [26]. The participant wears a Magic Leap AR goggle [27] so that they can see a virtual interviewer in the real world and interact with it. In [28], an AR smart glasses system was developed to coach social communication where gamified AR applications provided children with ASD coaching for emotion recognition, face directed gaze, and eye contact. The coaching intervention was found to be well tolerated, engaging, and fun at the same time. Although an emerging field, existing research to date has shown the potential of AR-based intervention systems to help children with ASD learn life skills by immersing them in real-life situations [29].

Many children with ASD experience difficulties in learning good daily living activity habits potentially due to their fine motor dexterity and learning impairments [30]. In particular, many children with ASD have problems with maintaining good oral health [31, 32] and find difficulties in learning how to brush their teeth, at times needing assistance when brushing [33, 34]. Lack of toothbrushing skills can lead to the need for increased dental visits, which is already extremely traumatizing for children and even adults with ASD due to non-compliance and sensitivity [35]. Besides health implications, poor daily living activity habits may also impede one's social interactions across their life-span [36].

In recent years, there have been a few technology-based systems that aim to help children with ASD to improve their toothbrushing skills. In [37], a toothbrushing training program on tablets shows the steps of toothbrushing to children with ASD, which resulted in some improvement. Another study used marker-based video triggering software to show children with ASD a clip of a peer brushing her teeth [38]. All participants learned how to brush their teeth and maintained the skill in their daily life for 9 weeks after the study. A picture exchange communication system (PECS) based toothbrushing program was used on gingival health in children with ASD. The parents of 37 children rated the program as useful in improving gingival health for children with ASD [39]. A cartoon game called "Brush Up" was utilized on children with ASD and it resulted in significantly reduction of visible plaque on post-intervention [40]. Although promising, the existing studies are limited to showing photos and video clips in an open-loop manner and did not teach toothbrushing skills in an interactive way to provide real-time feedback and instructions based on the performance of the children. We believe that a closed-loop system that can engage children through active participation and provide adaptive and individualized feedback in real-time will likely be more engaging and effective in imparting skill as is evidenced in other skill learning activities [41]. The focus of the current work is to help children with ASD with their oral health through an AR-based intervention system.

Thus, we present CheerBrush, a novel interactive augmented reality coaching system for toothbrushing skills, to help children with ASD learn independent toothbrushing skills. CheerBrush, which has both a coaching mode and an assessment mode, is an AR-based system that teaches children with ASD how to brush their teeth in a fun, playful way. In the coaching mode, CheerBrush combines a virtual realty task environment with an augmented reality motion tracking system to create an immersive practicing and learning experience for toothbrushing. The VR provides a toothbrush and toothpaste in a bathroom environment with an avatar that guides children through the whole brushing process in a step-by-step manner with appropriate verbal and gestural help and feedback. The system provides several levels of practice opportunities based on task difficulty. The AR mechanism using the Microsoft Kinect V2 camera projects the faces of the children on to the virtual environment and tracks and maps their actual brushing motion in real-time and overlays it on their face within the VR task environment. A supervisory controller monitors the whole process in a closed-loop manner and informs the avatar how to guide each participant in an individualized way. In the assessment mode, CheerBrush involves children brushing with an actual custom-designed mechatronic toothbrush equipped with sensors to measure her physical brushing motion instead of a virtual brush and overlays the motion on to her face in the VR environment without any guidance from the avatar. In addition, in both coaching and assessment modes, CheerBrush measures stress responses of children through a wearable physiological sensor, E4. The idea is to observe whether children experience stress during brushing and whether stress level reduces with coaching using CheerBrush.

The primary contributions of this work are the design and development of a novel AR-based daily living activity intervention system for children with ASD that can provide autonomous coaching for toothbrushing in an individualized and adaptive manner based on quantitative measurement of the brushing process. We also provide results from an initial feasibility study with children with ASD as well as typically developing (TD) children to demonstrate the potential of CheerBrush. The rest of this chapter is organized as follows: Section 4.3 describes the CheerBrush system design including the software development. Section 4.4 presents the feasibility study to validate the acceptability and potential of the system through human-participant experiments. It also presents the experimental results and data analysis. Finally, we conclude the chapter in Section 4.5.

## 4.3 CheerBrush System Design

CheerBrush shows the real-world surroundings, observes the toothbrushing movements of children, moves virtual objects based on real-time movements of children and provides virtual reality-based coaching to the children. It also measures physiological signals during coaching and movement data during the pre- and post-tests. There are several modules of CheerBrush: 1) an augmented reality platform (ARP) that projects the real-world surroundings around the children and demonstrates toothbrushing through the animation of virtual avatars as well as highlighting the target areas; 2) a Microsoft Kinect V2 sensor to capture the toothbrushing movements and facial expressions of the children in real-time; 3) a customized mechatronic toothbrush that can measure the toothbrushing motion and location of brushing on the face of children; 4) an E4 wearable sensor to capture physiological responses to infer stress experienced by the children during coaching; and 5) A supervisory controller to ensure communications among various system modules, to provide appropriate feedback to the children, and to collect data for analysis. The system architecture is shown in Fig. 1.

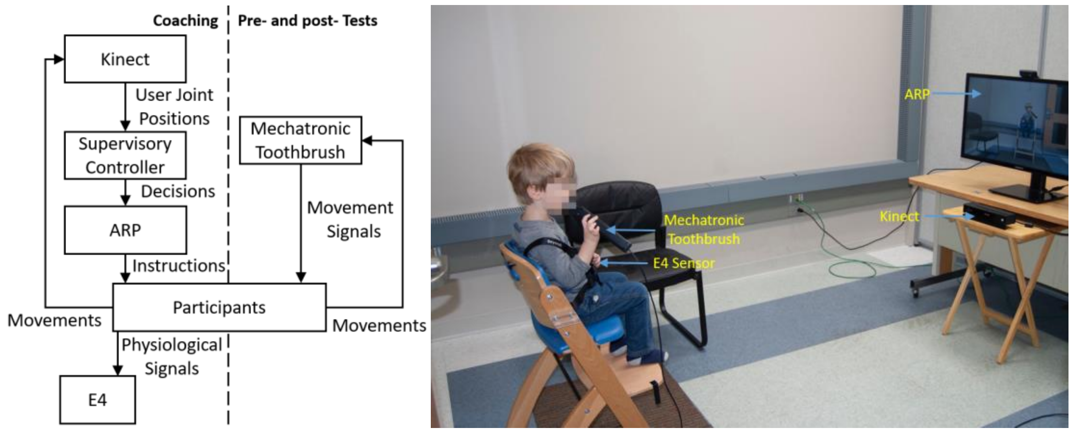



**Fig. 1.** System Architecture and Setup

### *4.3.1 ARP Development*

The ARP was designed to display the combined real-world surroundings with the virtual environment and objects. The virtual environment is a bathroom where the relevant virtual objects include a virtual toothbrush, a glass cup, and a toothpaste. Its graphical user interface (GUI) uses the video stream of the Kinect as the base display so that the GUI works as a mirror. Children can see themselves with their real-world surroundings and their body movement in real-time that are superimposed on to the virtual world. The children can manipulate the virtual objects on the ARP by grasping them. They move their arms, hands, and fingers in the air in front of them, which are then mapped on to the virtual space where the virtual objects lie.

The GUI is presented on a 40-inch computer screen. The ARP was designed and developed using the Unity 3D game engine [42], which is a developmental platform to help create VR and AR environments. The ARP gives real-time instructions and feedback to the children. A screen shot of the ARP and the real-world that it represents at the same moment during experiment are shown in Fig. 2. As shown, children can see themselves in an augmented mirror projected in front of them, and they can interact with the virtual objects. We have designed visual cues in the ARP to help children find the virtual objects and understand where and when to use them. For example, the red circle in Fig. 2 indicates where the hand is, and the green circle highlights the brush he needs to grab. These visual cues will transfer to the correct positions and orientations corresponding to the gestures of children. Within the ARP, the visual cues take shape as blinking arrows, semi-transparent circles and squares that highlight the areas of interest. The colors and the blink rate of these cues vary to keep the children engaged. There are also visual rewards to keep the children motivated and provide reinforcement for their behaviors. For instance, a virtual gold coin will make a jingle sound when a child completes a subtask and his score, which is displayed on the screen, is increased by 1 point.

**Fig. 2.** Real-world picture of the child in the experiment room (left) and his projection on the ARP (right)

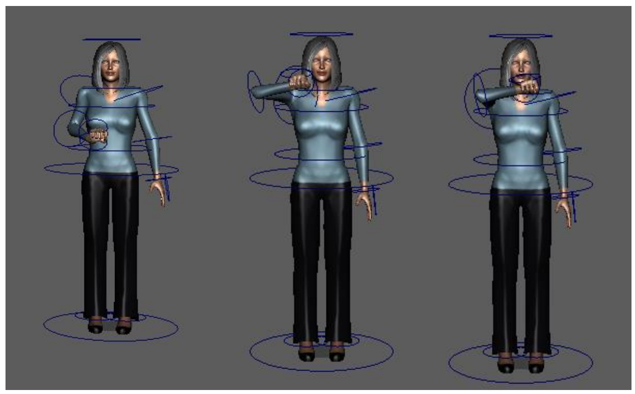


**Fig. 3.** The embedded avatar with animation capability in ARP

To demonstrate the appropriate toothbrushing movements, an avatar with several animations was created and embedded in the ARP as can be seen in Fig. 3. In order to clearly instruct and encourage the children, a few prior-recorded audio clips such as “Grab the brush”, “Go back and forth”, and “Good job” were delivered through the avatar at appropriate times using the supervisory controller.

### *4.3.2 Real-time Movement Detection*

In order to create a real-time closed-loop coaching system, CheerBrush needs to recognize the children and their brushing motion along with the location of their brushing within the facial area so that it can provide real-time instructions and feedback as well as assess the effectiveness of brushing. We have integrated a Microsoft Kinect V2 and its SDK to accomplish these objectives. The Kinect uses a color camera and a depth sensor to provide video stream and compute human joint positions and facial expressions. In particular, the joint positions are used to capture the hand and head positions of the children. We also use Kinect data to determine whether the hand of children is open or closed. The facial expressions data are used to determine whether the eyes and mouth of children are open. The approximate eye gaze of the child is detected by the computer vision algorithm of the Kinect to infer whether children are looking at the screen or looking away. The video stream of Kinect is updated at about 30 frames per second, which is sufficient for smooth display. The joint positions of children are updated at about 10Hz. We program the Kinect such that it tracks the first child that enters its field of view.

In order to smoothly combine the physical movement of children with the motion of virtual objects, we needed to calibrate the ARP. For example, when the children grab a virtual toothbrush and bring it to their mouth, the motion of the virtual toothbrush must be consistent with the motion of their arm and mouth to make it appear that the toothbrush is moving as grabbed and touching their mouth. In order to perform this calibration to map the motion of children with different sizes, we introduced the following coordinate transformation between the real-world coordinates of children detected by the Kinect and the ARP canvas coordinates as shown in Eqns. (1) and (2).

$$x_c = a_1 x_r + b_1 \tag{1}$$

$$y_c = a_2 y_r + b_2 \tag{2}$$

Here $x_c$ and $y_c$ are the canvas coordinates of children's joints on the ARP display and $x_r$ and $y_r$ are the real-world coordinates of children's joints detected by the Kinect. The coefficients $a_i$ and $b_i$ are related to the physical sizes of children. To conveniently and precisely acquire these coefficients to customize CheerBrush for children of different sizes, we conducted a human centered design approach where we invited 5 children of ages between 3 and 8 years. We measured their arm lengths, head circumferences, and heights since these were the key physical differences of the children that impacted the complementary motion of the virtual objects in our system. We then determined the coefficients in Eqns. (1) and (2) using regression analysis as functions of the above-mentioned key physical parameters through a task where the children were required to grab and move a virtual ball.

### *4.3.3 Mechatronic Toothbrush*

We need to measure and record how the children use a toothbrush in the real-world to both assess whether the coaching is effective and to provide feedback. In order to achieve these objectives, we designed a mechatronic toothbrush that had the shape of a regular electrical toothbrush with an integrated motion measurement unit. The case of the brush as shown in Fig. 4 was designed using the computer aided design (CAD) software SolidWorks [43] and 3D printed by a Stratasys 3D printer [44]. The electrical elements, which included an inertial measurement unit (IMU) [45], an Arduino Nano microcontroller [46], a wireless transmitter [47] and a 9V Lithium-ion rechargeable battery, were then integrated inside the 3D printed case. The electronics components and the communication scheme are shown in Fig. 5. The IMU measures the inertial forces along the 3 orthogonal axes and then computes the roll, pitch, and yaw angles of the toothbrush. It also measures the absolute orientation and the angular velocity at 100Hz and sends them to the Nano controller through I2C communication protocol. The Arduino Nano is a small microcontroller with general input and output ports that runs at 16MHz. The wireless transmitter forwards the data from the microcontroller to the supervisory controller on a workstation. The wireless transmitter is a single chip radio transceiver with a 2.4GHz radio band, a maximum baud rate of 2Mbps, and a maximum transmission range of 100 meters. The total weight of the prototype is 94.4 grams and the total cost is $71 excluding the computer.

We have developed software to gather and send the roll, pitch, and yaw angles from the IMU to the supervisory controller in the following manner. First, in the Arduino integrated development environment (IDE), we sent the data from the IMU to the radio transmitter. Then, another controller, an Arduino Mega [48], connected to the workstation, reads from the radio and sends the data to the serial port of the workstation. Next, the supervisory controller reads the data from the serial port and formats the data with time stamps and records them for future analysis.

From these signals, the positions of the tip of the mechatronic toothbrush are computed by Eqns. (3)-(5).

$$x_e = x_h + 0.5L_{MT}\sin\theta \tag{3}$$

$$y_e = y_h + 0.5L_{MT}\sin\psi\cos\theta \tag{4}$$

$$z_e = z_h + 0.5L_{MT}\cos\psi\cos\theta \tag{5}$$

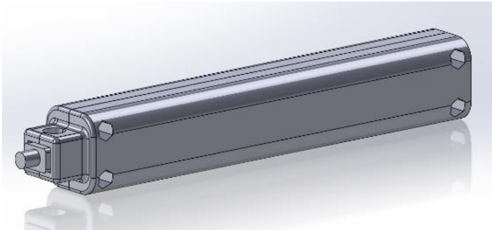


**Fig. 4.** The CAD design of the mechatronic toothbrush

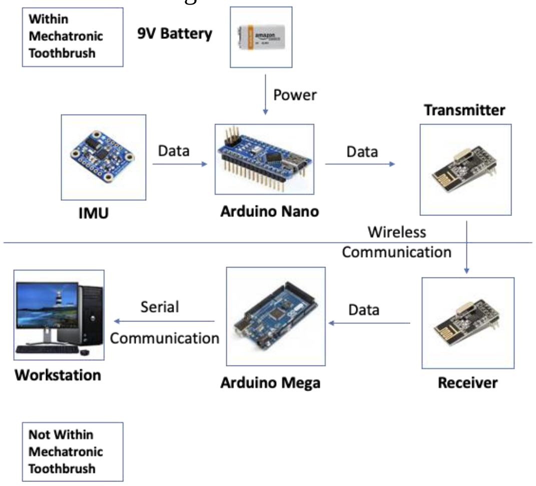



**Fig. 5** Electronic components and communications within the mechatronic toothbrush

Here $x_e$, $y_e$ and $z_e$ are the 3D position coordinates of the end of the mechatronic toothbrush; $x_h$, $y_h$ and $z_h$ are the 3D position coordinates of the hand of children detected by Kinect; $\psi$ and $\theta$ are the roll and pitch angles of the mechatronic toothbrush as detected by the IMU integrated in the mechatronic toothbrush; and $L_{MT}$ is the length of the mechatronics toothbrush. From the motion of the tip position, we compute several metrics such as the frequency of a child's brushing movement, the variances of the brushing motion and the areas and angles she has covered during brushing. We evaluated toothbrushing skill level by comparing these data to the data obtained from adults who are able to properly brush their teeth.

### *4.3.4 E4 Wristband*

In order to track the physiological responses and estimate the stress of the children during the coaching sessions, we used an E4 sensor [49]. The E4 wristband is a medical grade wearable device that collects physiological signals of blood volume pulse (BVP) and galvanic skin response (GSR), from which heart rate (HR) and skin conductance can be computed. It has the shape of a smartwatch and the children can wear it on their wrist. It has been demonstrated in the literature that peripheral physiological data can be used to infer human mental states [50, 51]. In this research, we use the physiological data of children to infer their stress levels. We log these data for offline analysis to understand whether the children were stressed during the activities. Such information will potentially guide us to improve task design and interaction protocols in the future.

### *4.3.5 Task Design*

In consultation with several ASD clinicians and parents of children with ASD, we deconstructed the process of toothbrushing into smaller subtasks, which were then used to demonstrate the toothbrushing process and guide the children. These subtasks were: "grab the brush", "put toothpaste on the toothbrush", "open your mouth", "put the toothbrush on your teeth", and "move the brush back and forth on your teeth". We designed and created four levels of toothbrushing tasks with varying difficulties: "beginner", "simple", "medium" and "hard", so that the children could first become familiar with CheerBrush and gradually get coached towards the complete toothbrushing process. The levels differ in multiple areas such as how children must grab the toothbrush, the number of times children must move the toothbrush back on forth on their teeth, and the inclusion of toothpaste. For example, in the beginner level, children are not required to grab the brush; instead, as soon as the level begins, the brush is virtually attached to their hand. In the simple level, children must actually grab the brush, and in the medium level, children must move the brush back and forth on their teeth five times as opposed to the three times which were required in the beginner and simple levels. After children master the toothbrushing tasks in the beginner, simple, and medium levels, we introduce the toothpaste in the hard level. In the former levels, the toothpaste was already on the brush so it was not necessary for children to grab the toothpaste and put it on the brush. In the hard level, however, the children must apply the toothpaste and move the brush back and forth on their teeth ten times to successfully complete the level.

### *4.3.6 Supervisory Controller*

The primary purpose of the supervisory controller is to coordinate the step-by-step demonstration of toothbrushing in the ARP. In addition, the supervisory controller needs to ensure communication between the Kinect and the ARP, capture toothbrushing data from the mechatronic brush, and provide real-time feedback. The steps of toothbrushing and transitions among them were formalized through a finite state machine (FSM). We implemented the FSM using the programming language C# in the integrated IDE of Microsoft Visual Studio. A FSM is a mathematical model of computation based on the ideas of a system changing state due to inputs supplied to it with conditions that govern state to state transitions with some states being the final states [52]. It is easy to use, provides powerful algorithms for synthesis and verification, and has been successfully used in game design. A FSM is defined by the quintuple $(Q, \sum, \delta, Q_0, F)$, where $Q$ is the set of finite number of states, $\Sigma$ is a non-empty set of symbols of input to the states, $\delta$ is the state transition function, $Q_0$ is the initial state, and $F$ is the set of final states.

For our FSM:

$$Q_0 - Q_4 = \{Start, GrabBrush, BrushOnTeeth, Brushing, LevelFinish\}$$

$$\sum_i = \{x_h, y_h, z_h, s_h, \cdots, s_m, t\}$$

$$Q_0 = \{Start\}$$

$$\delta_1 : Q_0 \times \sum(User\,\mathrm{Re}cognized) \to Q_1$$

$$\cdots$$

$$\delta_5 : Q_3 \times \sum(EnoughTimes) \to Q_4$$

$$F = \{LevelFinish\}$$

The $Q_0 - Q_4$ are the states of the FSM. The $\Sigma_i$ are the inputs of the FSM. The inputs are: positions of hand $x_h$, $y_h$ and $z_h$; whether the hand and mouth are open or not denoted by $s_h$ and $s_m$, respectively; positions of the mouth $x_m$, $y_m$ and $z_m$; and the time spent in a certain level, $t$. $\delta_i$ are the state transition functions expressing the input needed for a certain state to move on to the next corresponding state. The $Q_0$ and $F$ stand for the start and finish states, respectively. The start state and the finish state are shown as "Start" and "Level Finish" in Fig. 6, respectively. There are 5 state-transitions functions to cover the virtual tasks that help FSM proceed from the current state to the following state based on inputs. The supervisory controller runs the FSM for the whole coaching procedure, communicates among different system components, and makes the right moves for different situations. An example FSM for the medium difficulty level is shown in Fig 6. Each difficulty level has different number of state transition functions.

In Fig. 6, the FSM starts with the “Start” state when the Kinect will be actively looking for the children. If it recognizes one, it will stop looking and take that child as the participant and move to the next state “Grab Brush”. In the “Grab Brush” state, the virtual brush will show up on the ARP and the child is supposed to open her dominant hand, move it to the right position on the ARP and close the hand to grab the virtual brush as shown in Fig. 7. During the same time, multimodal cues will be delivered to the child, which are: a blinking arrow that continuously points from the hand of the child to the brush to indicate where to move to grab the brush, the avatar moving her arm and closing her hand to show how to grab the brush, and an audio cue saying, “Grab the brush”. If successfully grabbed, the virtual brush will move with the hand of the child. Rewards will also be given in the form of a falling gold coin with sound effects. A numerical score will be added and shown with the audio cues offering encouragement such as, “Good job,” or, “Way to go”. The score is shown in the upper right corner of the ARP. Then the FSM moves on to the state “Put Brush on Teeth” where the cues of opening the mouth and putting the brush will be given. A blinking black box around the mouth will show up for the child. After successfully putting the brush on her teeth, the FSM will move on to the state “Brushing” and two blinking boxes will show up on the left and right side of the mouth to provide guidance with regard to the brushing motion. The child is supposed to move the brush from the current red box to the target green box. Once the target box is hit, it will become red and the other box (i.e., the former current box) will become green so that the child can move the brush back and forth. During the same time, the supervisory controller computes the speed of the back and forth motion and there will be audio cues “brush faster” or “brush slower” for the child to keep the brushing speed within an appropriate range. When either the task is completed or the time spent during a certain level reaches the limit, the state will directly move on to “Level Finish”.

When the FSM of a single difficulty level reaches the final state, “Level Finish”, it moves on to a harder or an easier difficulty level based on whether the child completed the level. The difficulty levels are beginner, simple, medium, and hard. The levels differ on various parameters including the time to brush back and forth, objects to manipulate like the toothpaste and degree of difficulty to grab the brush. The FSM of different levels are similarly designed. The FSM runs with an average updating frequency of 48Hz, which is fast enough to provide a smooth real-time interaction. We carefully debugged the system and since the system has only a small number of states, we applied an extensive brute-force test to explore every possible scenario to ensure that the FSM ran robustly. We did not observe any system glitches during experiments.

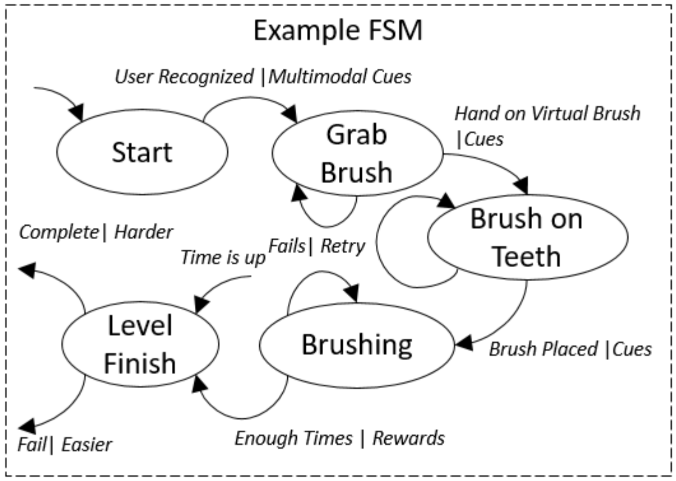



**Fig. 6.** An example FSM showing different states of the medium difficulty level

## 4.4 Feasibility Study and Data Analysis

### *4.4.1 Experiments*

We conducted a feasibility study with children, who did not know how to brush their teeth, to explore the acceptability and usefulness of CheerBrush. We recruited 8 children (6 males, 2 females) from 3-6 years old as participants. There were 4 children with ASD and 4 TD children. The means and standard deviations of the age of two groups are 4.98 and 1.157, and 4.25 and 0.957, respectively. The experimental protocol is shown in Fig. 8. It is to be noted that the purpose of this current experiment was to assess whether CheerBrush was acceptable to the target population and whether they were interested in interacting with the game, and whether there was any potential for skill learning.

We first put the E4 sensor on the wrist of children. Then we let the children play with the toys scattered around the experiment area in the presence of their parents. When the parent informed us that the children seemed comfortable in the new environment, we started recording their baseline physiological data. After five minutes of continued play, we asked the children to go into the experimental space and we put them in a Rifton chair (a supported chair with straps, appropriate for a variety of ages and developmental levels) facing the screen and Kinect. In the current set-up, for best detection accuracy by the Kinect, the children needed to be at least 1.0 meter in height and sitting 1.5 meters away from the Kinect. The experimental space is shown in Fig 9. We played an appropriate cartoon or comic video on the monitor to attract the attention of children and measure their arm lengths, head circumferences and heights while they watched the video. With these measurements, the ARP calibration was completed. Children first underwent a pre-test where they were asked to use the mechatronic brush to brush their teeth. All the data was recorded for analysis. Then the children went through the CheerBrush coaching, and finally they were asked to brush again using the mechatronic brush as a post-test to assess any improvement. The whole session lasted no more than 30 minutes. We obtained approval for this study from the Institutional Review Board (IRB) of Vanderbilt University. There were no known risks to children or parents regarding participation in the study and all procedures were in compliance with IRB approved procedures.

### *4.4.2 Data Analysis and Results*

All children managed to finish the experimental sessions. To quantitatively evaluate the interaction with the CheerBrush, we conducted the following data analysis.

During the game, we used the Kinect SDK to measure whether the children were engaged or looking away from the screen. From this data, we analyzed the percentage of their engaging time. In Fig. 7, it can be observed that the children with ASD looked at the relevant regions of interests (ROIs) such as the face, brush, and the avatar about 65% of the CheerBrush coaching time while their TD counterparts paid attention to those ROIs about 80% of the time. The children with ASD looked away from the screen about 22% of the coaching time, compared to 10% for the TD children. The rest of the time they looked at the screen but not at the relevant ROIs. While there were differences in engagement between the TD and ASD groups, they both were engaged for the majority of the interaction. From these results it can be inferred that the AR-based toothbrushing game was attractive and had the ability to engage children.

The performance of the children during different difficulty levels was also recorded as shown in Fig. 8. On average, the ASD group and the TD group completed 3.00 and 3.75 levels of the game, respectively. The average and ranges of time spent on different levels show that the ASD group required more time at each level although the difference was most pronounced at the most difficult (Hard) level, which was expected because children with ASD are likely to struggle with the combined aspects of toothbrushing. On the Beginner, Simple, and Hard levels they spent 6.14%, 34.18% and 33.49% longer, respectively. However, on the Medium level they were roughly equal.

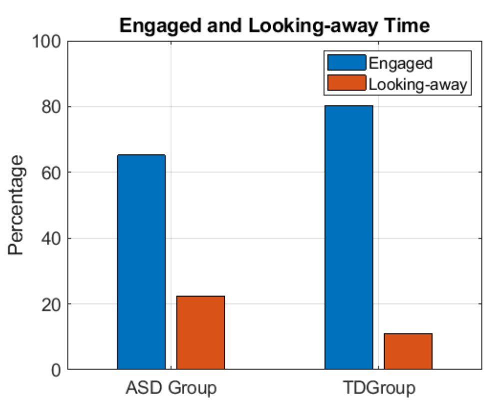



**Fig. 7.** Engagement Time

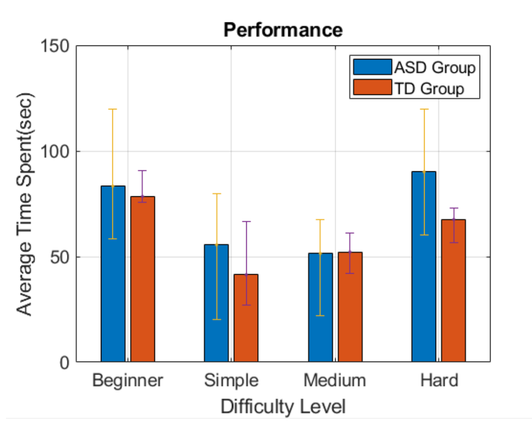



**Fig. 8.** Coaching Performance

During the pre- and post-tests, we measured the speed and amplitude of brushing of the children and compared them to the same features from adult participants who knew how to properly brush their teeth. Two typically developed healthy adult participants were asked to use the mechatronic toothbrush to brush their teeth so that we could obtain the standard brushing speed and amplitude. This comparison can quantitatively explore the potential toothbrushing skills' improvement in the children. The brushing speed and the distance covered by the brush on teeth are the two features that indicate the pattern of brushing [53]. Both brushing speed and distance were measured with respect to the yaw angle of the mechatronic brush and collected from its embedded IMU. Since we are trying to coach the basic toothbrushing skills, advanced toothbrushing skills such as covering occlusal and lingual surfaces are beyond the scope of this study [54]. It can be seen from Figs. 9 and 10 that both sets of children, ASD and TD, improved both their brushing speeds and distances between the pre-test and post-test, indicating the impact of our system's coaching. Though the TD children also improved, their changes were smaller than the children with ASD. These results, although based on a small sample size, indicate the potential of CheerBrush in improving the toothbrushing skills of the children.

We have also measured the physiological responses of the children as they went through the coaching using the E4 sensor. Heart rate (HR) and skin conductance level (SCL) were extracted from the BVP and GSR data obtained from the E4 sensor. These features are strong indicators of stress [55]. We compared the data collected during the coaching sessions to the baseline data. In Fig. 11, the HR levels during the baseline and coaching sessions are shown. The HR levels decreased by 12% for the participants with ASD and it decreased by 8.5% for the TD children showing that the children were less distressed while playing with the system than they were when playing with the toys before the experiment began. Fig. 12 shows the SCL at the baseline and coaching sessions. For children with ASD, the SCL decreased by 28.5% and for the TD children, the SCL stayed the same from the baseline period to the coaching session. The physiological data during coaching sessions had similar forms to the patterns of calmness [56].

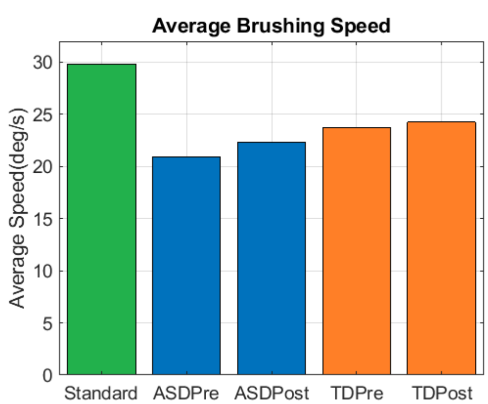



**Fig. 9.** Brushing speeds of children

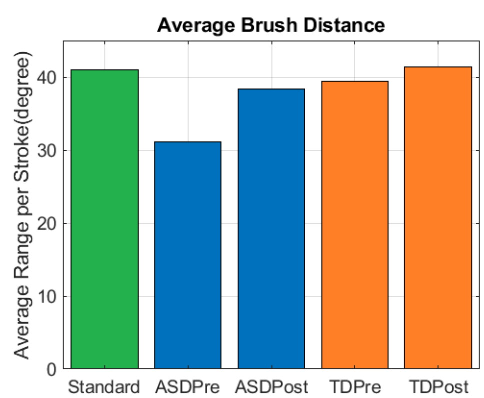
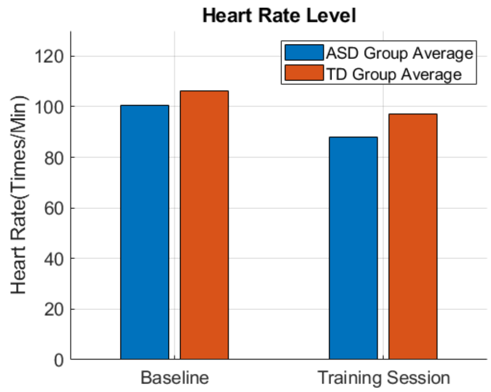



**Fig. 10.** Brushing distances of children



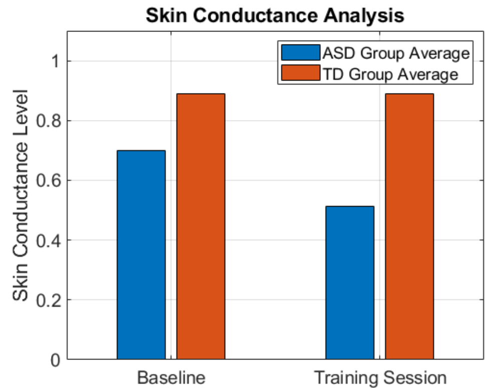




**Fig. 11.** Heart rate comparison

**Fig. 12.** Skin conductance level comparison

Finally, after the experiments, participants and their parents answered questions to evaluate the system feasibility on a 1-5 Likert scale, where 1 represents strong disagreement and 5 represents strong agreement. The questionnaire results are shown in Table 1. Questions 1-3 were asked of children by their parents. Questions 4-5 were answered by the parents themselves.

**Table 1. Questionnaire Results**

| | Question | Avg | Std |
|---|---|---|---|
| 1 | Did you like the game? | 4.18 | 0.60 |
| 2 | Were the instructions helpful? | 4.09 | 0.54 |
| 3 | Do you think this game will teach you (for the kids) how to brush? | 4.09 | 0.70 |
| 4 | Do you think this game will teach your kids how to brush? | 3.91 | 0.70 |
| 5 | Did you think this game made your kids more interested in toothbrushing? | 4.18 | 0.87 |

As seen from the responses, most children had enjoyed interacting with CheerBrush The parents enjoyed CheerBrush and found the instructions beneficial for their children. They also believed that the system would properly coach their children how to brush their teeth, and that it would make them more engaged in the interaction. Additionally, the children themselves reported that the game would be helpful for them to learn toothbrushing. Overall, the responses from the parents and children were positive in all aspects of the questionnaire, indicating that CheerBrush has the potential to impact toothbrushing skills in all children including children with ASD.

## 4.5 Conclusion

We have created CheerBrush, a novel coaching system with interactive AR, for imparting toothbrushing skills in children with ASD. CheerBrush projects the real-world surroundings and children on to a virtual environment and allows for manipulation of virtual objects with real hand and arm motion. The virtual toothbrushing task is broken into smaller subtasks, which are then demonstrated to the children via an avatar's animation and multimodal feedback. A mechatronic toothbrush is designed to measure brushing motion during pre- and post-tests to assess the impact of coaching.

In this work, we have presented the system design, development, and integration of different modules and its feasibility with a pilot study. We conducted a feasibility study of CheerBrush with 8 children. The system functioned as designed and the children with ASD as well as the TD children enjoyed interacting with CheerBrush. The physiological data pattern suggests that the coaching sessions of CheerBrush did not stress the children. A low-stress fun environment for coaching is important so that the coaching can be more effective and beneficial. Their brushing skills as measured from pre and post-test show some improvement, and the brushing movement patterns of the children became closer to the movement patterns of healthy adults.

Although CheerBrush was promising in engaging and coaching children to learn toothbrushing skills through AR technology, the current system and the presented study have several limitations. First, the current system only coaches brushing on mesial and facial surfaces and lacks complexity to include distal and lingual surfaces for complete, healthy toothbrushing. Although the current work is meant for concept demonstration, the system in its current version is not portable and waterproof for bathroom use. The experimental setup requires that children sit 1.5 meters away from the Kinect to obtain the best detection accuracy, which is also a limitation. The feasibility study involved only a small number of participants for one session and consequently the results are not sufficiently generalizable. A longitudinal study with more participants and trials must be conducted to establish long-term gain and transfer of skills into the real world, which is our future goal. Despite these limitations, CheerBrush is promising and demonstrates using quantitative and qualitative data the potential of an AR system in improving daily living activity habits for children with ASD. In the process, we have made contributions in both the intelligent system design and in the field of potential ASD intervention. We believe the proposed system can effectively teach children with ASD toothbrushing skills to not only improve the quality of life of the children, but also for their families and caregivers.

## References

[1] Biao, J., Prevalence of Autism Spectrum Disorders: Autism and Developmental Disabilities Monitoring Network, 14 Sites, United States, 2008. Morbidity and Mortality Weekly Report. Surveillance Summaries. Volume 61, Number 3. Centers for Disease Control and Prevention, 2012.

[2] Kim, Y.S., et al., *Prevalence of autism spectrum disorders in a total population sample.* American Journal of Psychiatry, 2011. **168**(9): p. 904-912.

[3] Johnson, C.P. and S.M. Myers, *Identification and evaluation of children with autism spectrum disorders.* Pediatrics, 2007. **120**(5): p. 1183-1215.

[4] Baio, J., et al., *Prevalence of autism spectrum disorder among children aged 8 years—Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2014.* MMWR Surveillance Summaries, 2018. **67**(6): p. 1.

[5] Maglione, M.A., et al., *Nonmedical interventions for children with ASD: Recommended guidelines and further research needs.* Pediatrics, 2012. **130**(Supplement 2): p. S169-S178.
[6] Blumberg, S.J., et al., *Changes in prevalence of parent-reported autism spectrum disorder in school-aged US children: 2007 to 2011–2012.* National health statistics reports, 2013. **65**(20): p. 1-7.
[7] Howlin, P., I. Magiati, and T. Charman, *Systematic review of early intensive behavioral interventions for children with autism.* American journal on intellectual and developmental disabilities, 2009. **114**(1): p. 23-41.
[8] Iovannone, R., et al., *Effective educational practices for students with autism spectrum disorders.* Focus on autism and other developmental disabilities, 2003. **18**(3): p. 150-165.
[9] Braddock, Barbara, and Kimberly Twyman. "Access to treatment for toddlers with autism spectrum disorders." *Clinical pediatrics* 53, no. 3 (2014): 225-229.
[10] Bonis, Susan. "Stress and parents of children with autism: A review of literature." *Issues in mental health nursing* 37, no. 3 (2016): 153-163.
[11] Maglione, Margaret A., Daphna Gans, Lopamudra Das, Justin Timbie, and Connie Kasari. "Nonmedical interventions for children with ASD: Recommended guidelines and further research needs." *Pediatrics* 130, no. Supplement 2 (2012): S169-S178.
[12] Ganz, M.L., *The lifetime distribution of the incremental societal costs of autism.* Archives of pediatrics & adolescent medicine, 2007. **161**(4): p. 343-349.
[13] Ploog, B.O., et al., *Use of computer-assisted technologies (CAT) to enhance social, communicative, and language development in children with autism spectrum disorders.* Journal of autism and developmental disorders, 2013. **43**(2): p. 301-322.
[14] Wainer, A.L. and B.R. Ingersoll, *The use of innovative computer technology for teaching social communication to individuals with autism spectrum disorders.* Research in Autism Spectrum Disorders, 2011. **5**(1): p. 96-107.
[15] Williams, C., et al., *Do children with autism learn to read more readily by computer assisted instruction or traditional book methods? A pilot study.* Autism, 2002. **6**(1): p. 71-91.
[16] Shane, H.C. and P.D. Albert, *Electronic screen media for persons with autism spectrum disorders: Results of a survey.* Journal of autism and developmental disorders, 2008. **38**(8): p. 1499-1508.
[17] Pennington, R.C., *Computer-assisted instruction for teaching academic skills to students with autism spectrum disorders: A review of literature.* Focus on Autism and Other Developmental Disabilities, 2010. **25**(4): p. 239-248.
[18] Parsons, S. and S. Cobb, *State-of-the-art of virtual reality technologies for children on the autism spectrum.* European Journal of Special Needs Education, 2011. **26**(3): p. 355-366.
[19] Strickland, D., *Virtual reality for the treatment of autism.* Studies in health technology and informatics, 1997: p. 81-86.
[20] Wade, Joshua, Dayi Bian, Lian Zhang, Amy Swanson, Medha Sarkar, Zachary Warren, and Nilanjan Sarkar. "Design of a virtual reality driving environment to assess performance of teenagers with ASD." In *International Conference on Universal Access in Human-Computer Interaction*, pp. 466-474. Springer, Cham, 2014.
[21] Zhao, Huan, Zhaobo Zheng, Amy Swanson, Amy Weitlauf, Zachary Warren, and Nilanjan Sarkar. "Design of a Haptic-Gripper virtual reality system (Hg) for analyzing fine motor behaviors in children with autism." *ACM Transactions on Accessible Computing (TACCESS)* 11, no. 4 (2018): 1-21.
[22] Kandalaft, Michelle R., Nyaz Didehbani, Daniel C. Krawczyk, Tandra T. Allen, and Sandra B. Chapman. "Virtual reality social cognition training for young adults with high-functioning autism." *Journal of autism and developmental disorders* 43, no. 1 (2013): 34-44.
[23] Carmigniani, Julie, Borko Furht, Marco Anisetti, Paolo Ceravolo, Ernesto Damiani, and Misa Ivkovic. "Augmented reality technologies, systems and applications." *Multimedia tools and applications* 51, no. 1 (2011): 341-377.
[24] Yuen, Steve Chi-Yin, Gallayanee Yaoyuneyong, and Erik Johnson. "Augmented reality: An overview and five directions for AR in education." *Journal of Educational Technology Development and Exchange (JETDE)* 4, no. 1 (2011): 11.
[25] Casas, Xavier, Gerardo Herrera, Inmaculada Coma, and Marcos Fernández. "A Kinect-based Augmented Reality System for Individuals with Autism Spectrum Disorders." In *Grapp/ivapp*, pp. 440-446. 2012.
[26] Hartholt, Arno, Sharon Mozgai, Ed Fast, Matt Liewer, Adam Reilly, Wendy Whitcup, and Albert" Skip Rizzo. "Virtual humans in augmented reality: A first step towards real-world embedded virtual roleplayers." In *Proceedings of the 7th International Conference on Human-Agent Interaction*, pp. 205-207. 2019.

[27] MagicLeap 1, MagicLeap, https://www.magicleap.com/en-us/magic-leap-1
[28] Liu, Runpeng, Joseph P. Salisbury, Arshya Vahabzadeh, and Ned T. Sahin. "Feasibility of an autism-focused augmented reality smartglasses system for social communication and behavioral coaching." *Frontiers in pediatrics* 5 (2017): 145.
[29] Khowaja, Kamran, Bilikis Banire, Dena Al-Thani, Mohammed Tahri Sqalli, Aboubakr Aqle, Asadullah Shah, and Siti Salwah Salim. "Augmented reality for learning of children and adolescents with autism spectrum disorder (ASD): A systematic review." *IEEE Access* 8 (2020): 78779-78807.
[30] Murshid, Ebtissam Z. "Parents' dental knowledge and oral hygiene habits in Saudi children with autism spectrum disorder." *Global J of Med Research* 14, no. 2 (2014): 11-18.
[31] American Academy of Pediatric Dentistry. "Guideline on caries-risk assessment and management for infants, children, and adolescents." *Pediatric dentistry* 35, no. 5 (2013): E157.
[32] Du, Rennan Yanlin, Cynthia KY Yiu, and Nigel M. King. "Oral health behaviours of preschool children with autism spectrum disorders and their barriers to dental care." *Journal of Autism and Developmental Disorders* 49, no. 2 (2019): 453-459.
[33] Capozza, Lauren E., and Enrique Bimstein. "Preferences of parents of children with autism spectrum disorders concerning oral health and dental treatment." *Pediatric dentistry* 34, no. 7 (2012): 480-484.
[34] Cuvo, Anthony J., Anna Godard, Rachel Huckfeldt, and Ronda DeMattei. "Training children with autism spectrum disorders to be compliant with an oral assessment." *Research in Autism Spectrum Disorders* 4, no. 4 (2010): 681-696.
[35] Lai, Bien, Michael Milano, Michael W. Roberts, and Stephen R. Hooper. "Unmet dental needs and barriers to dental care among children with autism spectrum disorders." *Journal of autism and developmental disorders* 42, no. 7 (2012): 1294-1303.
[36] McKeown, L. "Social relations and breath odour." *International journal of dental hygiene* 1, no. 4 (2003): 213-217.
[37] Cazaux, S. Lopez, Gaëlle Lefer, A. Rouches, and P. Bourdon. "Toothbrushing training programme using an iPad® for children and adolescents with autism." *European Archives of Paediatric Dentistry* 20, no. 3 (2019): 277-284.
[38] Cihak, David F., Eric J. Moore, Rachel E. Wright, Don D. McMahon, Melinda M. Gibbons, and Cate Smith. "Evaluating augmented reality to complete a chain task for elementary students with autism." *Journal of Special Education Technology* 31, no. 2 (2016): 99-108.
[39]Al-Batayneh, O. B., T. S. Nazer, Y. S. Khader, and A. I. Owais. "Effectiveness of a tooth-brushing programme using the picture exchange communication system (PECS) on gingival health of children with autism spectrum disorders." *European Archives of Paediatric Dentistry* (2019): 1-7.
[40]Cheung, Wing-yan, Hei-nok Fan, Kwong-chuen Lam, Ashley Lau, Wing-tung Leung, Ka-hei Tang, and Kim-wing Tse. "Brush up to a healthy smile: developing and testing an oral health education kit for children with autism spectrum disorder." *Community Health Project* (2018).
[41] Zhao, Huan, Zhaobo Zheng, Amy Swanson, Amy Weitlauf, Zachary Warren, and Nilanjan Sarkar. "Design of a Haptic-Gripper virtual reality system (Hg) for analyzing fine motor behaviors in children with autism." *ACM Transactions on Accessible Computing (TACCESS)* 11, no. 4 (2018): 1-21.
[42] Unity 3D, https://unity.com/
[43] SolidWorks, https://solidworks.com/
[44] F170, Stratasys, https://www.stratasys.com/3d-printers/f123
[45] BNO055, Bosch, https://www.bosch-sensortec.com/products/smart-sensors/bno055.html
[46] Nano, Arduino, https://store.arduino.cc/usa/arduino-nano/
[47] nRF24L01, Nordic semiconductor, https://www.sparkfun.com/products/691
[48] Mega, Arduino, https://store.arduino.cc/usa/mega-2560-r3/
[49] E4, Empatica, https://www.empatica.com/research/e4/
[50] Thayer, Julian F., Fredrik Åhs, Mats Fredrikson, John J. Sollers III, and Tor D. Wager. "A meta-analysis of heart rate variability and neuroimaging studies: implications for heart rate variability as a marker of stress and health." *Neuroscience & Biobehavioral Reviews* 36, no. 2 (2012): 747-756

[51] Fletcher, Anne C., Cheryl Buehler, Amy L. McCurdy, and Bridget B. Weymouth. "Skin conductance reactivity as a moderator of associations between youth perceptions of neighborhood stress and depressive symptoms." *The Journal of Early Adolescence* 39, no. 8 (2019): 1154-1176.
[52] Baheti, Radhakisan, and Helen Gill. "Cyber-physical systems." *The impact of control technology* 12, no. 1 (2011): 161-166.
[53] American Dental Association, "Mouth Healthy: Brushing Your Teeth", 2019
[54] Lee, Young-Jae, Pil-Jae Lee, Kyeong-Seop Kim, Wonse Park, Kee-Deog Kim, Dosik Hwang, and Jeong-Whan Lee. "Toothbrushing region detection using three-axis accelerometer and magnetic sensor." *IEEE Transactions on Biomedical Engineering* 59, no. 3 (2011): 872-881.
[55] Setz, Cornelia, Bert Arnrich, Johannes Schumm, Roberto La Marca, Gerhard Tröster, and Ulrike Ehlert. "Discriminating stress from cognitive load using a wearable EDA device." *IEEE Transactions on information technology in biomedicine* 14, no. 2 (2009): 410-417.
[56] Farrow, Tom FD, Naomi K. Johnson, Michael D. Hunter, Anthony T. Barker, Iain D. Wilkinson, and Peter WR Woodruff. "Neural correlates of the behavioral-autonomic interaction response to potentially threatening stimuli." *Frontiers in Human Neuroscience* 6 (2013): 349.

# Chapter 5: Smart Application to Establish Brevity Control in Functional Analysis

## 5.1 Motivation

Children and adolescents with autism spectrum disorders (ASD) who engage in problem behaviors often cause disruption to the classroom environment and pose safety risks to themselves, other students, and staff [1]. Functional analysis are commonly used to assess controlling variables for behavioral topographies of problem behaviors [2]. Visual analysis take repeated measurements of behaviors and it is a staple of the functional analysis field [3]. Changes in trend, level and variability will impact decisions and modifications during a functional assessment to achieve behavioral stability. When considering the specific relevance to practitioners, a visual analysis will determine if functional control has been established over a socially significant behavior, thus implicating the success or failure of a given behavioral assessment. Without a visual analysis of data depicted in a single-subject design, practitioners' interpretations of an assessment's efficacy will be difficult and often dependent on speculation from anecdotal reports or unsystematic observations.

Conducting a visual analysis first requires the collection of direct measures of behavior. This typically involves recording a count or duration. A recording of a count requires each response to be tallied during an observation period, whereas in a recording of duration a stopwatch is started and stopped with the onset and offset of a target response. Traditional methods for collecting direct measures of behavior can be somewhat cumbersome because the process requires the practitioner to (a) create data sheets; (b) complete data sheets by hand during an observation; (c) calculate post-hoc summary statistics; (d) graph results depicted in sessions and (e) physically store completed data sheets. Therefore, clinicians may find themselves in a difficult situation requiring multitasking while working with participants, which could impact treatment fidelity or service delivery, spending excessive amounts of time and resources towards indirect hours preparing materials. To maintain an efficient assessment period, some clinicians may base interpretations on direct observations of the sessions or examination of raw data, reserving any visual analysis until after the functional analysis is complete.

Computerized programs, such as BDataPro [4], have been developed that could improve data collection for practitioners and applied researchers. The BDataPro program includes practical functions that allow the individual collecting data to create profiles and design individualized keys of frequency or duration-based measures of behavior. Another software for behavioral intervention, Catalyst, allows collection and management of a wide variety of data for behavioral intervention, including skill acquisition and behavior reduction [5]. An annotation tool for problem for people with ASD was also developed to log data more conveniently [6]. While these computerized programs may reduce effort involved in collecting data, they do not necessarily improve visual analysis. That is because the programs do not automatically generate graphs or include supplemental structured criteria that can aid interpretations of recorded outcomes. This is especially important considering the fact that visual analysis can be somewhat subjective and result in inconsistent interpretations without proper training [7][8]. Hagopian et al. developed a set of structured criteria to be used to supplement visual inspection of data collected during a functional analysis [9][10]. Due to the common use of a multi-element design with multiple conditions rapidly alternating with one another, interpreting the results of a functional analysis can be notoriously difficult. Therefore, Hagopian et al. worked with a panel of two experts who had extensive experience conducting functional analyses to develop a set of rules to become the structured criteria. The structured criteria involved calculating a mean rate of responding during the control condition to establish an upper criterion line set at one standard deviation above the mean and a lower criterion line set at one standard deviation below the mean. These criterion lines were meant to be physically drawn on the visual representation of the data so that the points above and below the criteria lines can be counted to calculate a difference. The number of points above the upper criterion line, as opposed to below the lower criterion line, were then used to inform a binary determinant of control.

Jessel et al. [11] extended the structured criteria to provide a more fluid interpretation of functional control by including differing potential levels of control. The multilevel structured criteria included two relevant considerations: Overlap and problem behavior during the control. Strong control was represented by no overlap with the test and control conditions and no problem behavior during the control. Moderate control had some overlap between the test and control conditions or some problem behavior during the control. Whereas, weak control had both overlap and problem behavior during the control. The author suggested this multilevel structured criteria could improve the sensitivity of clinical interpretations beyond the binary, yes or no, conclusions regarding control, further improving and supporting visual analysis. The multilevel structured criteria has also been used to interpret within-session data to determine levels of control in a single 3-min, 5-min, or 10-min session [12]. Jessel et al. calculated rates of problem behavior during periods when the establishing operation (EO) was in place and compared them to rates of problem behavior when the reinforcer was present. The multiple, rapidly alternating intervals allowed for interpretations of differentiated outcomes with strong control obtained in the majority of analyses. It is important to note that greater control did tend to correspond to longer session duration. In fact, there was only one recorded case from the 18 participants in which a worsening in control was observed when the session duration was extended to 10 min. Nevertheless, this suggests that the level of control obtained can vary, and be sensitive to contingencies, on the order of minutes.

We propose to combine the functionality of a computerized program, the Problem Behavior Multilevel Interpreter (PB.MI), for collecting data, with automatically generated graphing and interpretations of control on a real-time ongoing basis. That way the clinicians need only collect data using the program to make informed decisions regarding the introduction of treatment following the functional analysis. The program was specifically designed to be incorporated with the state-of-the-art functional analysis format, the interview-informed synthesized contingency analysis (IISCA) [13]. Such smart application would augment the decision making process of practitioners on both brevity and robustness, having the potential to save significant amount of session time and therapeutic costs.

## 5.2 Application Design

PB. MI allows convenient and precise data collection of direct behavioral measurement. Through real-time graphing and structured control level computations, it offers quantitative and qualitative guidelines for behavioral analysts during sessions. The application is able to be deployed on PC, Mac, Android and IOS platforms.

### *5.2.1 Graphical User Interface (GUI) Design*

The PB.MI was developed through the Unity platform [14]. It has three main pages including setup, recording and summary. The app starts with the setup page, as shown in Fig. 1, which allows behavioral analysts to log in basic information for the session. Behaviors among children with ASD vary in types and behavioral topographies [15], and thus the app allows GUI customization for specific frequent problem behaviors. Behavior analysts enter anticipated problem behaviors and customize hotkeys on the recording page. Once the behavior analyst clicks the start button, the app advances to the recording page and generate a csv file to store the session information.

The recording page, as shown in Fig. 2, features buttons for behavioral observation inputs and texts for session information feedback. On the upper left side, the app shows the current interval from Establishing Operations (EO) and Reinforcement (SR). The behavior analyst can switch between these intervals to stay consistent with the ongoing assessment. There are multiple timers to help behavior analysts to keep track of the session, such as the total session time, current interval time, and total EO/SR time durations. The buttons on the left are for problem behavior registrations. They are customized by the inputs from the setup page (e.g. 'S' for self-injury) and can be triggered by both clicking the buttons and striking hotkeys on the keyboard. Behavior analysts are occupied during assessments [16] and the hotkeys can help them to keep their eyes still on the assessment and record the observations. The switching of intervals, the problem behaviors and their time stamps are all automatically recorded and stored within the PB.MI, precise the nearest millisecond. As a means of providing the behavior analyst with real-time feedback as to the level of control within the current session, three additional categories of data are displayed. The upper right corner of the recording page lists the rate of total problem behavior within each EO or SR interval. Below the rate data, the behavior analyst can see the current level of control for the current session according to the binary structured criteria, multilevel structured criteria, and percentage of non-overlapping data points [17]. At the bottom of the recording page the behavior analyst may refer to a real-time updating graph that depicts the rates of problem behavior for each interval of EO and SR. In addition, the green (EO) and blue (SR) data paths are visible on the recording page, allowing for a visual analysis of differentiation and control. The scale of the graph's y-axis automatically adjusts to fit the rates of problem behavior recorded.

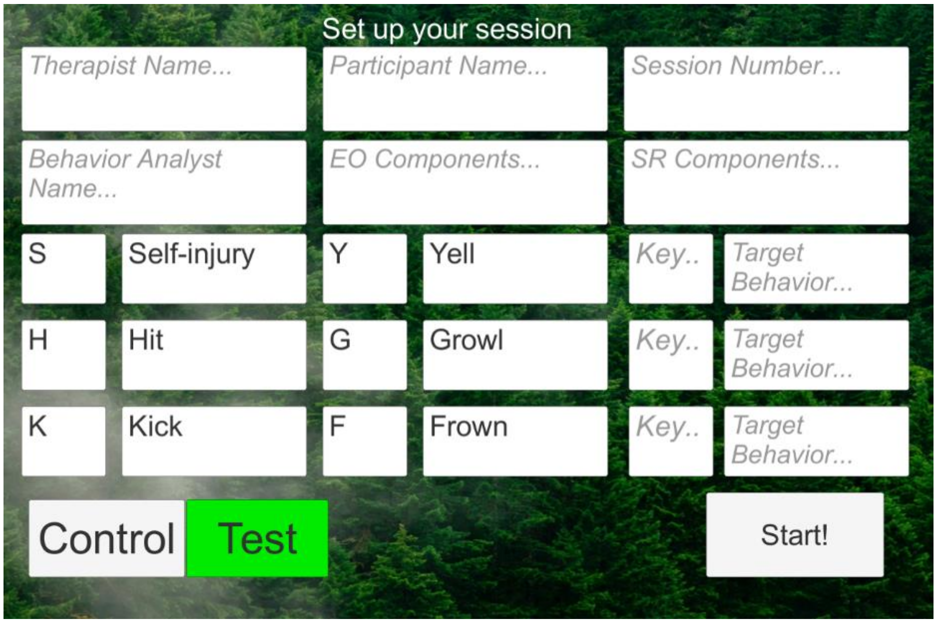



**Fig. 1.** Setup Page of the PB.MI App

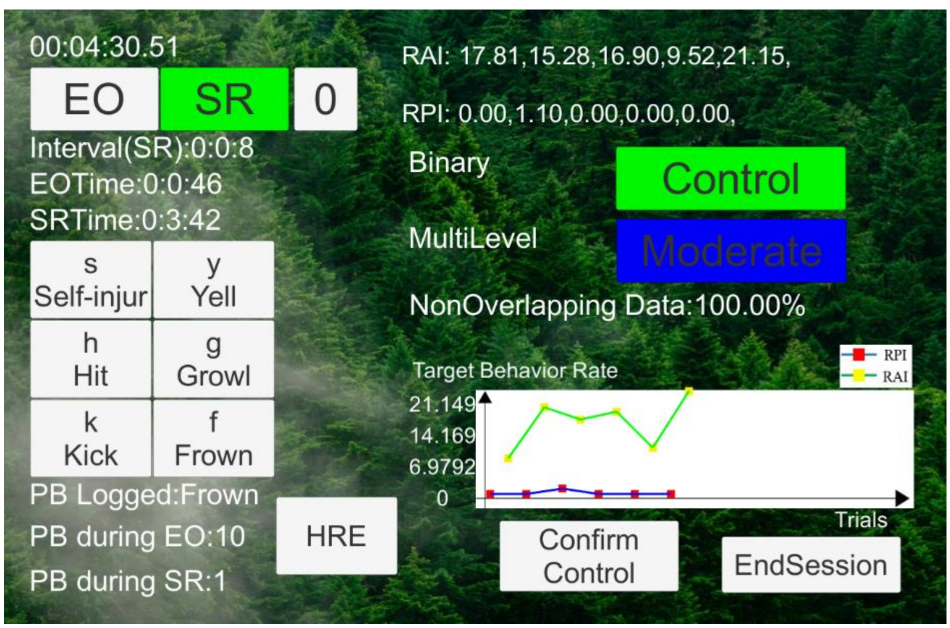



**Fig. 2.** Recording Page of the PB.MI App

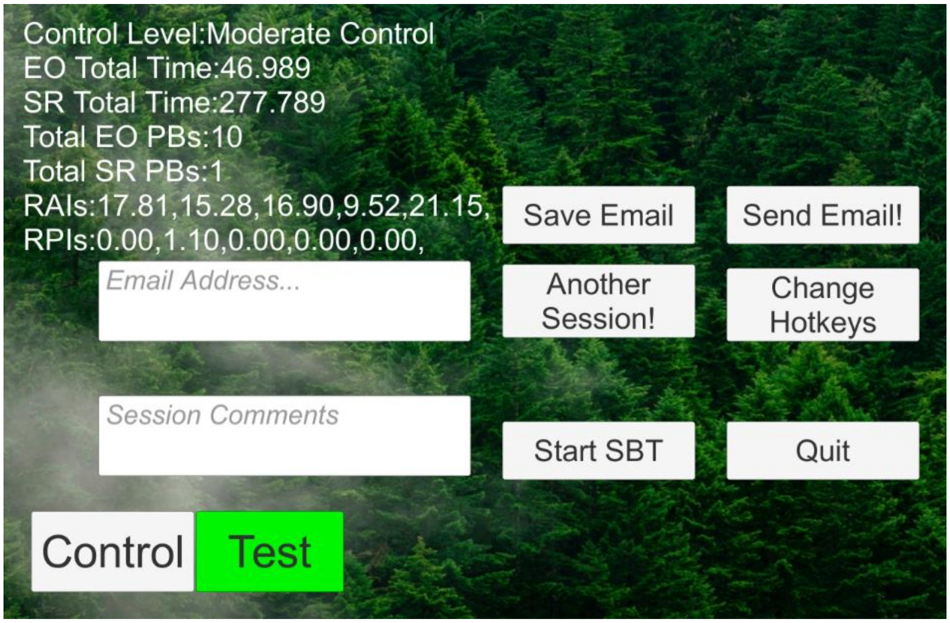



**Fig. 3.** Summary Page of the PB.MI App

When a session is completed, the user can click the "end session" button, advancing the PB.MI program to the summary page, as shown in Fig. 3. In the summary page, the user can send the session data file by email, observe past session information, adjust the hotkeys setting, or start a new session. Session data is exported in a .csv file format. This data output can be used to input data into other software platforms for graphing or analysis. In addition, the behavior analyst has the option to continue data collection during further assessment and treatment. If the behavior analyst were to click the "Start SBT" button, they would be taken to the PreMAC program that has been previously developed and validated for data collection during a multi-session IISCA or skill-based treatment, called the Behavior Data Collection Integrator (BDCI) [18]. Thus, the behavior analyst has the option to (a) continue to collect data during a traditional IISCA that compares performance between alternating control and test sessions if undifferentiated outcomes were obtained during the single-session IISCA or (b) begin collecting data during the treatment phase if differentiated outcomes were obtained during the single-session IISCA.

*5.2.2 Control Establishment Criteria*

As the data are being recorded and stored, the PB.MI program interprets these data to provide a real-time graphical display that allows the behavior analyst to visually analyze the data, as well as monitor the level of control suggested by each of the three criteria. First, rates of problem behavior are computed using the count over the duration of each EO and SR interval. These two intervals were mutually exclusive and one had to be in place at all times. Intervals during the EO were defined as reinforcer absent intervals (RAIs), whereas intervals with reinforcement were defined as reinforcer present intervals (RPIs). Then binary, multilevel control levels are established based on the existing work [11][12]. During a session, there are two data paths of RAIs and RPIs. The binary control is established if Equation (1) holds true, indicating that the percentage of RAIs with a rate of problem behavior higher than the value of the mean rate of problem behavior across all RPIs plus one standard deviation of the RPI mean is at least 50% more than the percentage of RAIs with a rate of problem behavior smaller than the mean rate of problem behavior across all RPIs minus the standard deviation of the RPI mean:

$$\frac{\sum_{i=1}^{Length(RAI)}\left[RAI_i > \overline{RPI} + s_{RPI}\right] - \sum_{i=1}^{Length(RAI)}\left[RAI_i < \overline{RPI} - s_{RPI}\right]}{Length(RAI)} > 0.5 \qquad \text{Equation (1)}$$

$$nonOverlapping = \frac{\sum_{i=1}^{Length(RAI)}\left[RAI_i > \max_{RPI}\right]}{Length(RAI)} \qquad \text{Equation (2)}$$

If the binary control is established, we can further compute the level of the control. The non-overlapping ratio is computed by Equation (2), which is the percentage of RAIs with rates of problem behavior higher than than the highest rate of problem behavior recorded in any of the RPIs. Then, different control levels are established by the multilevel control establishment algorithm, as shown in Table 1. The algorithm determines the level of control as being strong if there is no overlap between the RAIs and RPIs and no problem behavior during any RPIs. If there is some overlap or some problem behavior recorded during an RPI, the algorithm switches the level of control to moderate. Finally, the algorithm indicates weak control if both statements hold true (i.e., overlap and problem behavior during the RPI).

**Table 1.** Multilevel Control Establishment Algorithm

```
If (nonOverlapping =1) and (maxRPI = 0)
    ControlLevel = Strong
Else if (nonOverlapping =1) or (maxRPI = 0)
    ControlLevel = Moderate
Else
    ControlLevel = Weak
End
```

The binary control, multilevel control and the non-overlapping percentage are updated every time a new interval ends, recorded when the user toggles to the opposite condition or clicks the “End Session Button.” The computed and updated control levels are displayed on the recording page in real-time. There are different colors representing different levels of control. For example, when displaying the multilevel control strength, the text that states the level of control appears in green, blue, yellow and red for strong, moderate, weak and no control, respectively.

## 5.3 Validation of Application

We have performed tests to validate basic functioning such as accurate timing, event recording, and summary statistics output [4]. There is no glitches from the app and the timing and event recording were correctly recorded after manual checks. More importantly, tests were conducted to determine the reliability between computer-generated interpretations of control and human interpretations of control using the multilevel structured criteria. Furthermore, a test was conducted to determine the accuracy of the graphical representation created from the data input.

### *5.3.1 Evaluation of Multilevel Control Algorithm*

To investigate how well our algorithm can establish control compared to human behavioral analysts, we conducted a validation study with simulation. We generated virtual sessions for the program validation study. A MATLAB program was written for the simulation. To generate complete virtual sessions, the program has the following steps: 1) generate two random arrays for the time durations of the RAIs and RPIs; 2) for each RAI and RPI, generate a random number of problem behaviors; 3) compute the RAI and RPI and plot the virtual session for visualization; 4) compute the control level of the session. The parameters of random generation were informed by rates obtained from published data sets. For an example, the time durations of the RAIs and RPIs follow a normal distribution with a mean of 30 s and a standard deviation of 10s. The program automatically generates virtual sessions and determines its control levels at the end of the session. Thus, we generated 50 virtual sessions of each potential level of control, from no control, weak, moderate to strong control. The y-axis label of the graph indicated the response measure (i.e., responses per sec) with a range of 0 and 1, while the x-axis label indicated the unit of time (i.e., interval) with a range of 9 to 12 intervals. The ranges were determined based on previously collected, published data sets. The graph included the two distinct intervals (RAIs and RPIs) depicted using different symbols. An example virtual session is shown in Fig. 4.

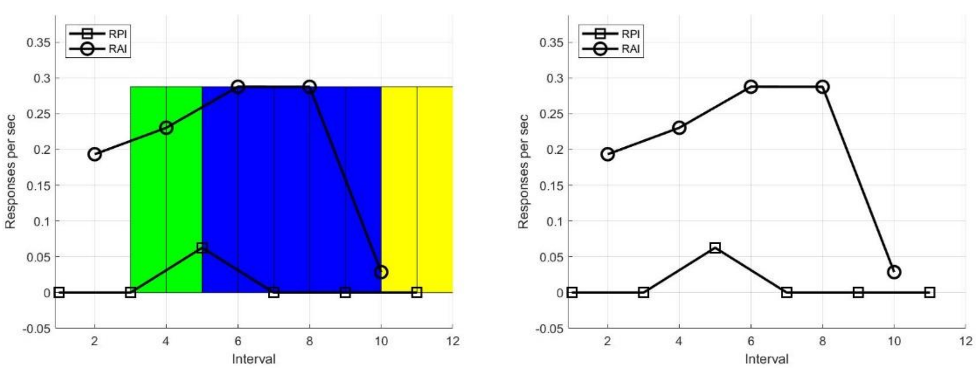



**Fig. 4.** Simulated Virtual Session Example

Indicators of level of control were removed from the graphs, such as the right figure of Fig. 4, and they were assigned a randomized code to ensure that evaluators could not identify the level of control without visually analyzing the data. A trained research assistant and a Board Certified Behavioral Analyst (BCBA) independently evaluated the level of control of all 200 graphs using the multilevel structured criteria. There were two levels of agreement that were analyzed. First, agreement was calculated between the automatically generated level of control and the trained research assistant. Second, agreement was calculated between the trained research assistant and the BCBA. This allowed for the evaluation of both the trained observers' ability to appropriately apply the multilevel structured criteria and the accuracy of the computer code to categorize the level of control. Any disagreements were discussed with the BCBA and deemed either human or computer error. Typically, the computer determined level of control was identified as an error and replaced if both evaluators agreed. In the case of computer error, new graphs were generated for replacement.

Agreement coefficients were calculated by dividing the number of agreements by 200 (the total number of graphs). Agreement between the trained research assistant and the computer generated level of control was 92%. Agreement between the trained research assistant and the expert BCBA was 96%. This result indicates that the multilevel control structure that the PB.MI app uses can establish behavioral control at a high accuracy, which is comparable to a trained research assistant.

*5.3.2 Chance Simulation*

To investigate the possibilities of the PB.MI program incorrectly interpreting moderate or strong levels of control due to random patterns of target behavior, we conducted a simulation experiment. The brevity control establishment proposed in this work usually involve about 5 min sessions, which are comprised of 5 RAIs and 5 RPIs. Thus, we have simulated random occurrences of target behavior across these 10 total intervals. First, we generated random normally distributed session times for the RAIs (*M* = 25 s; *SD* = 5) and RPIs (*M* = 35 s; *SD* = 5 s), separately. Then we generated the random uniformly distributed responses for these sessions, with varying response per second. After rounding the responses, we computed the random RAI/RPI as in Equation 4:

$$RAI\,/\,RPI = \frac{Round\left(rate \times t_{RAI/RPI}\right)}{t_{RAI/RPI}} \qquad \text{(Equation 3)}$$

In this way, we simulated a large number of sessions in which random behavior occurs in a way that could be a representative analogue to real-life response rates. For each response per second rate, we generated one million sessions and applied our multilevel control criteria to determine the level of control for each of these simulated sessions. As Table 2 displays, out of 1 million virtual sessions in which a simulated individual is engaging in behaviors randomly, the criteria will classify more than 90% of them as having no control and about 7% of them as having weak control. The chances of classifying sessions with random behavior as having moderate control are less than 0.4% and it is nearly impossible to classify sessions with random behavior as having strong control. These ratios are consistent across varying rates of responses per second.

**Table 2.** Results of the Chance Simulation

| Rate (per sec) | Strong | Moderate | Weak | No Control |
|---|---|---|---|---|
| 0.04 | 0.0062% | 0.37% | 4.71% | 94.92% |
| 0.1 | 0.0026% | 0.30% | 6.81% | 92.83% |
| 0.3 | 0% | 0.38% | 7.67% | 91.95% |
| 0.5 | 0% | 0.39% | 7.79% | 91.80% |

*5.3.3 PB.MI Data Input Evaluation*

The raw data from the randomly generated graphs were then input back into the PB.MI program by research assistants to determine the accuracy of the program's ability to generate an identical graph with the same data and maintain the interpretations of level of control. The raw data included the time in which each response occurred. The research assistant then started the PB.MI program and manually input each response following the guidance of the raw data file.

To calculate inter-rater reliabilities, we extracted the RAI and RPI recorded by the two raters of each session and analyzed their differences. For each interval in the same virtual session, the two raters recorded different problem behavior episodes and interval durations, resulting in different RAI/RPI values. We computed the difference between the recorded RAI/RPI values in percentage, for each interval, using the equation:

$$e = \frac{2 \times \left| RAI_1 - RAI_2 \right|}{RAI_1 + RAI_2} \qquad \text{Equation (4)}$$

Then, if the error of the recorded values was less than 3%, we would consider this data point to be reliable or inter-rater agreed. We then computed the percentage of reliable data points in each session, and computed the average reliable data percentages for each control level (strong, moderate, weak, no control) and interval type (RAI and RPI). The reliable data results are shown in Table 3. Data recorded for most control levels had greater than 90% reliability, which is well within the "almost perfect" (80%-100%) interpretation of the value of Kappa for inter-rater agreement in clinical research [19]. The data for RPIs have significantly better chances to be reliable, potentially due to the infrequent occurrence of problem behaviors. We then computed the p values of the error arrays for each control level using the MATLAB function two-sample t-test. Every control level except for the moderate level reached a p-value less than 0.05, statistical significance. This result indicates that the possibility of reaching these inter-rater reliabilities being a coincidence is very low.

**Table 3.** Reliability Results

| | Strong | Moderate | Weak | No Control |
|---|---|---|---|---|
| RAI | 98.7%*** | 76% | 89.3%* | 96%* |
| RPI | 100%*** | 90.7%*** | 97.3%* | 98%*** |

*Note*. Three asterisks indicate a <.001 p-value. Two asterisks indicate a <.01 p-value. One asterisk indicates a <.05 p-value. RAI refers to Reinforcer Absent Interval. RPI refers to Reinforcer Present Interval.

### *5.3.4 Machine Learning*

Despite the current multilevel control algorithm, it is interesting to investigate how accurate a machine learning model can establish behavioral control through observation data. With the 200 graphs of 4 different kinds of control levels, confirmed by human experts, we applied machine learning for a supervised classification. We utilized the Scikit-learn library on the Jupyter notebook platform for machine learning training. We compared popular machine learning algorithms including decision tree, random forest, k nearest neighbors, support vector machine (SVM), logistic analysis, Naïve Bayes, and multilayer perceptron. The SVM had the best performance in classification accuracy among these algorithms. For feature extraction, we used the RAI and RPI values of each interval. Thus, every training sample is a 1-by-12 vector. For sessions that had less than 12 intervals, we interpolated the last available RPI/RAI to the non-existing values. Then, we split the 200 graphs to training and test in a ratio of 80:20 and ran a 5-fold cross validation. The model could assess a test set graph's control level at an accuracy of 77.0%, compared to the random guessing baseline of 25%/ While this result is not close to human expert BCBA yet, it shows the potential of PB.MI collecting larger scale digital and easy-to-process data for machine learning. More importantly, due to the nature of machine learning, the prediction accuracy will most likely improve significantly with a larger sample size. This sample size will increase as more behavior analysts use the program and share their de-identified data with the author. We expect this machine learning approach could assist behavior analysts even more so over time to assess the level of control demonstrated within an assessment or treatment session in the future.

## 5.4 Conclusion

In this chapter, we presented the design and validation of PB.MI, a smart application to assist visual analysis and establish multilevel brevity behavioral control. The PB.MI program provided an accurate representation in a test of randomly generated data and did not impact interpretations of control. In fact, reliability measures remained high across all tests, indicating that the PB.MI program could be used to collect data, generate a graph of data for clinical use, and supplement real-time interpretations of control. These outcomes may have practical benefits for clinicians. For example, the graphs can be copied and pasted into any progress reports, reducing some preparation time that may be required to manually create the graphs using professional graphing programs such as Microsoft Excel or Graphpad Prism. In addition, no technical skills are required for applying the structured criteria because it is automatically calculated by the PB.MI program.

It is important to point out that the functionality of the computerized program is not intended to replace visual analyses conducted by the practitioner. The computer calculated interpretations are there to aid visual analysis. Although within a 90% agreement range with human evaluators, the current version of the program cannot entirely account for more nuanced changes in trend, level, variability that can impact interpretations and clinical judgements. This suggests that indicators of control (e.g., binary, multilevel, PND) that are reliant on strict structured criteria are currently not as sensitive tools as the individual properly trained in visual analysis. Therefore, practitioners are advised to continue to familiarize themselves with the multilevel structured criteria before using the PB.MI program. Much like any structured criteria, they are suggested to be used as supplemental supports to visual analysis [20]. The feature from the PB.MI program that automatically interprets level of control may be especially supportive during ongoing visual analysis. When conducting a single-session IISCA, ongoing visual analysis occurs on the order of minutes and the clinician is expected to make informed decisions based on the changes in rate of problem behavior between rapidly alternating RAIs and RPIs. The real-time generation of the graph along with the ongoing interpretation of control could improve the efficiency of decision-making, allowing the clinician to confidently begin function-based treatments designed to reduce problem behavior.

From a technical standpoint, the PB.MI program can be adopted in its current version. However, we have yet to complete a social validation of the program in real-world use with other clinicians. To do so would require a clinician to conduct a single-session IISCA with a patient in need of functional assessment services for problem behavior while using the PB.MI program to collect data. Questionnaires can then be provided following the completion of the single-session IISCA regarding whether the use of the program is intuitive, feasible, and acceptable. In addition, any open-ended feedback could help to further develop features of the program and improve practicality and adoption among clinicians. In addition, future researchers may want to consider advancing user-friendly integration of other programs designed for use with multi-session IISCAs (e.g., PreMAC program) by creating decision-making algorithms that help the clinician determine what procedures to use and when (i.e., single-session IISCA or multi-session IISCA). These suggested programmatic changes could help guide clinicians to the appropriate assessment format given the contextually relevant variables of each individual situation they encounter in practice.

Inviting a larger population of clinicians to use the PB.MI program could also more directly impact the accuracy in generating interpretations of control. That is because machine learning models usually have significant increased performance with larger sample sizes [21]. Thus, future research is warranted with the PB.MI program distributed more openly among clinicians who readily conduct functional analyses with their patients. Machine learning holds promise for revolutionizing and automating how educators and clinicians make decisions [22]. Published questions about behavioral function data were used to build a neural network to inform conditions that should be present in a functional analysis. Machine learning was also used to support decision making of clinicians and researchers on interpreting inter-rater reliabilities of single-case designs using visual analysis [23]. Despite these successes, applications of machine learning in the function analysis remain under explored. The machine learning function of PB.MI is a novel attempt to assist behavioral analysts to interpret control levels with knowledge from existing data.

## References

[1] C. J. Whelan, G. P. Hanley, R. Landa, E. Sullivan, K. LaCroix, and R. Metras, "Randomized controlled trial of seminar-based training on accurate and general implementation of practical functional assessments," *J. Appl. Behav. Anal.*, 2021, doi: 10.1002/jaba.845.

[2] M. Torres-Viso, C. W. Strohmeier, and J. R. Zarcone, "FUNCTIONAL ANALYSIS AND TREATMENT OF PROBLEM BEHAVIOR RELATED TO MANDS FOR REARRANGEMENT," *J. Appl. Behav. Anal.*, vol. 51, pp. 158–165, 2018, doi: 10.1002/jaba.437.

[3] J. D. Lane and D. L. Gast, "Visual analysis in single-case experimental design studies: Brief review and guidelines," *Neuropsychol. Rehabil.*, vol. 24(3–4), pp. 445–463, 2014.

[4] C. E. Bullock, W. W. Fisher, and L. P. Hagopian, "Description and Validation of a Computerized Behavioral Data Program: 'bDataPro,'" *Behav. Anal.*, vol. 40, no. 1, pp. 275–285, 2017, doi: 10.1007/s40614-016-0079-0.

[5] Catalyst, "DataFinch." https://datafinch.com/aba-data-collection/.

[6] A. Sano, J. Hernandez, J. Deprey, M. Eckhardt, M. S. Goodwin, and R. W. Picard, "Multimodal annotation tool for challenging behaviors in people with autism spectrum disorders," *UbiComp'12 - Proc. 2012 ACM Conf. Ubiquitous Comput.*, pp. 737–740, 2012, doi: 10.1145/2370216.2370378.

[7] W. W. Fisher, M. And, K. Krieger, M. E. Kelley, and J. E. Lomas, "VISUAL AIDS AND STRUCTURED CRITERIA FOR IMPROVING VISUAL INSPECTION AND INTERPRETATION OF SINGLE-CASE DESIGNS," 2003.

[8] S. E. Danov and F. J. Symons, "A Survey Evaluation of the Reliability of Visual Inspection and Functional Analysis Graphs," *Behav. Modif.*, vol. 32, no. 6, Nov. 2008, doi: 10.1177/0145445508318606.

[9] H. S. Roane, W. W. Fisher, M. E. Kelley, J. L. Mevers, and K. J. Bouxsein, "USING MODIFIED VISUAL-INSPECTION CRITERIA TO INTERPRET FUNCTIONAL ANALYSIS OUTCOMES," doi: 10.1002/jaba.13.

[10] L. P. Hagopian, W. W. Fisher, R. H. Thompson, J. Owen-Deschryver, B. A. Iwata, and D. P. Wacker, "TOWARD THE DEVELOPMENT OF STRUCTURED CRITERIA FOR INTERPRETATION OF FUNCTIONAL ANALYSIS DATA," 1997.

[11] J. Jessel, R. Metras, G. P. Hanley, C. Jessel, and E. T. Ingvarsson, "Evaluating the boundaries of analytic efficiency and control: A consecutive controlled case series of 26 functional analyses," *J. Appl. Behav. Anal.*, 2020, doi: 10.1002/jaba.544.

[12] J. Jessel *et al.*, "Does analysis brevity result in loss of control? A consecutive case series of 26 single-session interview-informed synthesized contingency analyses," 2019, doi: 10.1002/bin.1695.

[13] J. D. Slaton, G. P. Hanley, and K. J. Raftery, “Interview-informed functional analyses: A comparison of synthesized and isolated components,” *J. Appl. Behav. Anal.*, 2017, doi: 10.1002/jaba.384.
[14] Unity3D, “Unity Editor.” https://unity.com/.
[15] D. Spiker, L. J. Lotspeich, S. Dimiceli, R. M. Myers, and N. Risch, “Behavioral Phenotypic Variation in Autism Multiplex Families: Evidence for a Continuous Severity Gradient,” *Am. J. Med. Genet. (Neuropsychiatric Genet.*, vol. 114, pp. 129–136, 2002, doi: 10.1002/ajmg.10188.
[16] J. Jessel, G. P. Hanley, M. Ghaemmaghami, and R. Metras, “An evaluation of the single-session interview-informed synthesized contingency analysis,” *Behav. Interv.*, vol. 34, no. 1, pp. 62–78, 2019, doi: 10.1002/bin.1650.
[17] T. E. Scruggs, M. A. Mastropieri, and G. Casto, “The Quantitative Synthesis of Single-Subject Research: Methodology and Validation.”
[18] Z. K. Zheng *et al.*, “A Predictive Model to Alert Caregivers of Problem Behaviors for Children with ASD ( PreMAC ),” 2020.
[19] M. L. McHugh, “Lessons in biostatistics interrater reliability : the kappa statistic,” *Biochem. Medica*, vol. 22, no. 3, pp. 276–282, 2012, [Online]. Available: https://hrcak.srce.hr/89395.
[20] A. Dowdy, J. Jessel, V. Saini, and P. Corey, “Structured Visual Analysis of Single Case Experimental Design Data: Developments and Technological Advancements,” *J. Appl. Behav. Anal.*, 2021.
[21] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
[22] S. Turgeon and M. J. Lanovaz, “Tutorial: Applying Machine Learning in Behavioral Research,” 2020, doi: 10.1007/s40614-020-00270-y.
[23] M. J. Lanovaz, A. R. Giannakakos, and O. Destras, “Machine Learning to Analyze Single-Case Data: A Proof of Concept,” *Perspect. Behav. Sci.*, vol. 43, no. 1, pp. 21–38, 2020, doi: 10.1007/s40614-020-00244-0.

# Chapter 6: Real-time Prediction of Precursors of Problem Behaviors

## 6.1 Abstract

Affective computing has demonstrated the feasibility of detecting problem behaviors from multimodal behavioral responses. As shown in Chapter 3, the performance of the machine learning model was promising. Timely intervention can significantly lower the risk of behavioral escalations and the frequency of problem behaviors. Thus, the machine learning prediction of problem behaviors has the potential to offer an alternative for behavior monitoring that is low-cost, consistent and precise. However, there is not yet a proof-of-concept study to apply real-time prediction of problem behaviors in ASD. It remains unknown how accurate and robust an offline machine learning model could predict problem behaviors in real-time. The field also lacks a real-time wearable multimodal signal processing and feature extraction system for online problem behavior detections. Therefore, in this chapter, we are going to present the design of our real-time wearable precursor prediction system, the pilot study for data collection and system validation, as well as the machine learning model performance data analysis. The real-time prediction system could collect signals, extract the features and make classifications with a low latency. The data analysis compared the real-time predictions with human expert observations, which are the state-of-the-art ground truth for behavior annotations. The real-time predictions reached an average of 77.42% accuracy and it demonstrated the potential for automatic problem behavior prediction through real-time multimodal sensing and machine learning. We also performed a machine learning experiment to investigate the potential reasons for the decrease accuracy in real-time predictions, as compared to the offline machine learning results. The results indicated that with more data from different days, the real-time prediction is likely to increase prediction accuracies and will be more generalized on different behavioral patterns.

## 6.2 Motivation

There has been a variety of research utilizing behavioral response sensing to detect or predict problem behaviors in children with ASD. Peripheral physiological responses such as heart rate (HR) and electrodermal activity (EDA) are indicators for stress and frustration [1][2]. Such physiological data were collected and used to analyze imminent aggression [3]. With the same dataset, significantly improved prediction accuracies over varying prediction windows were achieved by a support vector machine model [4]. The results demonstrated that the individualized and group models were able to predict the onset of aggression one minute before occurrence with accuracies over 82%. Body motion also contains rich information about problem behaviors, and accelerometer data was used to recognize stereotypical hand flapping and body rocking behaviors, which may occur in some children with ASD [5]. In another study, researchers found out movement data along with annotated behaviors could build a machine learning model to predict episodes of self-injurious behaviors [6]. Multimodality is defined by the presence of more than one data channel, and multimodal data improved performance with more data modalities and potential correlations among them [7]. In Chapter 3, we presented the predictive multimodal framework to alert caregivers of problem behaviors for children with ASD (PreMAC), that utilizes motion, physiological, social orientations and behavioral observations to predict precursors of problem behaviors [8]. The existing literature has demonstrated the feasibility of using multimodal behavioral responses to detect or predict problem behaviors in children with ASD.

Existing literature also shows the efficiency of on-time intervention reducing the rate of problem behaviors. Jessel et al. applied the interview informed synthesized contingency analysis (IISCA) to achieve socially significantly reductions in problem behaviors [9]. A 90% or greater reduction in problem behaviors was obtained for every participant by the end of the treatment evaluation. IISCA is the state-of-the-art functional analysis that utilizes precursors of problem behaviors that is safer and faster [10]. IISCA was also used to treat elopement in children with ASD with good results [11]. In the questionnaire of this study, parents stated that they would be confident to apply the same strategy they have seen in the clinic to address elopement at home. Thus, through the advancement of functional analysis, it is very likely to reduce the rate and severity of problem behaviors, if precursors of problem behaviors can be observed in time. However, due to the heavy caregiving workloads on parents of autistic children [12], they do not have enough bandwidth to notice precursors of problem behaviors continuously. Prediction of precursors of problem behaviors in real-time has the potential to allow parents timings to intervene in time to reduce the impact of such behaviors.

Real-time artificial intelligence methodologies have become pervasive in edge-based affective computing for its low latency, portability and data security [13]. Edge-based is a distributed computing paradigm that has computation and data storage closer to the source of data, in comparison with cloud-based computing [14]. BodyEdge is a software structure to assess worker and athlete stress in real-time through ECG. The stress level can be estimated within a latency of 0.152 seconds. Ragev et al. trained deep learning models with three public datasets on stress estimation and ran the classifications on a Raspberry Pi 3 Model B [15]. The model could estimate emotions like arousal, dominance and valance at over 60% of accuracy. Guo et al. designed a light weight convolutional neural networks model to detect facial expressions on mobile devices [16]. The model could recognize emotions including neutral, happy, sad, surprise, and fear for 31 frames per second on an iPhone 5s but the classification accuracy for most emotions are around 60%. The existing research on real-time affective computing indicate the feasibility to detect and inform different emotions at high updating frequencies. However, many studies showed significantly decreased accuracies, compared to the offline machine learning. Overfitting in supervised learning is likely to occur in affective computing. That is because behavioral responses and observations are often difficult and expensive to get [17] and thus training data can be rare and only capture partial behaviors.

In this chapter, we present the design of a real-time prediction system of imminent precursors of problem behaviors. Building on the PreMAC, we designed and developed network sockets and communications to extract and synchronize features from real-time data. The system then makes prediction in real-time on whether an imminent precursor is going to take place. To validate the feasibility and precision of the real-time predictions, we conducted a pilot study with autistic children that display some problem behaviors. The pilot study has two sessions, where we collect the training data during the first session to build the individualized model and validate the model on the second session. **We hypothesize that** 1) a machine learning model from the collected training data is able to predict precursors of problem behaviors in real-time with good accuracies; and 2) with more data, the individualized models are going to be more generalized and robust for problem behavior detection.

### 6.3 Design of the Real-time Problem Behavior Prediction System

The proposed system will have three main system components: a data collection module, a signal processing module, and an online classification module. The system architecture is shown in Fig. 1. The system collects behavioral responses from children, then processes the signals and extracts the features. Based on the real-time features, the system makes online predictions of whether an imminent precursor is going to happen.

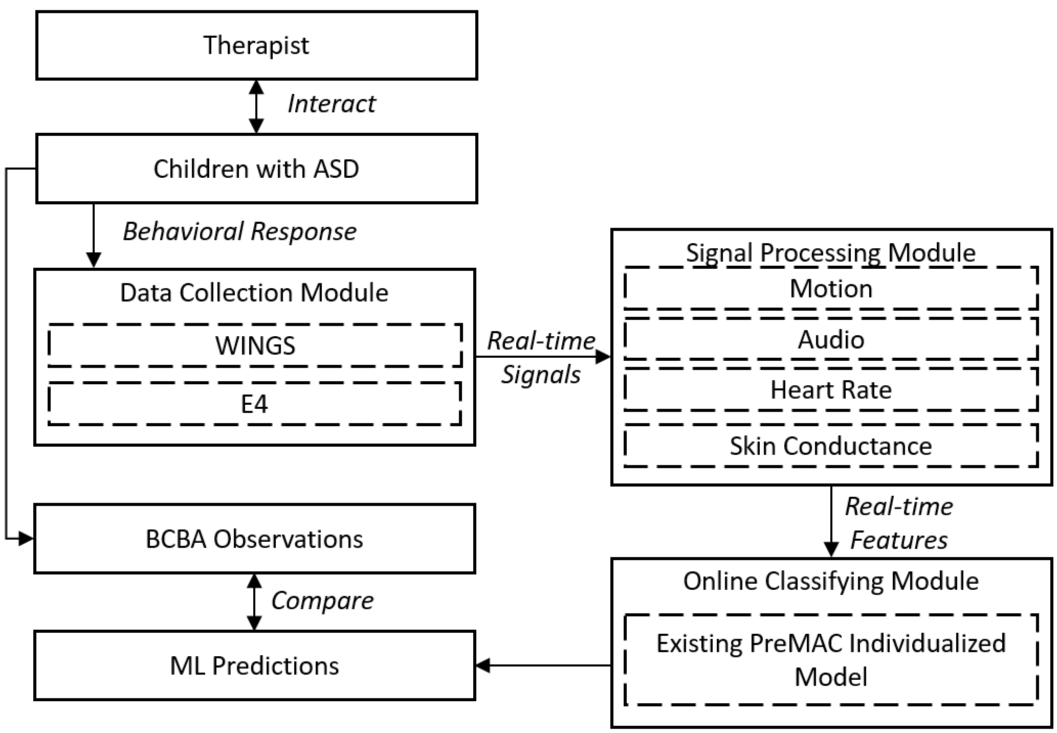



**Fig. 1.** System Architecture

### *5.2.1 Data Collection Module*

The data collection module was based on the Multimodal Wearable Data Capture System (M2P3), as shown in chapter 3. From the previous study and machine learning results, we have observed that the WINGS and E4 wristband are the most informative data modalities. Computer vision through Kinect had many missing data points because the participants were actively moving in the space. Therefore, we utilized the WINGS sensor and E4 wristband in the data collection module for the real-time system. Audio channel is an  potential data modality to contain rich predictive power for affective states [18]. Vocal precursors are also very common among the frequent precursors of autistic children [19]. Therefore, we integrated a Lapel microphone in the WINGS sensor to collect some preliminary data. The communication scheme for data collection is shown in Fig.2, where the dotted lines stand for wireless communication and the rigid lines stand for wired communication. The WINGS sensor integrated inertial measurements units (IMU) to measure the roll, pitch and yaw angles of the subject, and it sent the signals to the Arduino connected to the workstation through a wireless radio frequency (RF) communication. The Arduino then passed on the signals to the data collection controller through serial communication. The E4 wristband has a commercial E4 streaming server [20] and we wrote a client in the data collection controller to have a TCP socket [21]. The Lapel mic transmitter has a RF communication with the receiver, which sends the signals to the data collection controller.

### *5.2.2 Signal Processing Module*

The signal processing module is a C# script in Unity, and it extracts and synchronizes the real-time features. For physiological data, we recorded both photoplethysmographic (PPG) and EDA data in this study. PPG is used as a measure of blood volume pressure (BVP) and to compute heart rate (HR) by identification of local minima and inter-beat interval (IBI), as shown in Equation (1). Blood pressure and HR variability correlate with defensive reactions, pleasantness of stimuli and basic emotions.

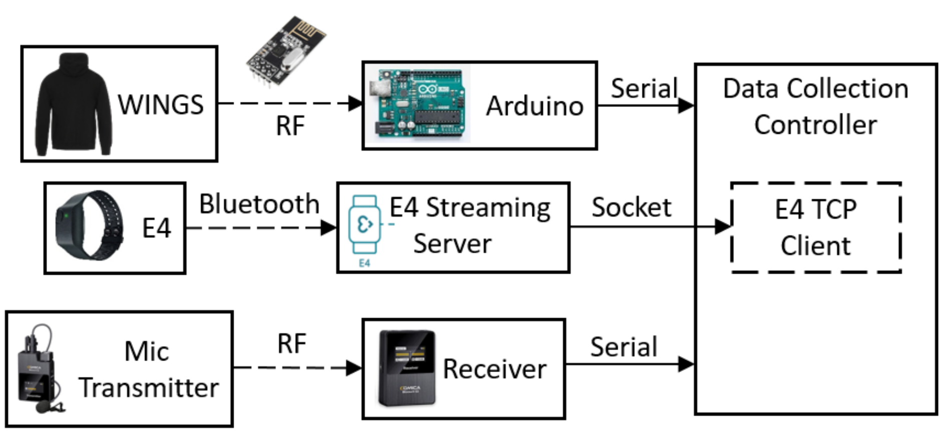



**Fig. 2.** Data Collection Module Communication

$$HR_I = \frac{60}{\left| t_{peak(I+1)} - t_{peak(I)} \right|} \tag{1}$$

From the 3 components of the accelerometer readings, $accl_x$, $accl_y$, and $accl_z$, and 3 components of the magnetometer readings, $mag_x$, $mag_y$, and $mag_z$, we can compute the roll, pitch and yaw angles ($\theta$, $\psi$, $\phi$) of the torso and limbs using Equations (2), (3), and (4) as shown below. The roll and pitch angles are computed by the IMU orientations with respect to the gravitational direction. The yaw angle is computed by the relative IMU orientations with respect to the earth's magnetic field.

$$\theta = \tan^{-1}\left( \frac{accl_y}{\sqrt{accl_y^2 + accl_z^2}} \right) \tag{2}$$

$$\psi = \tan^{-1}\left( \frac{accl_x}{\sqrt{accl_y^2 + accl_z^2}} \right) \tag{3}$$

$$\phi = \tan^{-1}\left( \frac{mag_z s\psi - mag_y c\theta}{mag_x c\theta + mag_y s\theta s\psi + mag_z c\psi s\theta} \right) \tag{4}$$

Then the signal processing module fuses the data. We used the feature level fusion due to its advantage in investigating the correlation between various multimodal features at an early stage [22]. The major technical challenge for feature level fusion is the signal synchronization because the updating frequencies of different data modalities usually vary. For our system, the updating frequencies for motion, EDA and HR are about 15Hz, 4Hz, and 1Hz, respectively. Our feature synchronization approach is shown in Fig. 3. The motion feature is the data modality with the highest updating frequency. Thus, every time the motion feature gets updated, the synchronization algorithm fetches the most recent value of other features and form a complete data entry. The time stamp of the complete data entry was computed as the timing of the motion feature updated. A complete data sample of our real-time prediction system is a 1-by-18 vector. Once a complete data entry is formed, it would be forwarded to the online classification module.

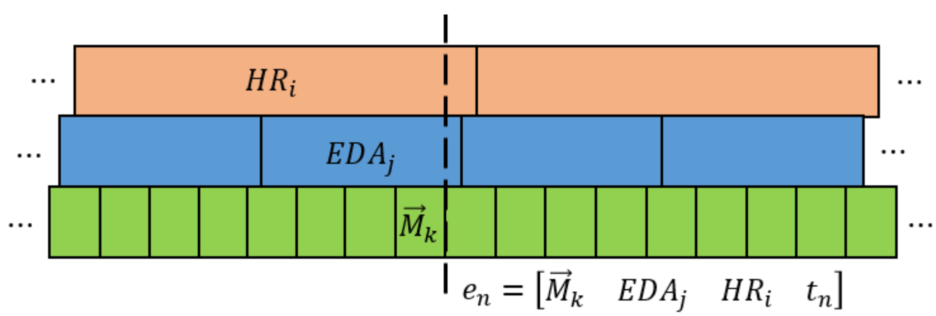



**Fig. 3.** Feature Synchronization

*5.2.3 Online Classification Module*

The online classification module takes real-time data samples from the signal-processing module and make classifications using the previously trained model. The online classification module is a Python script and it communicates with the signal-processing module through a local network socket. Through our previous  collected data, we had already trained individualized models for each child, and we serialized these models through the pickle library [23]. Upon real-time predictions, we de-serialized individualized models for the corresponding child and made predictions. We performed a latency test for our real-time prediction system, with a desktop with very basic components [24]. It has an Intel i3 processor (3.7GHz), Intel UHD 630 Graphics card and an 8GB RAM. We ran the predictions continuously and the system was able to make 12979 predictions in 1800 seconds, which equated 7.21Hz per prediction. In other words, the real-time prediction system would spend 0.138 seconds from catching a real-world behavior to a real-time classification. The latency can be further eliminated by upgrading the features of the workstation running the real-time prediction but it was already adequate for problem behavior detection.

## 6.4 Pilot Study

To validate the accuracy and feasibility of the real-time prediction system, we conducted a pilot study with autistic children with problem behaviors. We recruited four children with ASD (age mean: 4.56, std = 2.71; 4 male and 0 female). The experimental setup is shown in Fig. 4. The study was approved by the Institutional Review Board (IRB) at Vanderbilt University.

The children interact with our Board Certified Behavior Analyst (BCBA) therapist in the experimental space. We integrated the state-of-the-art functional analysis IISCA in our experimental protocol so that we can efficiently evoke precursors and then reinforce children's behaviors. With IISCA, we were able to have shorter and safer sessions. The average time for our data collection was only 45.92 minutes, compared to 4.35 hours in a related study [4]. The average precursor count for each visit was 26.87, and no children displayed problem behaviors. Therefore, we were efficiently reinforcing the behaviors of the children, with frequent episodes of precursors but they never escalated to problem behaviors, which may be harmful for both the children and the therapist. There were 5 sessions in each visit, consisting of 2 control sessions and 3 test sessions and aligned as CTCTT. In control sessions, the therapist does not evoke problem behaviors from children while the therapist switches between evoking and reinforcing in test sessions. The purpose of this design is to have baseline data for our data collection and allow necessary breaks for the children and the therapist [25]. During these breaks, the therapist could also utilize the summary page of our behavioral data collection integrator to make needed modifications to the following sessions.

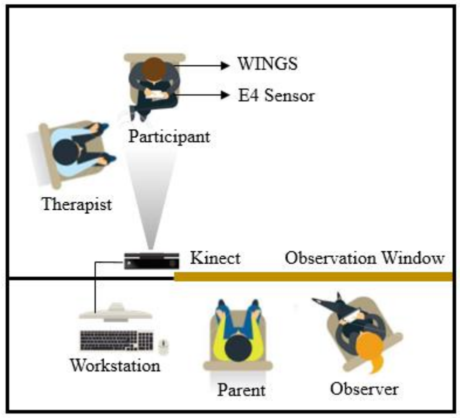



**Fig. 4.** Experimental Setup

We scheduled two visits for every child in this study. For the first visit, we collected their multimodal responses with behavioral observations, so we could train the machine learning model for individualized profiles. For the second visit, we brought the kid back and applied the learnt model to have real-time predictions. The observer collects the behavioral observations during the second visit to have ground truth for the real-time predictions. Four out of 5 children finished both sessions of our pilot study and they tolerated all the wearable sensors. One child withdrew from the study on the first visit because they did not enjoy the interaction and wanted to quit.

## 6.5 Data Analysis

From our pilot study, we have collected multimodal responses and behavioral observations for machine learning. Then we collected behavioral observations and the history of real-time predictions for the validation of real-time prediction accuracy. For data collection, there were an average of 20298 samples for each child. And for real-time prediction, there were an average of 18976 decisions made. First, we used the Jupyter Notebook platform [26] and scikit-learn library [27] to train the individualized models. The prediction accuracies, precision and recall of each child are shown in Table 1. The offline prediction accuracies remained very high, on an average of 77.42%, which is consistent from the results in PreMAC presented in Chapter 3.

**Table 1.** Offline Machine Learning Results

| CHILD | ACCURACY | PRECISION | RECALL |
|---|---|---|---|
| **1** | 98.53% | 0.996 | 0.961 |
| **2** | 99.6% | 0.991 | 0.990 |
| **3** | 96.98% | 0.983 | 0.910 |
| **4** | 98.97% | 0.998 | 0.959 |

It is more important to validate the accuracy of real-time prediction. From human expert BCBA observations, we were able to collect ground truth for whether an imminent precursor was going to happen. We compared the real-time machine learning results to the observer results and analyzed the percentage of them agreeing with each other. If the real-time system had the same prediction as the observer, we would count that as correct prediction. We also computed the recalls and precisions for the real-time predictions and these results are shown in Table 2. From these results, we can see that the prediction accuracies of real-time predictions have decreased, compared to the offline prediction. But an average prediction accuracy of 77.42% is still promising for real-time predictions. This decrease in prediction accuracy may be due to the overfitting of the initial models. Children with ASD display repetitive behaviors [28] and within one hour of data collection, it was unlikely for them to display a comprehensive range of their precursors. And for another session on a different day, significantly different behavior patterns may be present. Another observation is that the real-time models have rather stable recalls but significantly decreased precisions, as compared to the offline machine learning models. This result indicates that the real-time predictions had more false negatives than false positives. The reason for this maybe that our training data was slightly unbalanced and the positive samples were about 20%-40% of our dataset. We did not balance the classes because we had limited amount of total samples and the prediction accuracy drops if we discard portions of our data.

| **CHILD** | **ACCURACY** | **PRECISION** | **RECALL** |
|---|---|---|---|
| **1** | 86.52% | 0.989 | 0.622 |
| **2** | 79.83% | 0.951 | 0.807 |
| **3** | 70.76% | 0.896 | 0.697 |
| **4** | 72.55% | 0.953 | 0.686 |

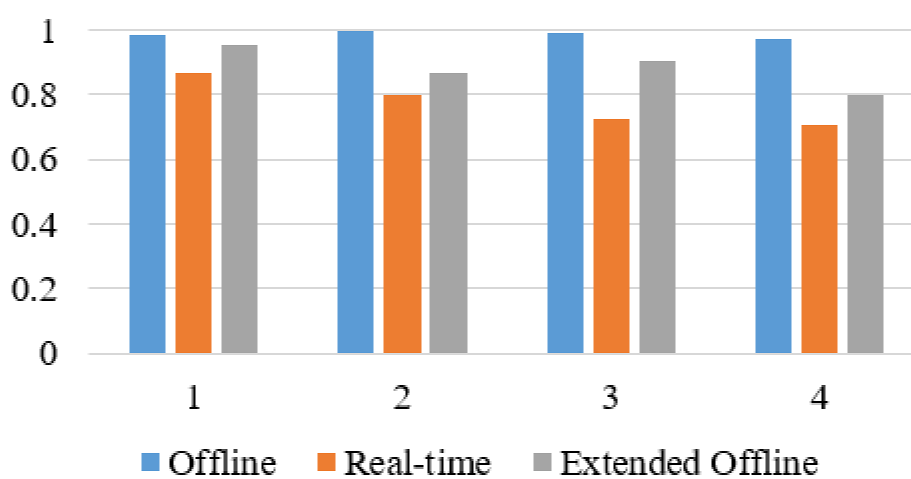



**Fig. 5.** Machine Learning Experiment Results

Therefore, one of the main limitations for real-time prediction may be that the individualized model could overfit onto specific behavioral patterns that already differed on a separate visit. To investigate how this limitation could be addressed in the future, we performed a machine learning experiment. We trained an extended offline model that combined the data for both visits. For the extended offline model, we used the whole data from the first visit and 80% of the second visit data as training, and the rest 20% of the second visit data as test. In this way, we could investigate whether obtaining some information from the second visit would contribute to the performance and whether behavior patterns in those visits differ. As Fig. 5 shows, the blue, orange and grey bars stand for the prediction accuracies of offline model, real-time predictions, and extended offline model, respectively. As shown in the figure, there are decrease in prediction accuracies from offline machine learning to real-time predictions. However, for the extended offline model, the prediction accuracies increased but did not exceed the performance of initial offline models. This result indicates that with more data from different visits, the combined machine learning model may reach a better prediction accuracy but more importantly, have better robustness and fit the general behavioral patterns of that child.

### 6.6 Conclusion and Future Work

In this chapter, we presented the design of a real-time prediction system for problem behaviors in children with ASD. The system captures real-time multimodal data, extracts features and make online decisions of whether an imminent precursor is going to happen. To validate the feasibility and accuracy of the presented system, we conducted a pilot study with 4 autistic children. Each child had two visits to first collect training data, and then validate the accuracy of real-time predictions. We compared the machine learning real-time predictions with the BCBA human observations and found promising accuracies. This result demonstrates the potential of using intelligent multimodal data fusion system with real-time machine learning to detect the precursors of problem behaviors in an automatic and consistent way. Through the investigation of decreased prediction accuracies between offline and online predictive models, we found results that support the theory that the machine learning models for problem behaviors may overfit to specific behaviors in data collection. But this problem is very likely to be solved by having a large data sample size with longitudinal visits.

There are some limitations of the current work. First, the data modalities are limited to motion and peripheral physiological data. Although they have been proven to have strong correlations with problem behaviors, some of the problem behaviors are not likely to be detected by them, such as screaming. Second, the real-time prediction accuracies are limited and may need more data. However, data collection for behavioral responses are expensive because extensive involvement of the BCBAs. The pilot study participant pool was small, so that the results are not generalizable at this stage. Therefore, there are many important future work for this work. First, we are going to incorporate audio channel into the multimodal data collection and processing system. This would significantly augment the system's ability to catch more types of precursors and hopefully increase the robustness of the real-time predictions. Second, using semi-supervised learning [30], we may be able to collect a significantly larger dataset on behavioral response without the presence of BCBAs. Therefore, we can collect much more data without the large labor cost and tight schedules while the unlabeled data is still likely to increase the accuracies of the real-time predictions. Finally, with more time and resources, a longitudinal study of real-time prediction data collection and validation would further prove this concept and have more generalized results.

## References


[1] C. Setz, B. Arnrich, J. Schumm, R. La Marca, G. Tröster, and U. Ehlert, "Discriminating stress from cognitive load using a wearable eda device," *IEEE Trans. Inf. Technol. Biomed.*, 2010, doi: 10.1109/TITB.2009.2036164.

[2] J. Taelman, S. Vandeput, A. Spaepen, and S. Van Huffel, "Influence of Mental Stress on Heart Rate and Heart Rate Variability," *IFMBE Proc.*, vol. 22, pp. 1366–1369, 2009, doi: 10.1007/978-3-540-89208-3_324.

[3] M. S. Goodwin *et al.*, "Predicting imminent aggression onset in minimally-verbal youth with autism spectrum disorder using preceding physiological signals," *ACM Int. Conf. Proceeding Ser.*, pp. 201–207, 2018, doi: 10.1145/3240925.3240980.

[4] T. Imbiriba, D. C. Cumpanasoiu, J. Heathers, S. Ioannidis, D. Erdomuş, and M. S. Goodwin, "Biosensor prediction of aggression in youth with autism using kernel-based methods," *ACM Int. Conf. Proceeding Ser.*, pp. 88–93, 2020, doi: 10.1145/3389189.3389199.

[5] F. Albinali, M. S. Goodwin, and S. Intille, "Detecting stereotypical motor movements in the classroom using accelerometry and pattern recognition algorithms," *Pervasive Mob. Comput.*, vol. 8, no. 1, pp. 103–114, 2012, doi: 10.1016/j.pmcj.2011.04.006.

[6] K. D. Cantin-Garside, Z. Kong, S. W. White, L. Antezana, S. Kim, and M. A. Nussbaum, "Detecting and Classifying Self-injurious Behavior in Autism Spectrum Disorder Using Machine Learning Techniques," *J. Autism Dev. Disord.*, 2020, doi: 10.1007/s10803-020-04463-x.

[7] S. Poria, E. Cambria, R. Bajpai, and A. Hussain, "A review of affective computing: From unimodal analysis to multimodal fusion," *Inf. Fusion*, vol. 37, pp. 98–125, Sep. 2017, doi: 10.1016/j.inffus.2017.02.003.

[8] Z. K. Zheng *et al.*, "A Predictive Model to Alert Caregivers of Problem Behaviors for Children with ASD ( PreMAC )," 2020.

[9] J. Jessel, E. T. Ingvarsson, R. Metras, H. Kirk, and R. Whipple, "Achieving socially significant reductions in problem behavior following the interview-informed synthesized contingency analysis: A summary of 25 outpatient applications," *J. Appl. Behav. Anal.*, vol. 51, no. 1, pp. 130–157, 2018, doi: 10.1002/jaba.436.

[10] H. Heath and R. G. Smith, "Precursor behavior and functional analysis: A brief review," *J. Appl. Behav. Anal.*, vol. 52, no. 3, pp. 804–810, 2019, doi: 10.1002/jaba.571.

[11] J. Jessel, E. T. Ingvarsson, R. Metras, R. Whipple, H. Kirk, and L. Solsbery, "Treatment of elopement following a latency-based interview-informed, synthesized contingency analysis," *Behav. Interv.*, vol. 33, no. 3, pp. 271–283, 2018, doi: 10.1002/bin.1525.

[12] C. N. Marsack-Topolewski and H. L. Church, "Impact of caregiver burden on quality of life for parents of adult children with autism spectrum disorder," *Am. J. Intellect. Dev. Disabil.*, vol. 124, no. 2, pp. 145–156, 2019, doi: 10.1352/1944-7558-124.2.145.

[13] R. K. Nath, H. Thapliyal, A. Caban-Holt, and S. P. Mohanty, "Machine Learning Based Solutions for Real-Time Stress Monitoring," *IEEE Consum. Electron. Mag.*, vol. 9, no. 5, pp. 34–41, 2020, doi: 10.1109/MCE.2020.2993427.

[14] P. Dong *et al.*, "Edge Computing Based Healthcare Systems: Enabling Decentralized Health Monitoring in Internet of Medical Things," *IEEE Netw.*, vol. 34, no. 5, pp. 254–261, Sep. 2020, doi: 10.1109/MNET.011.1900636.

[15] A. Ragav, N. H. Krishna, N. Narayanan, K. Thelly, and V. Vijayaraghavan, “Scalable deep learning for stress and affect detection on resource-constrained devices,” *Proc. - 18th IEEE Int. Conf. Mach. Learn. Appl. ICMLA 2019*, pp. 1585–1592, 2019, doi: 10.1109/ICMLA.2019.00261.
[16] Y. Guo, Y. Xia, J. Wang, H. Yu, and R. C. Chen, “Real-time facial affective computing on mobile devices,” *Sensors (Switzerland)*, vol. 20, no. 3, pp. 1–15, 2020, doi: 10.3390/s20030870.
[17] J. H. Tao, J. Huang, Y. Li, Z. Lian, and M. Y. Niu, “Semi-supervised Ladder Networks for Speech Emotion Recognition,” *Int. J. Autom. Comput.*, 2019, doi: 10.1007/s11633-019-1175-x.
[18] N. J. Shoumy, L. M. Ang, K. P. Seng, D. M. M. Rahaman, and T. Zia, “Multimodal big data affective analytics: A comprehensive survey using text, audio, visual and physiological signals,” *J. Netw. Comput. Appl.*, vol. 149, no. July 2019, p. 102447, 2020, doi: 10.1016/j.jnca.2019.102447.
[19] J. B. Plavnick and F. A. Vitale, “A Comparison of Vocal Mand Training Strategies for Children With Autism Spectrum Disorders,” *J. Posit. Behav. Interv.*, vol. 18, no. 1, pp. 52–62, 2016, doi: 10.1177/1098300714548800.
[20] Empatica, “E4 Streaming Server.” https://developer.empatica.com/windows-streaming-server.html.
[21] B. M. Shafiei, F. Idiranmanesh, and F. Iranmanesh, “Socket programming,” *Adv. Environ. Biol.*, 2012, doi: 10.1016/b978-155860826-9/50016-1.
[22] M. Haghighat, M. Abdel-Mottaleb, and W. Alhalabi, “Discriminant Correlation Analysis: Real-Time Feature Level Fusion for Multimodal Biometric Recognition,” *IEEE Trans. Inf. Forensics Secur.*, vol. 11, no. 9, pp. 1984–1996, 2016, doi: 10.1109/TIFS.2016.2569061.
[23] The Python Software Foundation, “pickle — Python object serialization — Python 3.7.3 documentation,” *09 June 2019*, 2019. .
[24] Dell, “XPS Desktop.” https://www.dell.com/en-us/shop/desktop-computers/xps-desktop/spd/xps-8940-desktop/xd8940msr10h.
[25] A. V. Balakrishnan, *Applied Functional Analysis*. Springer Science & Business Media, 2012.
[26] T. Kluyver *et al.*, “Jupyter Notebooks—a publishing format for reproducible computational workflows,” 2016, doi: 10.3233/978-1-61499-649-1-87.
[27] F. Pedregosa *et al.*, “Scikit-learn: Machine learning in Python,” *J. Mach. Learn. Res.*, 2011.
[28] A. S. Weitlauf, N. Sathe, M. L. McPheeters, and Z. E. Warren, “Interventions targeting sensory challenges in autism spectrum disorder: A systematic review,” *Pediatrics*. 2017, doi: 10.1542/peds.2017-0347.
[29] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
[30] Z. Zhang, F. Ringeval, B. Dong, E. Coutinho, E. Marchi, and B. Schuller, “Enhanced semi-supervised learning for multimodal emotion recognition,” 2016, doi: 10.1109/ICASSP.2016.7472666.

# Chapter 7: Contributions and Future Work

## 7.1 Overall Contributions

In the past decade, innovations such as sensing and machine learning have been enhancing interventions and supports for children with ASD. This dissertation describes my research on the design, analysis and validation of multimodal behavioral interventions. My research is expected to contribute to the science and engineering of technology-assisted ASD intervention. In this chapter, we are going to introduce the potential technical contributions and societal contributions of this dissertation.

## 7.2 Technical Contributions

The first set of technical contributions is in the design and development of a multisensory stimulation and data capture system (MADCAP). Research has shown that the different response patterns on audio-visual stimulus in infants can be a predictor of later diagnosis of ASD. However, the tactile responses remains under-researched in this context. The MADCAP system aims to investigate audio, visual, and tactile responses of high ASD risk infants with an aim to decipher patterns of responses that could be indicators of later diagnosis of ASD. In this work we present: 1) the design and development of a synchronized system that delivers audio, visual and tactile stimuli and collects multisensory responses from infants; 2) the design of a novel tactile stimulator that delivers affective touch with controlled speed and pressure; 3) the design of a soft robotic tendon driven sleeve, Soft-Brush, to improve experimental tolerance; and 4) a new machine learning approach on the collected data for response pattern differences to classify ASD/TD group.

The second set of technical contributions is in the design and development of a predictive multimodal framework to alert caregivers of problem behaviors for children with ASD (PreMAC). The existing studies have shown the physiological responses of children with ASD during problem behavior or self-injurious behavior episodes can be used to predict problem behaviors. However, most of the existing studies focus on unimodal sensing framework and demonstrated limited prediction accuracies. More importantly, all the existing work focus on offline analysis instead of real-time predictions. The PreMAC framework and the proposed work address these limitations and will have the following contributions: 1) the design of an integrated sensor framework to collect multimodal data with behavioral states from children with ASD and problem behaviors; 2) the design of the customized wearable intelligent non-invasive gesture sensor to capture upper body gestures; 3) a machine learning approach to classify whether an imminent precursor is going to happen, with significantly increased prediction accuracies than the state-of-the-art models; 4) analysis of feature importance and sensor contributions of the multimodal data; and 5) real-time prediction and its validation of problem behaviors in children with ASD.

The third set of technical contributions is in the design and development of interactive augmented reality coaching system for toothbrushing skills (CheerBrush). The existing VR based systems have shown better motivation for interventions and provide a low-cost and accessible alternative for ASD supports. However, application of VR and AR for training good personal hygiene habits remain under-explored. The existing systems on training personal hygiene habits are mostly open loop and lack immersive teaching environments. The CheerBrush system aims to address these limitations and will have the following contributions: 1) the design of a novel AR-based interaction coaching environment for toothbrushing skills; 2) integration of virtual avatars, audio and visual guidance and feedback; 3) the development of a mechatronic toothbrush to assess the toothbrushing movements before and after the coaching sessions; 4) physiological estimation of stress levels of children during AR coaching sessions.

## 7.3 Societal Contributions

Besides the technical contributions, this work will also contribute towards the science of ASD intervention. All completed work have been validated by feasibility and pilot study involving human subjects.

For the MADCAP work, we have designed a novel tactile stimulator, Soft-Brush, that can deliver affective touch in children with neurodevelopmental disorders. Soft-Brush provides the field with an automatic tool to trigger affective touch for this population. The Soft-Brush can also be applied on neuro-typical populations as well for somatosensory investigations. The machine learning models of MADCAP showed that the affective touch features contributed to the overall developmental risk detection and it along showed some potential for early screening of ASD developmental risk. Therefore, this work may pave the road for multisensory response analysis for early detection of ASD. The present of behavioral responses in early months of infancy may provide parents and clinicians valuable time to intervene for better long-term neural outcomes.

The PreMAC work has demonstrated a wearable motion sensor for this population with sensory differences. The presented WINGS sensor displayed high tolerability through experiments. Our machine learning models improved the state-of-the-art prediction accuracies, demonstrating the feasibility of utilizing multimodal data analysis on problem behavior prediction. The real-time PreMAC is the first work to develop and validate real-time prediction of problem behaviors. The results demonstrated promising prediction accuracies with an automatic, consistent, and low-cost problem behavior prediction approach.

The brevity app provides behavioral analysts a convenient and open-source program to establish brevity behavioral control levels in assessments. It will significantly decrease the amount of time needed for functional analysis assessments. It can also serve as a teaching tool for behavioral analyst trainings.

With the permission of users, the program can collect data from different assessments for machine learning. Thus, the control establishment criterion will be further improved.

For the CheerBrush project, we developed the first AR-based coaching system for personal hygiene skills. The system demonstrated positive improvements in toothbrushing movements, which is likely to transfer in real-life. In conclusion, the completed work of this dissertation provides important insights into how multisensory data collection and machine learning can assist and enhance the ASD intervention and research.

## 7.4 Future Work

One of the future directions of this research is to collect behavioral responses of autistic children in remote settings such as schools and clinics. These data may contain more predictive power related to precursors of problem behaviors. Thus, fully portable data collection and streaming system needs to be developed. As discovered by this research, portable sensing channels such as motion and peripheral physiological data can result in very promising prediction accuracies in real-time predictions. We hypothesize that through selected channels of the existing PreMAC system, we can collect data in remote settings that better describes behavioral responses in life-like situations. The presence of board certified behavioral analysts (BCBA) during remote data collections is not necessary. That would allow data collections over significantly longer periods of time because the major cost of data collection comes from the labor expense of BCBAs. In that way, we can significantly increase the sample size of our data collection by a large amount of unlabeled data. Semi-supervised learning is a machine learning approach that combines a small amount of labeled data with a large amount of unlabeled data during training. Unlabeled behavioral responses, when used in conjunction with a small amount of labeled data, has the potential to produce considerable improvement in learning accuracy. Furthermore, the audio channel is another obvious behavioral response and we will include vocal information in prediction as well. We are also going to have a longitudinal study on children with ASD on real-time problem behavior predictions so that our results can be more generalized.

MADCAP system has demonstrated the potential of investigating autism development risk by multimodal responses in toddlers. By a future longitudinal study on infants younger than 12 months, we can further validate the efficiency of predicting ASD in early infancy with behavioral responses and machine learning. We have presented the MRI-compatible Soft-Brush in chapter 2, which delivers affective touch in MRI environments. One potential future work would be investigating behavioral responses with MRI-compatible Soft-Brush in fMRIs, not only to validate the feasibility and tolerance of Soft-Brush in a MRI environment, but also to investigate fMRI responses of affective touch, that may be related to ASD development risk.

# Chapter 8: Ethical Aspects of This Dissertation

## 8.1 Ethics about Affective Computing and Wearable Sensing

Emerging artificial intelligence (AI) and wearable sensing technologies have been implementing new solutions and processes that can cut costs and increase convenience [1]. Popular applications include face authentication [2], banking risk management [3], medical diagnosis and healthcare [4], and personal smart assistants [5]. These technology-based advancements have the potential to improve quality of life by providing AI-based care and assistance, avoiding repetitive work, and making smart decisions from big data [6][7][8][9]. However, at the same time, these systems have raised concerns about invasion of privacy, constant behavioral monitoring, definitions of emotions, and inclusion of minorities [10]. The context of this dissertation focuses on affective computing and wearable sensing system design for human subjects. Thus, special ethical and risk considerations were required to avoid drawbacks and misuses of our research.

## 8.2 Ethical Considerations of this Dissertation

Wearable sensing and artificial intelligence technologies should be inclusive to minorities, such as individuals with neurodevelopmental disorders [11]. Individuals with autism spectrum disorders (ASD) show differences in behaviors, as compared to their typically developed peers [12] and many of them display hyper-sensitivity [13]. Bias in behaviors can be introduced if we train an affective computing model with data collected from typically developed individuals and apply it on autistic individuals. Therefore, we integrated and designed clinically grounded experimental protocols, with the advice from stakeholders in psychiatrics and functional analysis fields, to run data collections on individuals with ASD. With data collected from this population, the developed affective computing would not be biased to this group. Bias can also exist in unbalanced label distribution. If the target behavior has less than 5% in distribution, the machine learning model will be likely to have excessive false negatives, leading to failures to detect such behaviors [15]. Thus, we adjusted our data collection procedures so that the target behaviors and control behaviors have balanced distributions. In order to ensure that there were no induced distractions or stimuli [14], we designed and utilized non-invasive sensors and stimuli delivery devices to guarantee the safety and comfort of our data collections.

Appropriate definition of target emotions and behaviors is an important aspect of ethical affective computing. Emotions detected by the machine could be interpreted differently by the user themselves or due to different context [20]. Hence, problem behaviors in children with ASD need to be well defined and carefully treated. Our therapists consulted with the parents before the data collections on the most frequent precursors of problem behaviors, so that later on our observers and therapists could focus on those behaviors. Therefore, we were only investigating the precursors of problem behaviors that both the parents and our Board Certified Behavioral Analysts (BCBA) agreed upon. These highly trained professionals are certified to recognize unsafe behaviors so we were able to focus specifically on behaviors that would impede children's involvement in social communication, school and community, or cause physical harm [24]. It is suggested in the field that intervention on these problem behaviors augment the long-term behavioral outcomes [25][26].

Data security, privacy, and transparency are required so that users are comfortable with affective computing systems. To ensure these important aspects, we followed critical practice for all our data collections in this dissertation. First, the experimental protocol and sensors used were approved by Institutional Review Board (IRB) at Vanderbilt University so that our experiments were beneficial and of minimal risk for the children. Then, we introduced the purpose of our study, experimental procedure and setup to both the children and parents. We also verbally told the children and the parents that they could withdraw from the experiments any time they wanted. Finally, we obtained written consent from the parents because most children were not of age to consent. We introduced the mechanism of sensors to the parents and children before turning them on, and informed them when the sensors were recording data. Our collected data was safely stored on password-protected workstations and the access was strictly limited only to the key personnel on our IRB approved protocols.

As a conclusion, we have followed current guidelines and practices of ethical research in affective computing and wearable sensing field. We took precautions to protect the privacy of autistic individuals, kept them informed of the investigated behavior and sensor use, and made sure the target behavior was properly defined. We hope that these considerations could avoid misuses and ethical risks of this dissertation research.

## References


[1] M. Yousef Shaheen, "Applications of Artificial Intelligence (AI) in healthcare: A review Applications of Artificial Intelligence (AI) in healthcare: A review," no. September, pp. 0–2, 2021, doi: 10.14293/S2199-1006.1.SOR-.PPVRY8K.v1.

[2] A. Bud, "Facing the future: the impact of Apple FaceID," *Biometric Technol. Today*, vol. 2018, no. 1, pp. 5–7, Jan. 2018, doi: 10.1016/S0969-4765(18)30010-9.

[3] M. Leo, S. Sharma, and K. Maddulety, "Machine learning in banking risk management: A literature review," *Risks*, vol. 7, no. 1, 2019, doi: 10.3390/risks7010029.

[4] J. Waring, C. Lindvall, and R. Umeton, "Automated machine learning: Review of the state-of-the-art and opportunities for healthcare," *Artif. Intell. Med.*, vol. 104, p. 101822, Apr. 2020, doi: 10.1016/J.ARTMED.2020.101822.

[5] R. Knote, A. Janson, L. Eigenbrod, and M. Söllner, "The what and how of smart personal assistants: Principles and application domains for IS research," *MKWI 2018 - Multikonferenz Wirtschaftsinformatik*, vol. 2018-March, pp. 1083–1094, 2018.

[6] S. Lalmuanawma, J. Hussain, and L. Chhakchhuak, "Applications of machine learning and artificial intelligence for Covid-19 (SARS-CoV-2) pandemic: A review," *Chaos, Solitons & Fractals*, vol. 139, p. 110059, Oct. 2020, doi: 10.1016/J.CHAOS.2020.110059.

[7] A. A. Alfayez, H. Kunz, and A. G. Lai, "Predicting the risk of cancer in adults using supervised machine learning: A scoping review," *BMJ Open*, vol. 11, no. 9, pp. 1–11, 2021, doi: 10.1136/bmjopen-2020-047755.

[8] S. Nosratabadi *et al.*, "Data science in economics: Comprehensive review of advanced machine learning and deep learning methods," *Mathematics*, vol. 8, no. 10, pp. 1–25, 2020, doi: 10.3390/math8101799.

[9] A. Bannach-Brown *et al.*, "Machine learning algorithms for systematic review: Reducing workload in a preclinical review of animal studies and reducing human screening error," *Syst. Rev.*, vol. 8, no. 1, pp. 1–12, 2019, doi: 10.1186/s13643-019-0942-7.

[10] C. Reynolds and R. Picard, "Affective sensors, privacy, and ethical contracts," *Conf. Hum. Factors Comput. Syst.* -

*Proc.*, pp. 1103–1106, 2004, doi: 10.1145/985921.985999.
[11] C. Mattingly, “Autism and The Ethics of Care: A Phenomenological Investigation Into the Contagion of Nothing,” *Ethos*, vol. 45, no. 2, pp. 250–270, 2017, doi: 10.1111/etho.12164.
[12] W. Jones and A. Klin, “Attention to eyes is present but in decline in 2-6-month-old infants later diagnosed with autism,” *Nature*, 2013, doi: 10.1038/nature12715.
[13] L. Balasco, G. Provenzano, and Y. Bozzi, “Sensory Abnormalities in Autism Spectrum Disorders: A Focus on the Tactile Domain, From Genetic Mouse Models to the Clinic,” doi: 10.3389/fpsyt.2019.01016.
[14] M. Taj-Eldin, C. Ryan, B. O’flynn, and P. Galvin, “A review of wearable solutions for physiological and emotional monitoring for use by people with autism spectrum disorder and their caregivers,” *Sensors (Switzerland)*, vol. 18, no. 12, 2018, doi: 10.3390/s18124271.
[15] C. M. Bishop, *Pattern Recognition and Machine Learning*. Springer, 2006.
[16] E. Yadegaridehkordi, N. F. B. M. Noor, M. N. Bin Ayub, H. B. Affal, and N. B. Hussin, “Affective computing in education: A systematic review and future research,” *Comput. Educ.*, 2019, doi: 10.1016/j.compedu.2019.103649.
[17] M. Le Feber, T. Jadoenathmisier, H. Goede, E. Kuijpers, and A. Pronk, “Ethics and privacy considerations before deploying sensor technologies for exposure assessment in the workplace: Results of a structured discussion amongst dutch stakeholders,” *Ann. Work Expo. Heal.*, vol. 65, no. 1, pp. 3–10, 2021, doi: 10.1093/annweh/wxaa093.
[18] M. Malekzadeh, R. G. Clegg, A. Cavallaro, and H. Haddadi, “Privacy and utility preserving sensor-data transformations,” *Pervasive Mob. Comput.*, vol. 63, p. 101132, 2020, doi: 10.1016/j.pmcj.2020.101132.
[19] G. Greene, “The Ethics of AI and Emotional Intelligence,” *Partnersh. AI*, [Online]. Available: https://www.partnershiponai.org/the-ethics-of-ai-and-emotional-intelligence/.
[20] E. Kanjo, L. Al-Husain, and A. Chamberlain, “Emotions in context: examining pervasive affective sensing systems, applications, and analyses,” *Pers. Ubiquitous Comput.*, vol. 19, no. 7, pp. 1197–1212, 2015, doi: 10.1007/s00779-015-0842-3.
[21] K. H. Choi, J. Kim, O. S. Kwon, M. J. Kim, Y. H. Ryu, and J. E. Park, “Is heart rate variability (HRV) an adequate tool for evaluating human emotions? – A focus on the use of the International Affective Picture System (IAPS),” *Psychiatry Res.*, vol. 251, no. July 2016, pp. 192–196, 2017, doi: 10.1016/j.psychres.2017.02.025.
[22] M. A. Gianfrancesco, S. Tamang, J. Yazdany, and G. Schmajuk, “Potential Biases in Machine Learning Algorithms Using Electronic Health Record Data,” *JAMA Intern. Med.*, vol. 178, no. 11, pp. 1544–1547, 2018, doi: 10.1001/jamainternmed.2018.3763.
[23] J. Jessel, E. T. Ingvarsson, R. Metras, R. Whipple, H. Kirk, and L. Solsbery, “Treatment of elopement following a latency-based interview-informed, synthesized contingency analysis,” *Behav. Interv.*, vol. 33, no. 3, pp. 271–283, 2018, doi: 10.1002/bin.1525.
[24] S. M. Myers *et al.*, “Management of children with autism spectrum disorders,” *Pediatrics*, vol. 120, no. 5, pp. 1162–1182, 2007, doi: 10.1542/peds.2007-2362.
[25] E. Emerson, “Challenging behaviour: Analysis and intervention in people with severe intellectual disabilities,” 2001.
[26] S. L. Parish, K. C. Thomas, R. Rose, M. Kilany, and P. T. Shattuck, “State medicaid spending and financial burden of families raising children with autism,” *Intellect. Dev. Disabil.*, vol. 50, no. 6, pp. 441–451, 2012, doi: 10.1352/1934-9556-50.06.441.