PSYCHOPHYSIOLOGICAL ANALYSIS OF AFFECTIVE STATES IN HUMAN-COMPUTER INTERACTION FOR CHILDREN WITH AUTISM SPECTRUM DISORDERS

By

Karla Conn Welch

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

Electrical Engineering

December, 2009

Nashville, Tennessee

Approved:

Professor Nilanjan Sarkar

Professor Robert Bodenheimer

Professor Wendy Stone

Professor Zachary Warren

Professor Mitch Wilkes

To my family

ACKNOWLEDGEMENTS

Many thanks are extended to my dissertation committee for their support. Particularly my advisor, Dr. Nilanjan Sarkar, has been a great mentor and is perpetually prepared with guidance. This work would not have been possible without the initial commitment from Dr. Wendy Stone to join in this cross-disciplinary project as well as Dr. Zachary Warren's continued guidance of the work. Also, I thank Dr. Bobby Bodenheimer and Dr. Mitch Wilkes for their helpful suggestions that substantially improved this research. I am grateful for the National Science Foundation Graduate Research Fellowship, Marino Autism Research Institute grant, Autism Speaks grant, and Vanderbilt Institute for Clinical/Translational Research grant that financially supported this work.

I would like to thank Gary Walker, Robin Midgett, and Tony Costi for their technical support throughout the user studies. Also, many thanks go to Amy Swanson for serving as a reliable clinical observer. Working with my fellow lab members of the Robotics and Autonomous Systems Laboratory has been both enjoyable and educational. In particular, I thank Changchun Liu, Uttama Lahiri, Duygun Erol, and Bibhrajit Halder for their hard work and support.

Mostly, I thank my family; Ann, David, Kelly, Danny, Will, Kathy, Pierce, Rose, and Tommy; and friends; Kat, Kaci, Candi, Katie, Galen, Angela, Danielle, Ginny, Jennifer, Bill, Nick, MaryAnne, Carolyn, Bethany, Melinda, Stacy, and Bianca; who have seen me through both bright and dark days of this journey. I could not have accomplished this goal without their encouragement. In particular, I thank my husband, Aaron, for his love and understanding throughout this work.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ADI-R	Autism Diagnostic Interview-Revised
ADOS-G	Autism Diagnostic Observation Schedule-Generic
ASD	Autism Spectrum Disorder
BVP	Blood Volume Pulse
DSM-IV	Diagnostic and Statistical Manual of Mental Disorders-4th edition
EC	Experiment Condition
ECG	Electrocardiogram
EDA	Electrodermal Activities
EMG	Electromyogram
GSR	Galvanic Skin Response
HCI	Human-computer Interaction
HRI	Human-robot Interaction
IBI	Inter Beat Interval
ICG	Impedance Cardiogram
IRB	Institutional Review Board
K	Kappa coefficient
PCG	Phonocardiogram
PEP	Pre-ejection Period
PPG	Photoplethysmogram
PPVT	Peabody Picture Vocabulary Test
PTT	Pulse Transit Time
SCQ	Social Communication Questionnaire
SD	Standard Deviation
SRS	Social Responsiveness Scale
SVM	Support Vector Machines
VR	Virtual Reality

CHAPTER I

INTRODUCTION

Introduction

This dissertation covers the application of psychophysiological analysis to children with Autism Spectrum Disorders (ASD) during human-computer interaction (HCI) and human-robot interaction (HRI). Investigation into technology-assisted intervention for children with ASD has gained momentum in recent years. Clinicians[1] involved in interventions must overcome the communication impairments generally exhibited by children with ASD by adeptly inferring the affective cues of the children to adjust the intervention accordingly. Similarly, an intelligent system, such as a computer or robot, must also be able to understand the affective needs of these children - an ability that the current technology-assisted ASD intervention systems lack - to achieve effective interaction that addresses the role of affective states in HCI, HRI, and intervention practice. Affective cues are indicators, external or internal, of the manifestations of emotions and feelings experienced in a given environment. This research utilizes and merges recent technological advances in the areas of (i) robotics, (ii) virtual reality (VR), (iii) physiological signal processing, (iv) machine learning techniques, and (v) adaptive response technology in an attempt to create a tool for understanding various physiological aspects of social communication in children with ASD. The individual, familial, and

---

[1]The terms "clinician," "clinical observer," and "therapist" are used interchangeably in this dissertation to mean an expert with skill in making judgments, such as rating affective states, about the meaning of observable behaviors from individuals with autism.

societal impact associated with the presumed core social impairments of children with ASD is enormous. Thus, there is a need to better understand the underlying mechanisms and processes associated with these deficits as well as develop tools that can be used to create optimal intervention strategies.

Autism is a neurodevelopmental disorder characterized by core deficits in social interaction, social communication, and imagination (American Psychiatric Association, 2000). These characteristics often vary significantly in combination and severity, within and across individuals, as well as over time. Research suggests prevalence rates as high as approximately 1 in 150 for the broad autism spectrum (CDC, 2007). While, at present, there is no single universally accepted intervention, treatment, or known cure for ASD (NRC, 2001; Sherer and Schreibman, 2005); there is an increasing consensus that intensive behavioral and educational intervention programs can significantly improve long term outcomes for individuals and their families (Rogers, 1998).

In response to this need, a growing number of studies have been investigating the application of advanced interactive technologies to address core deficits related to autism, namely computer technology (Bernard-Opitz et al., 2001; Moore et al., 2000; Swettenham, 1996), virtual reality environments (Parsons et al., 2004; Strickland et al., 1996; Tartaro and Cassell, 2007), and robotic systems (Dautenhahn and Werry, 2004; Kozima et al., 2009; Michaud and Theberge-Turmel, 2002; Pioggia et al., 2005; Scassellati, 2005). Computer- and VR-based intervention may provide a simplified but exploratory interaction environment for children with ASD (Moore et al., 2000; Parsons et al., 2004; Strickland et al., 1996). Robots have been used to interact with children with ASD in common imitation tasks and can serve as social mediators to facilitate interaction with other children and caregivers (Dautenhahn and Werry, 2004; Kozima et al., 2009). In the rest of the dissertation, the term "computer" is used to imply both computer- and robot-assisted ASD interventions.

Even though there is increasing research in technology-assisted autism intervention, there is a paucity of published studies that specifically address how to automatically detect and respond to affective cues of children with ASD. Such ability could be critical given the importance of human affective information in HCI (Picard, 1997; Prendinger et al., 2005) and HRI (Fong et al., 2003) and the significant impacts of the affective factors of children with ASD on the intervention practice (Ernsperger, 2003; Seip, 1996; Wieder and Greenspan, 2005). A computer that can detect the affective states of a child with ASD and interact with him/her based on such perception could have a wide range of potential impacts. Interesting activities likely to retain the child’s attention could be chosen when a low level of engagement is detected. Complex social stimuli, sophisticated interactions, and unpredictable situations could be gradually, but automatically, introduced when the computer recognizes that the child is comfortable or not anxious at a certain level of interaction dynamics for a reasonably long period of time. A clinician could use the history of the child’s affective information to analyze the effects of the intervention approach. With the record of the activities and the consequent emotional changes in a child, a computer could learn individual preferences and affective characteristics over time and thus could alter the manner in which it responds to the needs of different children. The research in this dissertation assesses what effects there are on physiological response for children with ASD during performance-oriented and socially-oriented tasks. The ability to detect the physiological processes that are a part of impairments in social communication may prove an important tool for understanding the physiological mechanisms that underlie the presumed core impairments associated with ASD.

Background

Physiology for Affect Recognition of Children with ASD

There are several modalities such as facial expression (Bartlett et al., 2003), vocal intonation (Lee and Narayanan, 2005), gestures and postures (Asha et al., 2005; Kleinsmith et al., 2005), and physiology (Kulic and Croft, 2007; Mandryk et al., 2006; Nasoz et al., 2004; Rani et al., 2004) that can be utilized to evaluate the affective states of individuals interacting with computer. This work evaluates affective states based on physiological data for several reasons. Children with ASD often have communicative impairments (both nonverbal and verbal), particularly regarding expression of affective states (American Psychiatric Association, 2000; Green et al., 2002; Schultz, 2005). These vulnerabilities place limits on computerized affective modeling based on traditional conversational and observational methodologies. For example, video has been used to teach children with ASD to recognize facial expressions and emotions of *others* (Stokes, 2000), but no published studies were found that used visual recognition through video to autonomously determine the affective states of people with ASD. A facial recognition algorithm could be designed to detect certain expressions but would have to accommodate when expressions are abnormal (e.g., smiling under mild pain, etc.) or lack variability (Schultz, 2005). Physiological signals, however, are continuously available and are not necessarily directly impacted by these difficulties (Ben Shalom et al., 2006; Groden et al., 2005; Toichi and Kamio, 2003). As such, physiological modeling may represent a methodology for gathering rich data despite the potential communicative impairments of children with ASD. In addition, physiological data may offer an avenue for recognizing aspects of affect that may be less obvious for humans but more suitable for computers by using signal processing and pattern recognition tools. Furthermore, there is evidence that the transition from one affective state to another state is accompanied by dynamic shifts in indicators of Autonomic Nervous System (ANS) activity (Bradley, 2000). More than one physiological signal, judged as a favorable approach (Bethel et al., 2007), is examined in this research, and the set of signals consists of various cardiovascular, electrodermal, electromyographic, and skin temperature signals, all of which have been extensively investigated in psychophysiology literature (Bradley, 2000).

One of the prime challenges of this work is attaining reliable subjective reports. There have been reports that adolescents could be better sources of information than adults when it comes to measuring some psychiatric symptoms (Cantwell et al., 1997), but researchers are generally reluctant to trust the responses of adolescents on self-reports (Barkley, 1998). One should be especially wary of the dependability of self-reports from children with ASD, who may have deficits in processing (i.e., identifying and describing) their own emotions (Hill et al., 2004). While there have been some criticisms on the use of subjective report (i.e., self-assessment or the reports collected from observers) and its effect on possibly forcing the determination of emotions, the subjective report is by and large regarded as an effective way to evaluate affective responses. Due to the unresolved debate on the definition of emotion (e.g., objective entities or socially constructed labels), researchers in affective computing often face difficulties obtaining the ground truth to label the natural emotion data accordingly. As suggested by Cowie et al. (2001) and Pantic and Rothkrantz (2003), the immediate implication of such a controversy is that pragmatic choices (e.g., application and user-profiled choices) must be made to develop an automatic affect recognizer. As a result, subjective report is widely used for affective modeling and endowing a computer with the recognition abilities similar to those of the reporters (Picard, 1997; Silva et al., 2006).

An important question when estimating human affective response is how to operationalize the affective state. Although much existing research on affective modeling categorizes physiological signal data into "basic emotions," there is no consensus on a set of basic emotions among the researchers (Cowie et al., 2001). This fact implies that practical choices are required to select target affective states for a given application. Anxiety, engagement, and enjoyment/liking are chosen as the target affective states in this dissertation research. Anxiety is chosen for two primary reasons. First, anxiety plays an important role in various human-machine interaction tasks that can be related to task performance (Brown et al., 1997). Second, anxiety frequently co-occurs with ASD and plays an important role in the behavior difficulties of children with autism (Gillott et al., 2001). Engagement, defined as "sustained attention to an activity or person" (NRC, 2001), has been regarded as one of the key factors for children with ASD to make substantial gains in academic, communication, and social domains (Ruble and Robson, 2006). With "playful" activities during the intervention, the liking of the children (i.e., the enjoyment they experience when interacting with the computer) may create urges to explore and allow prolonged interaction for the children with ASD, who are susceptible to being withdrawn (Dautenhahn and Werry, 2004; Papert, 1993).

Literature in the human factors and psychophysiology fields provide a rich history in support of physiology methodologies for studying stress (Groden et al., 2005; Zhai et al., 2005), engagement (Pecchinenda and Smith, 1996), operator workload (Kramer et al., 1987), mental effort (Vicente et al., 1987), and other similar mental states based on physiological measures such as those derived from electromyogram (EMG), galvanic skin response (GSR; i.e., skin conductance), heart rate variability (HRV), and blink rates. Meehan et al. (2005) reported that changes in physiological activity are evoked by different amounts of presence in stressful VR environments. Prendinger et al. (2005) demonstrated that the measurement of GSR and EMG can be used to discriminate a user's instantaneous change in levels of anxiety due to sympathetic vs. unconcerned reactions from a life-like virtual teacher. In general, it is expected that higher physiological activity levels can be associated with greater stress levels (Smith, 1989). Therefore, developing technologies for exploration of physiological signals and the target affective states of anxiety, engagement, and enjoyment/liking that may be associated with core social deficits for children with ASD is both scientifically valid and technologically feasible.

Technology in the Treatment of ASD

Interventions often focus on social communication, including social-problem solving and social skills training, so that participants can gain experience and exposure to various situations representative of everyday living. The ultimate goal of such interventions is for some generalization of these skills to carry over into real-life situations. A growing number of studies have been exploring the application of interactive technologies for future use in interventions to address the social deficits of children with ASD. Initial results indicate that such technologies hold promise as a potential alternative intervention approach with broad accessibility. Various software packages and VR environments have been developed and applied to address specific deficits associated with autism, e.g., understanding of false belief (Swettenham, 1996), attention (Trepagnier et al., 2006), social problem-solving (Bernard-Opitz et al., 2001), and social conventions (Parsons et al., 2005). Research on applying robotics to ASD intervention has suggested that robots can allow simplified but embodied social interaction that is less intimidating or confusing for children with ASD (Robins et al., 2005). By employing HCI and HRI technologies, interactive intervention tools can partially automate the time-consuming, routine behavioral intervention sessions and may allow intensive intervention to be conducted at home (Dautenhahn and Werry, 2004). For the purpose of using affective computing tools, computers or robots could be the mode of technology for assisted ASD interventions.

Dautenhahn and colleagues have explored how a robot can become a playmate that might serve a therapeutic role for children with autism in the Aurora project. Dautenhahn et al. (2003) emphasize the importance of robot adaptability in autism rehabilitation. Research showed that children with ASD are engaged more with an autonomous robot in the "reactive" mode than with an inanimate toy or a robot showing rigid, repetitive, non-interactive behavior (Dautenhahn and Werry, 2004; Robins et al., 2004). Michaud and Theberge-Turmel (2002) investigated the impact of robot design on the interactions with children with ASD and pointed out that systems need to be versatile enough to adapt to the varying needs of different children. Pioggia et al. (2005) developed an interactive life-like facial display system for enhancing emotion recognition in people with ASD. Robotic technologies pose the advantage of furnishing robust systems that can support multimodal interaction and provide a repeatable, standardized stimulus while quantitatively recording and monitoring the performance progress of the children with ASD to facilitate autism intervention assessment and diagnosis (Scassellati, 2005).

There are numerous reasons why a VR-based intervention system may be particularly relevant for children with ASD. The strength of VR technology for ASD intervention includes malleability, controllability, reduced sensory stimuli, individualized approach, safety, and a reduction of human interaction during initial skill training (Strickland, 1997). VR does not necessarily include direct human-to-human interaction, which may work well for an initial intervention to remove the difficulties common in ASD related to mere human interaction that is part of a typical intervention setting involving a child and a clinician (Chen and Bernard-Opitz, 1993; Tartaro and Cassell, 2007). However, VR should not be considered an isolating agent, because dyadic communication accomplished between a child and a VR environment can lead into triadic communication including a clinician, caregiver, or peer and in due course potentially accomplish the intervention goals of developing social communication skills between the child with ASD and another person (Bernard-Opitz et al., 2001). Furthermore, the main sensory output of VR is auditory and visual, which may represent a reduction of information from a real-world setting but also represents a full description of a setting without need for imagined components (Sherman and Craig, 2003; Strickland, 1997). Individuals with ASD can improve their learning skills related to a situation if the proposed setting can be manifested in a physical or visual manner (Kerr and Durkin, 2004). Since VR mimics real environments in terms of imagery and contexts, it may allow for efficient generalization of skills from the VR environment to the real world (Cromby et al., 1996). However, since limited social insight and social cognition are vulnerabilities that are often part of the core deficits associated with ASD, individuals may lack the skills to envision abstract concepts or changes to situations on their own. Virtual environments can easily change the attributes of, add, or remove objects in ways that may not be possible in a real-world setting but could be valuable to teach abstract concepts. Therefore, VR can offer the benefit of representing abstract concepts through visual means (e.g., thought bubbles with text descriptions of a virtual character's thoughts) and seamlessly allows for changes to the environment (e.g., changing the color of a ball or making a table disappear) that may be difficult or even impossible to accomplish in a real-world setting (Sherman and Craig, 2003; Strickland, 1997). Furthermore, the spectrum nature of autism means an individual approach is appropriate, and computers can accommodate individualized treatment (Strickland, 1997). The highly versatile VR environment can illustrate scenarios which can be changed to accommodate various situations that may not be feasible in a given therapeutic setting because of space limitations, resource deficits, safety concerns, etc. (Parsons and Mitchell, 2002). Therefore, VR represents a medium well-suited for creating interactive intervention paradigms for skill training in the core areas of impairment for children with ASD (i.e., social interaction, social communication, and imagination). However, to date the capability of VR technology has not been fully explored to examine the factors that lead to difficulties in impairments such as social communication, which could be critical in designing an efficient intervention plan.

Consensus statements from both the American Academy of Pediatrics (Myers et al., 2007) and the National Resource Council (NRC, 2001) underscore that effective intervention for children with ASD includes: provision of intensive intervention, individual instruction tailored to the qualities of the child, promotion of a generalization of skills, and incorporation of a high degree of structure/organization. Despite the urgent need and societal import of intensive treatment (Rutter, 2006), appropriate intervention resources for children with ASD and their families are often difficult to access and extremely costly when accessible (Jacobson et al., 1998; Sharpe and Baker, 2007; Tarkan, 2002). Therefore, an important direction for research on ASD is the identification and development of technological tools that can make application of effective intensive treatment more readily accessible and cost effective (Parsons and Mitchell, 2002; Rogers, 2000). VR has also shown the capacity to ease the burden, both time and effort, of trained clinicians in an intervention process as well as the potential to allow untrained personnel (e.g., parents or peers) to aid a participant in the intervention (Standen and Brown, 2005). As such, the future creation of a VR-assisted affect-sensitive tool for autism intervention could meet all of the core components of effective intervention, while at the same time increasing the ability of the intervention provider to systematically control and promote intervention related skills.

Affective cues are insights into the emotions and behaviors of children with ASD. The ability to utilize the power of these cues may permit a smooth, natural, and more productive interaction process (Gilleade et al., 2005; Kapoor et al., 2001; Picard, 1997; Prendinger et al., 2005), especially considering the core social and communicative vulnerabilities that limit individuals with ASD to accurately self-identify affective experiences (Hill et al., 2004). Common in autism intervention, clinicians who work with children with ASD intensively monitor affective cues of the children in order to make appropriate decisions about adaptations to their intervention and reinforcement strategies. For example, "likes and dislikes chart" is recommended to record the children's preferred activities and/or sensory stimuli during interventions that could be used as reinforcers and/or "alternative behaviors" (Seip, 1996). Children with autism are particularly vulnerable to anxiety and intolerant of feelings of frustration, which requires a clinician to plan tasks at an appropriate level of difficulty (Ernsperger, 2003). The engagement of children with ASD is the ground basis for the "floor-time therapy" to help them develop relationships and improve their social skills (Wieder and Greenspan, 2005). Given the importance of affective cues in ASD intervention practice (Ernsperger, 2003; Seip, 1996; Wieder and Greenspan, 2005), using affective information as a means of implicit and bidirectional communication may be critical for allowing a computer to respond to a child's affective states. The design of affect-sensitive interaction, an area known as affective computing, is an increasingly important discipline within the HCI and HRI communities (Picard, 1997). However, to date little work has been done to explore this approach for technology-assisted intervention of individuals with ASD. Furthermore, no existing technology specifically addresses how to autonomously detect and flexibly respond to affective cues of children with ASD within an intervention paradigm (Bernard-Opitz et al., 2001; Dautenhahn and Werry, 2004; Kozima et al., 2009; Michaud and Theberge-Turmel, 2002; Mitchell et al., 2007; Parsons et al., 2005; Pioggia et al., 2005; Scassellati. 2005; Strickland, 1997; Swettenham, 1996; Tartaro and Cassell, 2007; Trepagnier et al., 2006). The primary contribution of this dissertation is to address this deficiency. The research develops HCI technologies capable of eliciting affective changes in individuals with ASD. We investigate how to augment HRI to be used in affect-sensitive interaction by endowing the technology with the ability to recognize and flexibly respond to the affective states of a child with ASD based on his/her physiological responses. The research also assesses the efficacy of measuring affect in VR.

Summary

This dissertation is presented as follows. Chapter II demonstrates the feasibility of affect modeling for children with ASD using proof-of-concept performance-based tasks on a computer. Chapter III builds on results from Chapter II by applying the affective models in an online investigation of affect recognition and affect-sensitive adaptation during interactions between children with ASD and a robot. Since social communication is one of the core deficits for children with ASD, Chapter IV represents a shift from performance-based tasks to socially-oriented tasks. A social interaction system is developed using virtual reality environments. Experiments with two groups of participants engaging with the social interaction system are presented in Chapter V. Physiological signal detection and VR technology are combined to explore realistic social situations within the VR environment that can be used to understand the physiological mechanisms that underlie the presumed core social impairments associated with ASD.

Conclusions and an overview of the contributions of this work are covered in Chapter VI along with a synopsis of future work.

CHAPTER II

PHYSIOLOGY-BASED AFFECTIVE MODELING OF CHILDREN WITH AUTISM SPECTRUM DISORDERS

Introduction

The primary objective of this chapter is to investigate the feasibility of affective modeling for children with ASD via a physiology-based affect recognition technique. There are several challenges that need to be addressed to develop a robust affect recognizer for children with ASD (e.g., obtaining reliable subjective reports on affective states, developing quantitative models for affect recognition). As such, there is a dearth of literature on quantitative modeling results of affect recognition of children with ASD (e.g., affective model with reliable prediction capability). Several researchers in the human–machine interaction community have focused on physiology-based affect-recognition for typical adults. Picard et al. (2001) have employed a combination of Sequential Floating Forward Search and Fisher Projection methods to classify eight emotions with 81% accuracy. K-Nearest Neighbors (KNN), Discriminant Function Analysis, and Marquardt Backpropagation algorithms were applied to differentiate among six emotions by Nasoz et al. (2004), and the correct classification accuracies – 71%, 74%, and 83%, respectively – were achieved for the three algorithms. Rani et al. (2006) compared several machine learning algorithms, namely, KNN, Bayesian Network, Support Vector Machines (SVM), and Regression Tree for determining the intensity of the affective states, and the best prediction accuracy rate 85.8% was achieved using SVM. The results in this chapter demonstrate for the first time that reliable affective models can be developed for children with ASD based on their physiological signals while they play interactive computer games.

Children with ASD are recommended to undergo at least 25 hours-per-week of year-round intensive autism intervention (i.e., one-on-one therapy with a trained therapist) outside of school and extracurricular activities (NRC, 2001; Tarkan, 2002). The developed affective model can be used in the computer-assisted autism interventions to detect the children's affective states on-line, move them toward the intervention goals in an affective manner, and make the treatment more accessible (e.g., possibly allowing intensive intervention to be conducted at home). The novelty of the presented affective model is that it is individual-specific, and it consists of an array of recognizers, each of which determines the intensity (e.g., high/low level) of one target affective state for each individual. An affect recognizer for children with ASD may need to be individual-specific to accommodate the differences encountered in emotional expression and the spectrum nature of autism (American Psychiatric Association, 2000). Therefore, the model takes into account evidence that the affective state could be an aggregate of various affective categories at different arousal levels (Vansteelandt et al., 2005) and that within a given context different individuals express the same emotion with different characteristic response patterns (i.e., phenomenon of person stereotypy) (Lacey and Lacey, 1958). Even though physiology has been successfully employed to build affect recognizers for typical individuals in several research groups (Kulic and Croft, 2007; Mandryk and Atkins, 2007; Picard et al., 2001; Rani et al., 2006), the studies of the correlation of the physiological signals and the affective states of people with ASD are relatively few (Ben Shalom et al., 2006; Groden et al., 2005) and no quantitative modeling results have been reported.

This work included an autism therapist who has five years of experience in therapeutic and diagnostic interventions for children with ASD and each participant's parent. The therapist and the parent observed the experiments and provided subjective reports based on their expertise/experience in inferring the presumable underlying affective states from the observable behaviors of a child with ASD. The therapist and the parent did not use the participant's physiological signals to recognize affective states, but these signals were recorded for eventual affective modeling (i.e., a mapping between the objective physiological signals and the subjective reports). In this study, the therapist's reports on perceived intensity of the affective states of a participating child and the extracted physiological indices are employed to build therapist-like affect recognizers. In autism interventions, a therapist continuously monitors the affective cues of children with ASD based on behavioral observations. In this work, the "therapist-like affect recognizers" were developed to emulate the therapist's affect-recognition capability, however, based on the children's physiological signals. With the incorporation of the therapist's reports, the recognizers will be capable of autonomously delivering similar assessments of the affective states of the children with ASD in real time even when the therapist is not available. Ultimately, integrating the affective models with current interactive intervention approaches may allow for automating the intensive, repetitive aspects of the existing behavioral therapy techniques and possibly steer the individual towards the intervention goal in an affect-sensitive manner.

Physiological Signal Acquisition and Indices

There is good evidence that the physiological activity associated with affective states can be differentiated and systematically organized (Bradley, 2000). The cardiovascular and electromyogram activities have been used to examine the positive and negative affective states of people (Cacioppo et al., 2000; Papillo and Shapiro, 1990). Electrodermal activities have been shown to be associated with task engagement (Pecchinenda and Smith, 1996). The variation of peripheral temperature due to emotional stimuli was studied by (Kataoka et al., 1998). In this chapter, we exploited the dependence of physiological responses on underlying affective states to develop affective models for children with ASD by using the machine learning method as described in Appendix A. The physiological signals examined were: various cardiovascular activities including electrocardiogram (ECG), impedance cardiogram (ICG), photoplethysmogram (PPG), and phonocardiogram (PCG)/heart sound; electrodermal activities (EDA) including tonic and phasic responses from skin conductance; electromyogram (EMG) activities from corrugator supercilii, zygomaticus major, and upper trapezius muscles; and peripheral temperature. Relevant features were derived from the physiological signals using various signal-processing techniques such as Fourier transform, wavelet transform, thresholding, and peak detection. The physiological signals that were examined with the features derived from each signal are described in Table 2.1.

Acquisition of Physiological Signals

The physiological signals were acquired using the Biopac MP150 physiological data acquisition system (biopac.com). ECG was measured from the chest using the

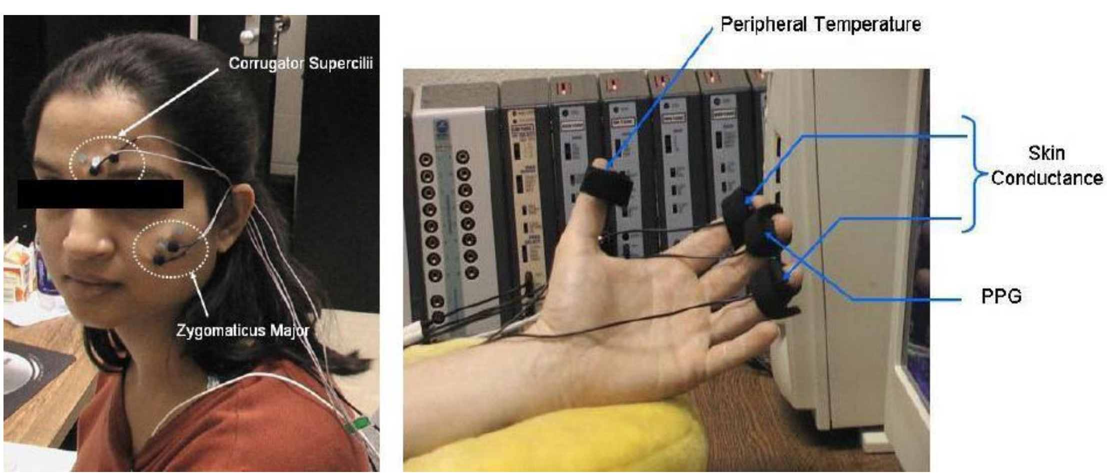
Table 2.1 Physiological Indices

Physiological Response	**Features Derived**	**Label Used**	**Unit of Measurement**
Electrocardiogram	Sympathetic power from ECG	Sym	unit/square second (unit/s^2)
	Parasympathetic power from ECG	Para	unit/s^2
	Very Low Frequency Power from ECG	VLF	unit/s^2
	Ratio of powers	Para/VLF Para/Sym VLF/Sym	No unit
	Mean of IBI	IBI_ECGmean	milliseconds (ms)
	SD of IBI	IBI_ECGstd	Standard Deviation (SD, ms)
Photoplethysmogram	Mean amplitude of the peak values of the PPG signal	PPG_Peakmean	microvolt (µV)
	Maximum amplitude of the peak values of the PPG signal	PPG_Peakmax	µV
	Mean of IBI of PPG	IBI_PPGmean	ms
	SD of IBI of PPG	IBI_PPGstd	ms
	Mean Pulse Transit Time (PTT)	PTTmean	ms
	SD Pulse Transit Time (PTT)	PTTstd	ms
Heart Sound	Mean of the 3rd, 4th, and 5th level coefficients of the Daubechies wavelet transform of heart sound	D3_HSmean D4_HSmean D5_HSmean	No unit
	SD of the 3rd, 4th, and 5th level coefficients of the Daubechies wavelet transform of heart sound	D3_HSstd D4_HSstd D5_HSstd	No unit
Bioimpedance	Mean Pre-Ejection Period (PEP)	PEPmean	ms
	SD Pre-Ejection Period (PEP)	PEPstd	ms
	Mean of IBI of ICG	IBI_ICGmean	ms
	SD of IBI of ICG	IBI_ICGstd	ms
Electrodermal activity	Mean tonic activity level	Tonicmean	microsiemens (µS)
	Slope of tonic activity	Tonicslope	µS/s
	Mean amplitude of skin conductance response (phasic activity)	Phasicmean	µS
	Maximum amplitude of skin conductance response (phasic activity)	Phasicmax	µS
	Rate of phasic activity	Phasicrate	peaks/min
Electromyographic activity	Mean of Corrugator Supercilii activity	Cormean	µV
	SD of Corrugator Supercilii activity	Corstd	µV
	Slope of Corrugator Supercilii activity	Corslope	µV/s
	Mean of IBI of blink activity	IBI_Blinkmean	s
	Mean amplitude of the peak values of blink activity	Blink_Peakmean	µV

Table 2.1 (continued) Physiological Indices

Physiological Response	Features Derived	Label Used	Unit of Measurement
Electromyographic activity	Mean amplitude of blink activity	Blinkmean	μV
	SD of blink activity	Blinkstd	μV
	Mean of Zygomaticus Major activity	Zygmean	μV
	SD of Zygomaticus Major activity	Zygstd	μV
	Slope of Zygomaticus Major activity	Zygslope	μV/s
	Mean of Upper Trapezius activity	Trapmean	μV
	SD of Upper Trapezius activity	Trapstd	μV
	Slope of Upper Trapezius activity	Trapslope	μV/s
	Mean and Median frequency of Corrugator, Zygomaticus, and Trapezius	Cfreqmean Zfreqmean Tfreqmean Cfreqmedian Zfreqmedian Tfreqmedian	Hertz
Temperature	Mean temperature	Tempmean	Degree Fahrenheit (°F)
	Slope of temperature	Tempslope	°F/s
	SD of temperature	Tempstd	°F

standard two-electrode configuration. ICG describes the changes of thorax impedance due to cardiac contractility and was measured by four pairs of surface electrodes that were longitudinally configured on both sides of the body. The top pair of ICG electrodes was placed on the neck parallel to and about 3 cm above the second pair, located at the base of the neck; the bottom electrodes were placed parallel to and about 5 cm below the third ones, which were placed on the sides of the chest at the level of the xiphisternal junction. A microphone specially designed to detect heart sound waves was placed on the chest to measure PCG. PPG, peripheral temperature, and EDA were measured from the middle finger, the thumb, and the index and ring fingers of the non-dominant hand, respectively, using surface electrodes sewn in stretchy Velcro straps. EMG was measured by placing surface electrodes on two facial muscles (corrugator supercilii and zygomaticus major) and an upper back muscle (upper trapezius). All the physiological sensors were extensions of the Biopac physiological data acquisition system. The sampling rate was fixed at 1000 Hz for all the channels. Appropriate amplification and band-pass filtering were performed. Before each session, a three-minute baseline recording was done that was later used to offset day-variability. During the baseline recording, participants were asked to relax in a seated position and read age-appropriate leisure material. Subsequently, emotional stimuli induced by cognitive tasks were applied in epochs of up to four minutes in length. Previous research (Pecchinenda and Smith, 1996; Rani et al., 2006) has shown that physiological signals (e.g. electrodermal activity, electromyographic activity, and cardiovascular activity) of 2-4 minutes in length were adequate for detecting affective states (e.g., anxiety, anger, engagement, etc.) from similar computer-based tasks. Each child with ASD took part in six one-hour sessions containing 13-15 epochs each. Each session took place on a different day to avoid bias in data due to habituation. Figure 2.1 shows the sensor setup.



(a) (b)

Figure 2.1 The sensor setup shows the position of facial EMG sensors (a) and the placement of sensors on the non-dominant hand (b).

Cardiovascular Activity

ECG measures the heart activity through the electrical signal of the heart muscle. The number of beats per minute (bpm) is called the heart rate and is typically 70-80 bpm at rest. Inter beat interval (IBI) is the time interval in milliseconds between two “R” peaks in the ECG waveform. The R-peak detection algorithm performed band-pass filtering on the raw ECG signal and the signal was then smoothed by a 10 ms moving average window. Peaks were then detected in the resulting signal, and detection heuristic rules were applied to avoid missing R peaks or detecting multiple peaks for a single heart beat. These rules included obtaining the amplitude threshold (the difference between a peak and the corresponding inflection point) at which a peak should be considered a beat, enforcing a minimum interval of 300 ms and maximum interval of 1500 ms between peaks, checking for both positive and negative slopes in a peak to ensure that baseline drift is not misclassified as a peak, and backtracking with reexamination/interpolation when peak missing was detected. Generally, the average change for heart rate is expected to be within the range of 2-15 bpm (Bradley, 2000). The chosen interval threshold between peaks was well above the rate of change of heart rate due to genuine heart acceleration. Time-domain features of IBI, such as the mean and standard deviation (SD), can be computed from the detected R peaks. IBI variability was explored by performing power spectral analysis on the IBI data to localize the sympathetic and parasympathetic nervous system activities associated with different frequency bands. “Sym” was the power associated with the sympathetic nervous system activity of the heart (in frequency band 0.07-0.14 Hz). “Para” was the power associated with the parasympathetic nervous system activity of the heart (in frequency band 0.15-0.5 Hz). “VLF” was the power associated with the frequency band from 0.02-0.06 Hz. The ratios of different frequency components were also computed as the input features for affective modeling.

The PPG signal measures changes in the volume of blood in the fingertip associated with the blood volume pulse (BVP) cycle, and it provides an index of the relative constriction versus dilation of the blood vessels in the periphery. The raw PPG signal was smoothed by a 10 ms moving average window, and the baseline drift was accounted for by subtracting the average value of the signal. Pulse transit time (PTT) is the time it takes for the pulse pressure wave to travel from the heart to the periphery, and it was estimated by computing the time between systole at the heart (as indicated by the R-wave of the ECG) and the peak of the BVP wave reaching the peripheral site where PPG was measured. Besides the mean and SD of PTT, the mean and maximum values of BVP peak amplitudes and mean and SD of the IBI of BVP peaks were also extracted as features.

ICG analysis measures the impedance or opposition to the flow of an electric current through the body fluids. The ICG signal was first filtered by a 5th order Butterworth filter (low-pass: 10 Hz) to clean up any residual noise on the waveform and was then differentiated. Pre-ejection period (PEP), derived from ICG and ECG signals, measures the latency between the onset of electromechanical systole and the onset of left-ventricular ejection and is most heavily influenced by sympathetic innervation of the heart. The time intervals between the successive peaks of ICG time-derivative and “R” peaks of ECG were calculated to obtain the value of PEP. The indices obtained were the mean and SD of PEP and the average and SD of the time interval between two peaks of the ICG time-derivative (i.e., IBI of ICG). The peak detection mechanisms used to determine the peaks of BVP and ICG time-derivative were similar to the ECG R-peak detection algorithm, while additional heuristic rules were added to reduce the degradation of the signal quality due to motion artifacts and avoid spurious peak detection with backtracking. Unlike ECG signals, the peak amplitudes of PPG and ICG showed a larger deviation over a given period of time. An adaptive thresholding rule was integrated in the peak detection algorithm to address this deviation, which continuously changed/updated the threshold value to determine whether candidates for peaks qualified as valid peaks.

The heart sound signal measured sounds generated during each heartbeat. These sounds are produced by blood turbulence primarily due to the closing of the valves within the heart. The features extracted from the heart sound signal consisted of the mean and SD of the 3rd (138-275 Hz), 4th (69-138 Hz), and 5th (34-69 Hz) level coefficients of the Daubechies wavelet transform.

Electrodermal Activity

Electrodermal activity consists of two main components - tonic response and phasic response. Tonic skin conductance refers to the ongoing or the baseline level of skin conductance in the absence of any particular discrete environmental events. Phasic skin conductance refers to the event related changes that are caused by a momentary increase in skin conductance (resembling a peak superimposed on tonic skin conductance). The raw EDA signal was smoothed by a 25 ms moving average window and then down-sampled by 10 to remove the high frequency measurement noise. The phasic skin conductance detection algorithm used the following heuristics for considering a particular peak as a valid skin conductance response: (i) the slope of the rise to the peak should be greater than 0.05 microsiemens/minute (μS/min); (ii) the amplitude should be greater than 0.05 μS; and (iii) the rise time should be greater than 0.25 seconds. Once the phasic responses were identified, we determined the rate of the responses and the mean and maximum phasic amplitude. All the signal points that were not included in the response constituted the tonic part of the skin conductance signal. The slope of tonic activity was obtained using linear regression. Another feature derived from tonic response was the mean tonic amplitude.

Electromyogram Activity

EMG measures the electrical activity in the muscle during contraction. The EMG signal from corrugator supercilii muscle (eyebrow) captures a person's frown and detects the tension in that region, and the EMG signal from the zygomaticus major muscle captures the muscle movements while smiling. Upper trapezius muscle EMG activity measures the tension in the shoulders, one of the most common sites in the body for developing stress. Time-domain features, such as the mean, SD, and slope were calculated from the EMG signals after performing a band-pass filtering operation (10-500 Hz). The analysis of the EMG activities in the frequency domain involved applying Fast Fourier Transform (FFT) on a given EMG signal, integrating the EMG spectrum, and normalizing it to [0,1] to calculate the two features of interest - the median frequency and mean frequency for each EMG signal. The blink-related features were determined from the corrugator supercilii EMG signals after being preprocessed by a low-pass filter (10 Hz).

Peripheral Temperature

Variations in the peripheral temperature mainly come from localized changes in blood flow caused by vascular resistance or arterial blood pressure and reflect the autonomic nervous system activity. The signal was down-sampled by 10 and filtered to remove high-frequency noise, from which the mean, SD, and the slope were calculated as features.

Experimental Investigation

Participants

Due to the fact that autism is a spectrum disorder (American Psychiatric Association, 2000), no one intervention technique will work for the entire population (NRC, 2001; Sherer and Schreibman, 2005). The research on autism intervention assistive tools is generally guided by the individual characteristics, needs, and preferences of the children (i.e., individual-specific approach) and focuses on one sector of the population to develop a method with the flexibility to allow future modifications for a wider part of the population (Pioggia et al., 2005; Robins et al., 2005; Robins et al., 2004; Werry et al., 2001). The spectrum nature of autism and the phenomenon of person stereotypy (Lacey and Lacey, 1958) led us to choose an individual-specific approach to work on a long-term basis with a small group of children with autism in order to evaluate the affect recognition tool to be used in computer-assisted autism intervention.

Six participants in the age range of 13 to 16 years volunteered to participate in the experiments with the consent of their parents. Each had a diagnosis on the autism spectrum, either autistic disorder, Asperger's Syndrome, or pervasive developmental disorder not otherwise specified (PDD-NOS), according to their medical records. Participants were recruited using standard referral procedures that included (i) newsletters distributed through the Vanderbilt Treatment and Research Institute for Autism Spectrum Disorders, (ii) flyers placed in the Vanderbilt Center for Child Development, and (iii) website advertisements through the Vanderbilt Kennedy Center and the Autism Society of Middle Tennessee. The Institutional Review Board (IRB) approval was sought and received for conducting the experiment. Interested parents throughout middle Tennessee contacted the research office by phone or e-mail to set up an initial telephone screening. Monetary compensation (a $10 gift card per session) was given for the children's voluntary participation. Due to the nature of the designed cognitive tasks, the following criteria were considered when choosing the participants: (i) a minimum competency level of age-appropriate language and cognitive skills (i.e., "high functioning") and (ii) no history of mental retardation. Each child with ASD underwent the Peabody Picture Vocabulary Test III (PPVT-III) to assess cognitive function (Dunn and Dunn, 1997). The PPVT-III is a measure of single-word receptive vocabulary that is often used as a proxy for intelligence quotient (IQ) testing because of its high correlations with standardized tests such as the Wechsler Intelligence Scale for Children (Bee and Boyd, 2004). It provides standard scores with a mean of 100 and a standard deviation of 15, and the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV; American Psychiatric Association, 2000) classifies full scale IQ's above 70 as non-retarded. Participants in our study obtained a standard score of 80 or above on the PPVT-III measure. Table 2.2 shows the characteristics of the six children who participated in the experiments.

Table 2.2 The characteristics of the participants

Child ID	Gender	Age	Diagnosis	PPVT-III Score
A	Male	15	Autistic Disorder	99
B	Male	15	Asperger's Syndrome	80
C	Male	13	Autistic Disorder	81
D	Male	14	PDD-NOS	92
E	Male	16	PDD-NOS	93
F	Female	14	PDD-NOS	83

Several conditions posed challenges in recruitment of participants who matched the inclusion-exclusion criteria (e.g., cognitive skills and age range) and in coordination of schedules between the autism therapist and the parent of the participating child who were also involved in the experiment. First, autism may often co-occur with varying levels of mental retardation (American Psychiatric Association, 2000), which reduces the possible participant pool. Second, the IRB stipulates cutoffs between participants in different age ranges (e.g., 7-12 years, 13-17 years, 18 years and above, etc.), and autism intervention studies usually focus on one sector of the population within a certain age range (Gaylord-Ross et al., 1984; IRB, 2004; Parsons et al., 2005). Third, the responsibilities of raising a child with ASD are vast; therefore, willing parents often had to bring their child to the laboratory on weekends or after school on days without conflicts with other activities (e.g., social skills therapy) or family obligations. The group sizes and the cardinality of participant age range of many studies on computer-assisted autism intervention are commensurate with our work when an individual-specific approach was used (Pioggia et al., 2005; Robins et al., 2005; Robins et al., 2004; Werry et al., 2001). It is worth noting that this individual-specific study was based on a large sample size of observations for each child with ASD, which is comparatively more favorable than many other works (Groden et al., 2005; Pioggia et al., 2005; Robins et al., 2004). Each child completed approximately 85 epochs over 6 sessions, which represents 6 different days and yields 6 hours of data for each child. This preliminary study focused on high-functioning children with ASD between 13 and 16 years old.

Cognitive Tasks for Affect Elicitation

Two computer-based cognitive tasks were designed and implemented to invoke varying intensities of the following three affective states: anxiety, engagement, and enjoyment/liking, in the participants. Physiological data from participants were collected during the experiment. The two tasks consisted of an anagram solving task and a Pong playing task. The anagram solving task has been previously employed to explore relationships between physiology and anxiety (Pecchinenda and Smith, 1996). Emotional responses were manipulated in this task by presenting the participant with anagrams of varying difficulty levels, based on vocabulary tests and established through pilot work. Five-letter words were collected from the Dolch sight word list (Dolch, 1948) and grade-level vocabulary tests (Barnhart and Barnhart, 1984) and used to create the different levels of difficulty in the anagram task. A long series of trivially easy anagrams caused less engagement. An optimal mix of solvable and difficult anagrams caused liking and engagement at times. Unsolvable or extremely difficult anagrams and giving time deadlines generated anxiety.

The Pong task consisted of a series of trials/epochs each lasting up to four minutes, in which the participant played a variant of the early, classic video game "Pong." This game has been used previously by researchers to study anxiety, performance, and gender differences (Brown et al., 1997). Various parameters of the game were manipulated to elicit the required affective responses. These included: ball speed and size, paddle speed and size, sluggish or over-responsive keyboard, random keyboard response, and the level of the computer opponent player. Very low speeds and large sizes of the ball and the paddle made games less engaging after a while; whereas high ball and paddle speeds along with smaller sizes of the two made the game engaging. Very high ball speeds and sluggish or over-responsive keyboard caused anxiety at times. Games with a moderate-level computer opponent player usually generated liking. The game configurations were established through pilot work.

Each task sequence was subdivided into a series of discrete epochs that were bounded by the subjective affective state assessments. These assessments were collected using a battery of questions about the target affective states and perceived task difficulty level rated on an eight-point Likert scale, where 1 indicated the lowest rating and 8 indicated the maximum rating. Each participant took part in six sessions – three one-hour sessions of solving anagrams and three one-hour sessions of playing Pong – on six different days. No more than one, one-hour session with an individual participant took place per day.

Experimental Setup

Figure 2.2 shows the setup for the experiment. The child with ASD was involved in the cognitive tasks on computer C1 while his/her physiological data was acquired via the Biopac system (biopac.com). Physiological signals were transferred from the Biopac transducers to C2 through an Ethernet link at 1000 Hz after being amplified, digitized, and stored. C1 was also connected to the Biopac system via a parallel port, through which

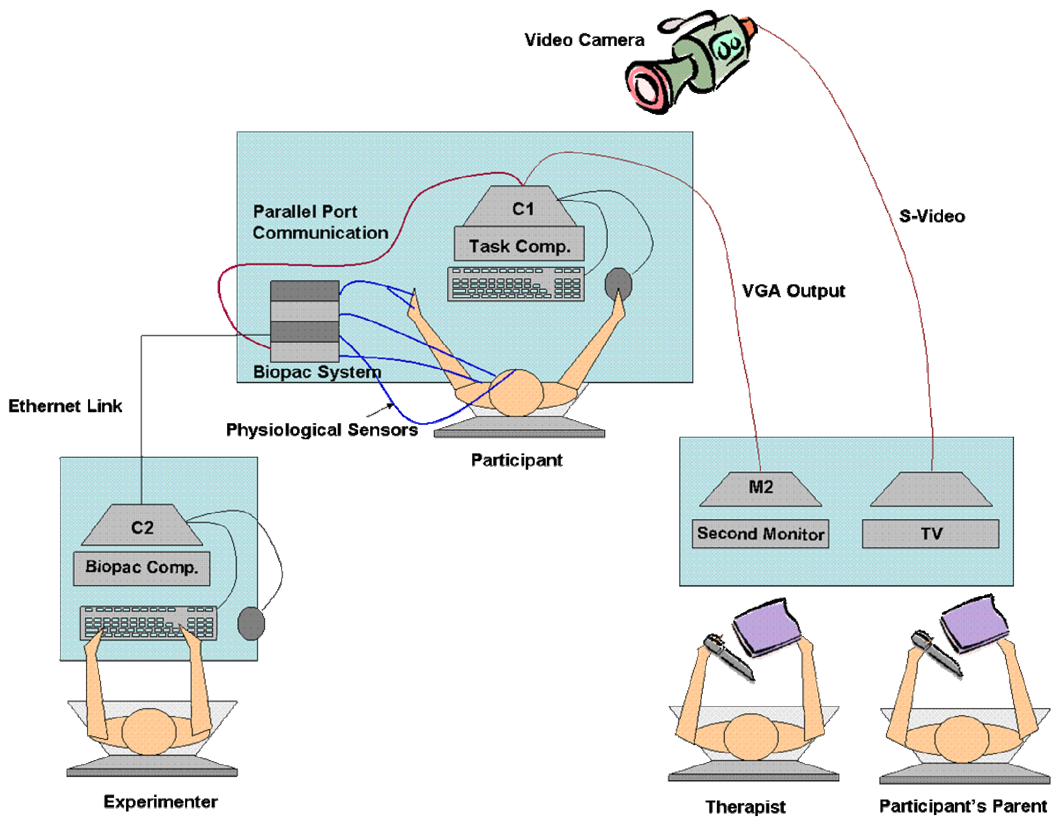


Figure 2.2 Experimental setup for collecting physiological data and subjective reports in the computer-based tasks

the task-related markers were recorded along with the physiological data in a time-synchronized manner. Different markers were defined to indicate the following events: start/end of game, performance events (right/wrong answer in anagram, hitting/missing ball in Pong), start/end of each epoch, and self-report logging.

To gain perspective from different sources and enhance the reliability of the subjective reports on the target affective states, a therapist with experience in autism intervention for children with ASD and each participant's parent were also involved in the study, who may best know the participant. We video recorded the sessions to cross-reference observations made during the experiment. The signal from the video camera was routed to a television, and the signal from the participant's computer screen where the task was presented was routed to a separate computer monitor M2. The therapist and the participant's parent were seated at the back of the experiment room, watched the experiment on the TV from the view of the video camera, and observed how the task (anagrams or Pong) progressed on the separate monitor.

Procedure

On the first visit, participants completed the PPVT-III measurement to determine a standardized measure of receptive vocabulary and eligibility for the experiments. After initial briefing regarding the tasks, physiological sensors from a Biopac system were attached to the participant's body and a three-minute baseline recording was performed. Each session lasted about an hour and consisted of a set (13-15) of either 3-minute epochs for anagram tasks or up to 4-minute epochs for Pong tasks. Each epoch was followed by subjective report questions rated on an eight-point Likert scale. The participants reported their perceived subjective affective states through a pop-up dialog window presented on C1. The therapist and the participant's parent also answered the questions about how they thought the participant was feeling during the finished epoch on an eight-point Likert scale based on their audio/visual observations from the viewing monitors (TV and M2). These three sets of subjective reports related to the target affective states, from the therapist, the participant's parent, and the participant, were used as the possible reference points to link the concurrently collected objective physiological data to the participant's affective states.

For developing affective models, we built mappings to determine the intensity (i.e., high/low) of a particular affective state from the physiological features. It resembles a binary classification problem where the attributes are the physiological features (listed in Table 2.1) and the target function is the degree of arousal. In this work we employed SVM to determine the underlying affective state of a child with ASD given a set of physiological indices, based on previous work (Rani et al., 2006) which showed SVM gave the best classification accuracy compared to KNN, Bayesian Network, and Regression Tree as applied to the domain of affect recognition using physiological signals for typical adults. Details of the theory and learning method of SVM can be found in (Vapnik, 1998) and are briefly described in the Appendix A.

Each participant had a data set that was comprised of both the objective physiological features and corresponding subjective reports on intensity of target affective states from the therapist, the participant's parent, and the participant. The subjective report forms instructed that 1-4 indicates the low level, 5-8 indicates the high level, and the different values represent the variation within each level. The physiological features were extracted using the approaches described previously in the Physiological Signal Acquisition and Indices section. Each feature was baselined to take into account day-variability, which is not commonly performed when testing statistically significant changes in physiological signals (Groden et al., 2005; Mandryk et al., 2006; Neumann and Waldstein, 2001) but has shown to be a useful technique when building physiology-based affective models (Picard et al., 2001; Rani et al., 2005; Zhai and Barreto, 2006). The following equation from Rani et al. (2005) was employed:

$$F_{baselined_{ijk}} = \frac{F_{epoch_{ijk}} - F_{base_{ik}}}{F_{base_{ik}}} \tag{2.1}$$

Each feature i (i = 1,2,...51; listed in Table 2.1) obtained in epoch j (j=1,2,...approximately 15), $F_{epoch_{ijk}}$, was baselined with respect to the feature obtained during the baseline recording for the corresponding day k (k=1,2,...6), $F_{base_{ik}}$. If $F_{base_{ik}}$ is equal to zero then the respective value of $F_{epoch_{ijk}}$ is automatically baselined. Also, each subjective report was scaled to [0, 1] using the following min-max equation (Zhai and Barreto, 2006).

$$R_{scaled_{ae}} = \frac{R_{epoch_{ae}} - R_{min_{ae}}}{R_{max_{ae}} - R_{min_{ae}}} \tag{2.2}$$

where the minimum subjective rating of each affective state a (a = anxiety, engagement, or enjoyment/liking) over all of a participant's epochs e (e=1,2,...approximately 85), $R_{min_{ae}}$, is subtracted from each individual epoch rating, $R_{epoch_{ae}}$, and divided by the difference between the maximum rating from the corresponding reporter of the affective state, $R_{max_{ae}}$, and the minimum rating, $R_{min_{ae}}$. All three affective states were partitioned so that there were two levels for each affective state. The scaled report ratings were discretized such that 0–0.50 was labeled as low level and 0.51–1 was labeled as high level. The reports from each rater (therapist, participant's parent, or child) for each of the three affective states were partitioned separately.

Each participant's data set contained approximately 85 epochs. Multiple subjective reports were analyzed, and one was chosen as the possible reference points to link the physiological measures to the participant's affective state. As illustrated in Figure 2.3, a therapist-like affect recognizer (i.e. a recognizer that captures the therapist's ability

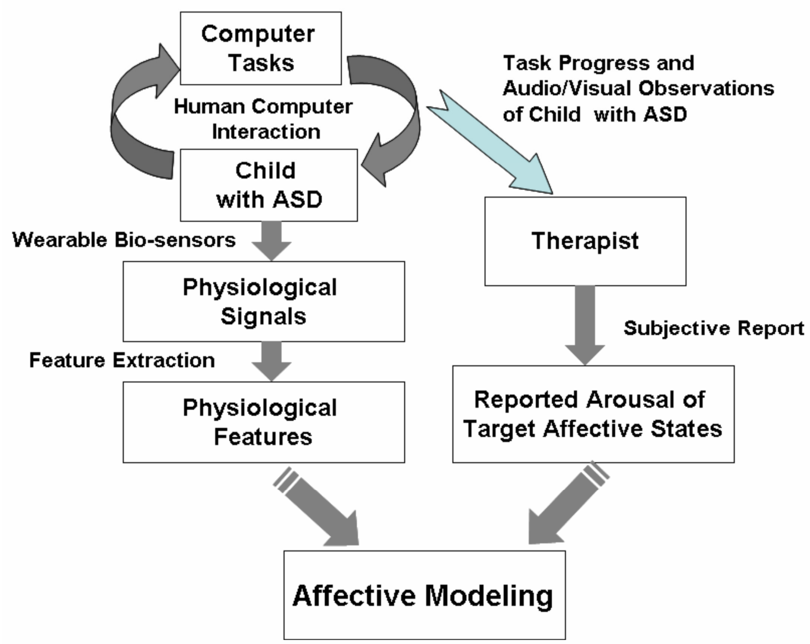


Figure 2.3 Affective modeling overview when the therapist's subjective reports are used

to assess affective states) can be developed when the therapist's reports are used. Current therapeutic settings do not retain quantitative records of the affective states of the children with ASD. A therapist generally uses qualitative affective evaluations suitable for binary (high/low) assessments to make intervention adjustments (e.g., using likes/dislikes charts (Seip, 1996). This study of differentiating high/low levels of the target affective states from physiological signals attempts to emulate the present autism intervention practice and to experimentally demonstrate the feasibility of affective modeling for these children with ASD via psychophysiological analysis.

A SVM-based recognizer was trained on each participant's data set for each target affective state. In this work, to deal with the nonlinearly separable data, soft-margin classifiers with slack variables were used to find a hyperplane with less restriction (Eqn. A.1, Appendix A). RBF (Radial Basis Function) was selected as the kennel function because it often delivers better performance (Vapnik, 1998). A ten-fold cross-validation was used to determine the kernel parameter and regularization parameter (Eqn. A.2, Appendix A) of the classifier.

Results

To measure the amount of agreement among the different reporters, the kappa statistic was used (Siegel and Castellan, 1988). The kappa coefficient (*K*) measures pair-wise agreement among a set of reporters making category judgments, correcting for expected chance agreement:

$$K = \frac{P(A) - P(E)}{1 - P(E)} \tag{2.3}$$

where $P(A)$ is the proportion of times that the reporters agree and $P(E)$ is the proportion of times that we would expect the reporters to agree by chance. When there is complete agreement, then $K = 1$; whereas, when there is no agreement other than that which would be expected by chance, then $K = 0$.

The results of the values of *K*, averaged across three target affective states, are shown in Figure 2.4. From the results, we can see that among the three possible pairs for each child (Therapist-Parent (T/P), Therapist-Child (T/C), and Parent-Child (P/C)) the agreement between the therapist and each participant's parent (T/P) shows the largest mean of the *K* statistic values (mean = 0.62, $p < 0.05$, paired *t*-test). The means of the *K*-statistic values between the children and either the parent or the therapist were relatively small (0.40 and 0.37, respectively). Note that the *K* agreement between therapist and

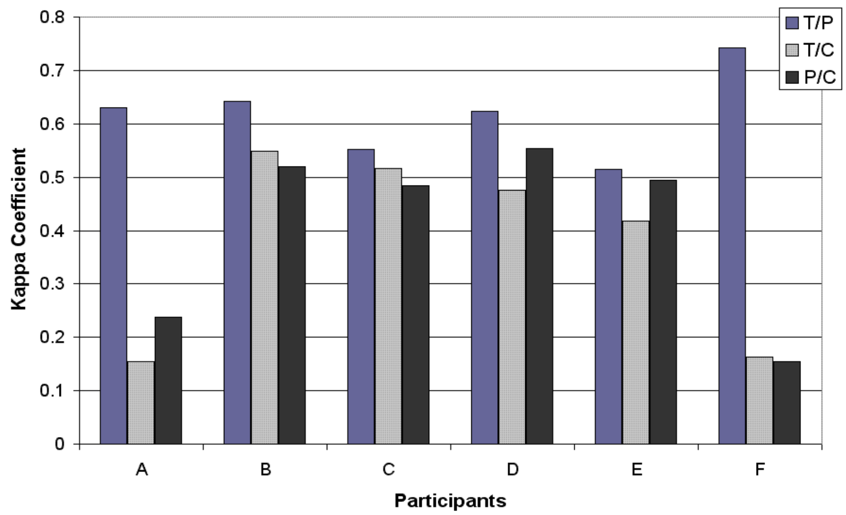


Figure 2.4 Average Kappa Statistics between Reporters for Affective States. Kappa coefficients averaged across affective states measure the agreement between the different subjective reports (T-Therapist, P-Parent, C-Child) corresponding to each participant (Child ID A-F).

parent is substantial for Child A, Child B, Child D, and Child F and moderate for Child C and Child E. Such results might stem from the fact that it could be difficult for the therapist or parent to distinguish certain emotions for a particular child with ASD. For example, the agreement between therapist and parent for the anxiety level of Child C and Child E (K coefficient: 0.35 and 0.37, respectively) are considerably less than the average level (mean K coefficient of T/P: 0.62). In the experiment, Child A and Child F's self-ratings for anxiety, engagement, and liking were almost constant which resulted in lower K statistic values for the therapist and child pair (T/C) and the parent and child pair (P/C) than those of the other participants. This may be due to the fact that the spectrum developmental disorder for children with autism manifests in different abilities to recognize and report their own emotions. Lack of agreement with adults does not necessarily mean that the self-report of children with ASD is not dependable. However, given the fact that therapists' judgment based on their expertise is the state-of-the-art in most autism intervention approaches and the reasonably high agreement between the therapist and the parents for all of the six children, the subjective report from the therapist was used as the reference points linking the objective physiological data to the children's affective states. To enhance the consistency of the subjective reports, the same therapist was involved in all of the sessions. This choice allows for building a *therapist-like* affective model. In the rest of the chapter, unless otherwise specified, the terms liking, anxiety, and engagement imply the target affective states as discerned by the therapist. Once the affect modeling is completed, the recognizers will be capable of autonomously inferring the affective states of the child with ASD from the physiological signals in real time even when the therapist is not available, which is investigated in Chapter III.

Figure 2.5 shows a comparison of the therapist's average ratings for liking, anxiety, and engagement when the children with ASD played easy or difficult epochs in the anagram and Pong computer games, and the small bars indicate the standard error of the mean. When averaged across all participants, liking decreased, anxiety increased, and engagement decreased with increasing task difficulty. Table 2.3 shows the correlation

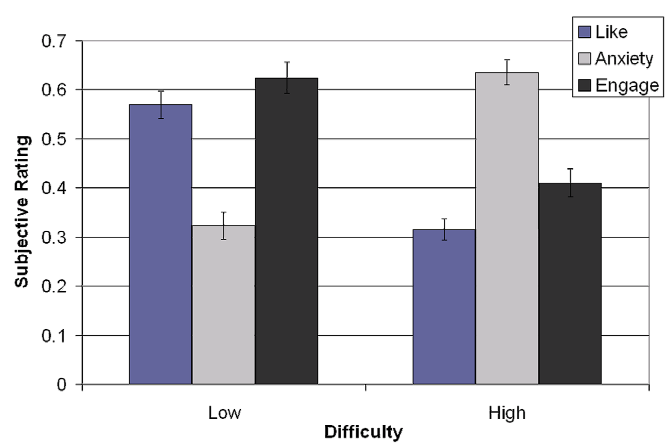


Figure 2.5 Rated average affect response from therapist's reports

Table 2.3 Results of Correlation Analysis from Therapist's Reports

	Anxiety	Engagement	Difficulty
Liking	-0.521	0.885	-0.616
Anxiety		-0.401	0.731
Engagement			-0.486

analysis between the reported affective states and the task difficulty. For each set of variables, the probability value (p-value) was computed from a two side t-test. Due to the large sample size (approximately 85 epochs for each participant), the p-value for all correlations was less than 0.005. Through point biserial correlation analysis, it was found that difficulty is strongly positively correlated with anxiety and negatively correlated with liking and engagement. By examining Pearson correlation coefficients (r), it was observed that there is strong positive correlation between liking and engagement and negative correlation between liking and anxiety, and there also exists a weak correlation between the reported anxiety and engagement. The results in Figure 2.5 and Table 2.3 present findings across all the children. However, when each child is examined individually, different trends could arise. For example, for Child A, anxiety is positively correlated with engagement ($r = 0.45$); for Child F, no significant correlation is observed ($r = -0.15$, $p > 0.05$); while for the four other children (B, C, D, and E) anxiety negatively correlated with engagement (r equals -0.50, -0.39, -0.61, and -0.58, respectively), which revealed diverse affective characteristics of the children with ASD.

The performance of the developed affective models based on the therapist's reports for each child (i.e., individual-specific approach) is shown in Figure 2.6. The cross-validation method, “leave-one-out,” was used, and the small bars represent the measured mean error rate of the trained machine, also called the empirical risk (Burges,

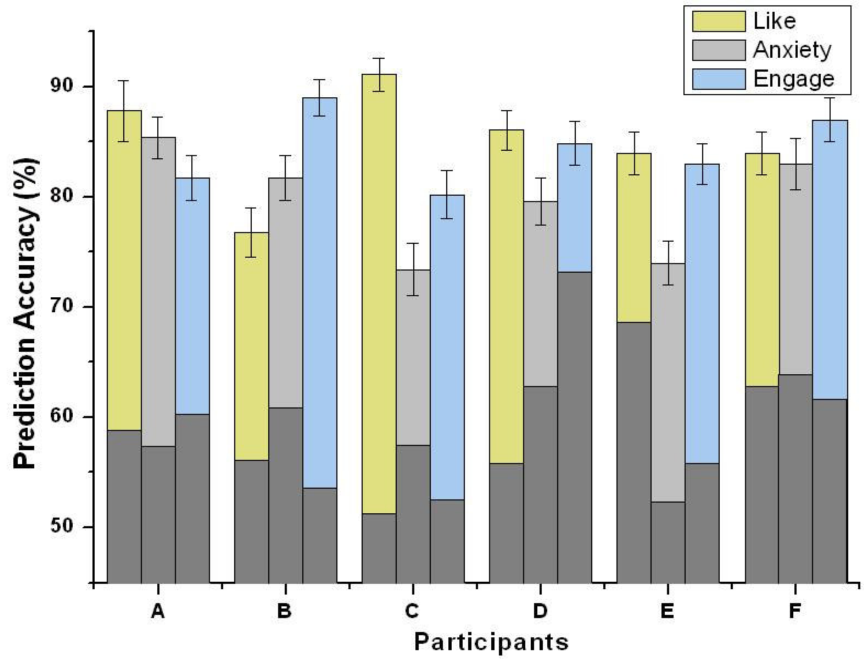


Figure 2.6 Prediction Accuracy of the Affective Model

1998). The affective model produced high recognition accuracies for each target affective state of each participant. The average correct prediction accuracies across all participants with ASD were: 85.0% for liking, 79.5% for anxiety, and 84.3% for engagement, which are comparable to the best results achieved for typical adults (Nasoz et al., 2004; Picard et al., 2001; Rani et al., 2006). Figure 2.6 shows that for Child C and Child E, the prediction accuracy for anxiety is lower; moreover, as mentioned previously for these two participants there is also considerably less agreement between the therapist and the parent (T/P) on the subjective reports with respect to the anxiety level. The comparatively low (approximately 5% less) average prediction accuracy of anxiety may be due to the fact that the intensity of anxiety of a particular child with ASD (e.g., Child C and Child E) could be more difficult for the therapist to distinguish based on the observations than the other two affective states (i.e., liking and engagement).

We also compared the performance of affective modeling to a control method that represents random chance. Suppose we had an equal number of subjective reports that rated a particular affective state level (low/high) for a participant, then the chance probability would be 50%. However, the prevalence of each level could be different. For example, in 48 out of 86 epochs the engagement of Child E was rated as low, where a random classification could assign all test epochs to this category and make accurate classifications $(48/86)\times100 = 55.8\%$ of the time. We thus considered the level with a majority of epochs and used the average of these higher numbers (across the participants' affective states) to represent the chance condition, which is denoted by dark gray bars in Figure 2.6. While the physiology-based affective modeling alone did not provide perfect classification (i.e., 100%) of affective states of children with ASD, they did yield reliable matches with the subjective rating and significantly outperformed a random classifier (averaging 82.9% vs. 59.2%). This was promising considering that this task was challenging in two respects: (i) the reports were collected from the therapist who was observing the children with ASD as opposed to having typical adults capable of differentiating and reporting their own affective states and (ii) varying levels of arousal of any given affective state (e.g., low/high anxiety) were identified instead of determining between two discrete affective states.

To explore the effects of reducing the number of physiological signals and the possibility of achieving more economical modeling (i.e., reducing the set of signals to be measured), we examined the performance of the affect recognizers when cardiovascular, electrodermal, and electromyographic activities and their combinations were used. As shown in Table 2.4, all the recognizers delivered better predication than random guess

Table 2.4 Prediction Accuracy of the Affective Modeling based on Different Physiological Signals (%)*

Physiological Signals	Liking	Anxiety	Engagement	Mean
Cardiovascular	75.7	68.5	76.2	73.5
Electrodermal	73.4	72.3	73.3	73.0
Electromyographic	73.1	65.8	70.1	69.7
Electrodermal + Electromyographic	75.0	69.4	71.4	71.9
Cardiovascular + Electromyographic	79.6	70.2	79.9	76.6
Cardiovascular + Electrodermal	79.9	74.3	81.9	78.7
All	85.0	79.5	84.3	82.9

(mean prediction rate: 52.9%), and with more information from physiological activities the performance of the affective models tends to improve (except the combination of electrodermal and electromyographic activities). The improved prediction accuracy of the models with increased physiological features may be due to the fact that the inherent kernel representation and soft-margin optimization endow SVM the capability to work effectively in the high-dimensional feature space (Burges, 1998). While electromyographic/EMG signals have been used as indicators of affective response for typical individuals (Kulic and Croft, 2007; Rani et al., 2006), in this study we observed that it is less discriminatory than the cardiovascular and electrodermal activities. As suggested by the American Psychiatric Association (2000) and Green et al. (2002), children with ASD often have nonverbal communicative impairments regarding expression of affective states (e.g., abnormal body postures and gestures and absence of

* Peripheral temperature has relatively few features derived as shown in Table 2.1 and was not examined independently. Instead, it was studied conjunctively with the electrodermal activity, both of which were acquired from the non-dominant hand of a participant.

facial expression), which might reduce the discriminatory capability of EMG signals (e.g., muscle activities from both the corrugator supercilii and the zygomaticus major) to reveal affective cues of the participants. While no combination of physiological activity surpassed the percent accuracy achieved when all signals were used, the results in Table 2.4 suggest that it may be possible to selectively reduce the set of signals and obtain nearly-as-good performance (e.g., using a combination of cardiovascular and electrodermal signals).

With post-hoc analysis, we found the prediction accuracy generally tends to be higher when the therapist and the participant's parent agree more on the subjective reports about how they thought the participant was feeling during the finished epoch. As shown in Table 2.5, the *K* statistic of the therapist and parent is positively correlated with the prediction accuracy of the developed affect recognizer (r = 0.71, $p < 0.001$). In this experiment, the *K* statistic could indicate whether it is relatively easy or difficult to

Table 2.5 Therapist-Parent (T/P) Kappa Statistics and Prediction Accuracy

Child ID		Liking	Anxiety	Engagement
A	Kappa Statistics (T/P)	0.566	0.831	0.494
	Prediction Accuracy (%)	87.8%	85.4%	81.7%
B	Kappa Statistics (T/P)	0.585	0.634	0.708
	Prediction Accuracy (%)	76.8%	81.7%	89.0%
C	Kappa Statistics (T/P)	0.753	0.352	0.551
	Prediction Accuracy (%)	91.1%	73.4%	80.2%
D	Kappa Statistics (T/P)	0.698	0.562	0.611
	Prediction Accuracy (%)	86.1%	79.6%	84.9%
E	Kappa Statistics (T/P)	0.721	0.372	0.449
	Prediction Accuracy (%)	83.7%	74.1%	83.2%
F	Kappa Statistics (T/P)	0.884	0.528	0.814
	Prediction Accuracy (%)	84.8%	82.6%	87.2%

differentiate the affective states of a child by observation. The autism therapist used in this work had no previous interaction with the participants. The prediction accuracy could likely improve if the therapist interacts with a particular child with ASD for a significant amount of time and gains more knowledge of his/her affective expression before making the reports regarding the presented interaction tasks, which is generally the case for ASD intervention.

Discussion

We have designed and implemented two computer-based cognitive tasks – solving anagrams and playing Pong – to elicit the affective states of liking, anxiety, and engagement, which considered important in autism intervention, for children with ASD. To have reliable reference points to link the physiological data to the affective states, the reports from the child, the therapist, and the child's parent were collected and analyzed. We have investigated a large set of physiological indices that may correlate with the affective states of children with ASD. A SVM-based affective model yielded reliable prediction with approximately 82.9% success when using the therapist's reports. This is the first time, to our knowledge, that the affective states of children with ASD have been experimentally detected via a physiology-based affect recognition technique.

It should be noted that due to the phenomenon of person stereotypy and the spectrum nature of autism, an individual-specific approach has been employed for affective modeling based on a large sample size of observations of each of the six participating children with ASD. The methodology for inducing, gathering, and modeling the experimental data in this chapter is not dependent on the participants. The group sizes and the cardinality of participant age range of many related studies are commensurate with our work and the sample size of observations in this work is comparatively extensive. The consistently reliable prediction accuracy for each participant demonstrated that it was feasible to model the affective states of these children with ASD via psychophysiological analysis.

CHAPTER III

AFFECT-SENSITIVE ADAPTATION DURING HUMAN-ROBOT INTERACTION WITH CHILDREN WITH AUTISM SPECTRUM DISORDERS

Introduction

The primary objective of this chapter is to investigate how to augment HRI to be used in autism intervention by endowing a robot with the ability to recognize and respond to the affective states of a child with ASD. The work in Chapter II developed affective models through psychophysiological analysis. This chapter employs those models to investigate affect sensitivity during the closed-loop interaction between a child with ASD and the robot. A proof-of-concept experiment was designed wherein a robot learns individual preferences based on the predicted liking level of the children with ASD as discerned by the therapist and selects an appropriate behavior accordingly.

Once affective modeling was completed in Chapter II, the therapist-like recognizers could be applied to a robot, equipping it with the capability to detect the affective states of the children with ASD in real time from on-line extracted physiological features, which could be utilized in future interventions even when a therapist is not available. As stated by Dautenhahn et al. (2003), it is important to have robots maintain characteristics of adaptability when applied to autism intervention. We designed and implemented a proof-of-concept experiment (i.e., robot-based basketball) wherein a robot adapts its behaviors in real time according to the preference of a child with ASD, inferred from the interaction experience and the predicted consequent liking level. This work is the first time, to our knowledge, that the feasibility and the impact of affect-sensitive closed-loop interaction between a robot and a child with ASD have been demonstrated experimentally. While the results are achieved in a non-social interaction task, it is expected that the real-time affect recognition and response system described in this work will provide a basis for future research into developing technology-assisted intervention tools to help children with ASD explore social interaction dynamics in an affect-sensitive and adaptive manner.

The overview of the affect-sensitive closed-loop interaction between a child with ASD and a robot is presented in Figure 3.1. The physiological signals from the children with ASD are recorded when they are interacting with the robot. These signals are processed in real time to extract features, which are fed as input into the models developed in Chapter II. The models determine the perceived affective cues and return this information as an output. The affective information, along with other environmental inputs, is used by a controller to decide the next course of action for the robot. The child who engages with the robot is then influenced by the robot's behavior, and the closed-loop interaction cycle begins anew.

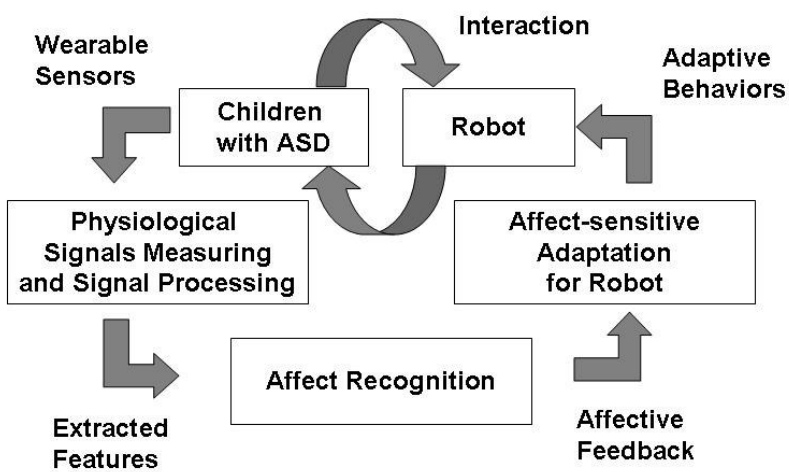


Figure 3.1 Framework overview

While the framework is implemented during affect-sensitive HRI, the affective models were built using physiological data gathered from two HCI tasks. Agrawal et al. (2008) showed that affective models built through HCI tasks could be successfully employed to achieve affect recognition in HRI for typical individuals. This observation suggests that it is possible to broaden the domain of tasks for affective modeling, thus reducing the habituation effect due to continuous exposure to the same system.

Experimental Investigation

Task Design

A closed-loop HRI task, RBB (robot-based basketball), was designed. The main objective was two-fold: (i) to enable the robot to learn the preference of the children with ASD implicitly using physiology-based affective models as well as select appropriate behaviors accordingly; and (ii) to observe the effects of such affective-sensitivity in the closed-loop interaction between the children with ASD and the robot.

The affective models developed in Chapter II are capable of predicting the intensity of liking, anxiety, and engagement simultaneously. However to designate a specific objective for the experiment in this chapter without compromising its proof-of-concept purpose, one of the three target affective states was chosen to be detected and responded to by the robot in real time. As has been emphasized by Dautenhahn and Werry (2004), the liking of the children (i.e., the enjoyment they experience when interacting with a robot) is a goal as desirable as skill learning for autism intervention.

Therefore, liking was chosen as the affective state around which to modify the robot's behaviors in the HRI experiment.

In the RBB task, an undersized basketball hoop was attached to the end-effector of a robotic manipulator, which could move the hoop in different directions (as shown in Fig. 3.2) with different speeds. The children were instructed to shoot a required number of baskets into the moving hoop within a given time. Three robot behaviors were designed as shown in Table 3.1. For example, in behavior 1 the robot moves towards and away from the participant (i.e., in the *X*-direction) at a slow speed with soft background music, and the shooting requirement for successful baskets is relatively low. The parameter configurations were determined based on a pilot study to attain varied impacts on affective experience for different behaviors. From this pilot study, the averaged performance of participants for a given behavior was compiled and analyzed. The threshold of shooting requirement (TSR) was defined as 10% lower than the average

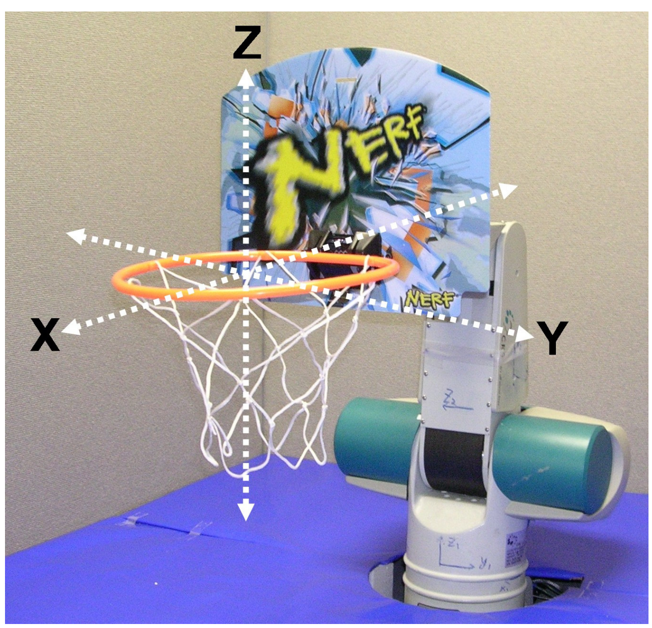


Figure 3.2 *X*-, *Y*-, and *Z*-directions for behaviors used in RBB

Table 3.1 Robot Behaviors

Behavior ID	Motion Direction	Speed (seconds/period)	Threshold (baskets/epoch)	Background Music
1	X	8	12	Serene
2	Y	4	20	Lively
3	Z	2	30	Irregular

performance. At the end of each epoch, the participant's performance was rated as excellent (successful baskets ≥ 1.2*TSR), above average (0.8*TSR ≤ successful baskets < 1.2*TSR), or below average (successful baskets < 0.8*TSR). Behavior transitions occurred between but not within epochs. As such, each robot behavior extended for the length of an epoch (1.5 minutes in duration) to have the participant fully exposed to the impact of that behavior.

Each of the six participants, who also participated in the experiment in Chapter II, took part in two robot basketball sessions (RBB1 and RBB2). In RBB1 (non-affect based) the robot selected its behavior randomly (i.e., without any regard to the liking information of the participant), and the presentation of each type of behavior was evenly distributed. This session was designed for two purposes: (i) to explore the state space and action space of the QV-learning algorithm used in RBB2 for behavior adaptation; and (ii) to validate that the different robot behaviors have distinguishable impact on the child's level of liking. In RBB2 (liking-based), the robot continues to learn the child's individual preference and selects the desirable behavior based on interaction experiences (i.e., records of robot behavior and the consequent liking level of a participant predicted by the affective model). The idea is to investigate whether the robot can autonomously choose the most-liked behavior of each participant as observed from RBB1 by means of employing the physiology-based affective model and QV-learning.

Experimental Setup

The real-time implementation of the robot-based basketball system is shown in Figure 3.3. The setup included a 5 degree-of-freedom robot manipulator (CRS Catalyst-5 System) with a small basketball hoop attached to its end-effector. Two sets of infrared (IR) transmitter and receiver pairs were attached to the hoop to detect small, soft foam balls going through the hoop. The set-up also included the biological feedback equipment (Biopac system) that collected the participant's physiological signals and the digital output from the IR sensors. The Biopac system was connected to a PC (C1) that: (i) acquired physiological signals from the Biopac system and extracted physiological features on-line, (ii) predicted the probable liking level by using the affective model

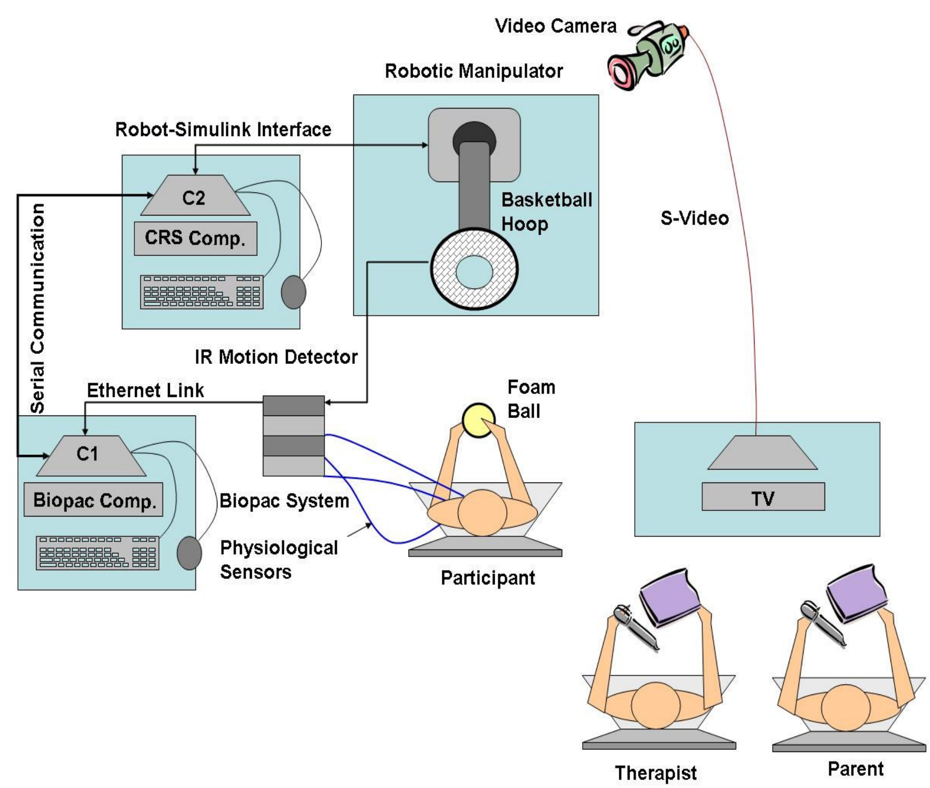


Figure 3.3 Experimental setup for robot basketball

developed in Chapter II, (iii) acquired IR data through the analog input channels of the Biopac system, and (iv) ran a QV-learning algorithm that learns the participant's preference and chooses the robot's next behavior accordingly. Computer C1 was connected serially to the CRS computer (C2), which ran Simulink software. The behavior switch triggers were transmitted from C1 to C2 via a RS232 link. The commands to control the robot's various joints were transmitted from C2 to the robot. There was a communication protocol established between C1 and C2 that ensured the beginning and end of the basketball task was appropriately synchronized with the physiological data acquisition on C1. As in the HCI tasks in Chapter II, the therapist and a parent were also involved, watching the experiment from the TV that was connected to a video camera.

Experimental Procedure

Each basketball session (RBB1 or RBB2) was approximately 1 hour long and included 27 minutes of active HRI (i.e., 18 epochs of 1.5 minutes each). The remaining time was spent attaching sensors, guiding a short practice, taking a baseline physiological recording, collecting subjective reports, and pausing for scheduled breaks. During the experiment, the participant was asked to take a break after every four epochs and the participant could request a break whenever he/she desired one. During each RBB epoch, the participant received commands and performance assessments from pre-recorded dialogue via a speech program running on C1 and the interaction proceeded as follows:

1. The participant was notified of the goal (i.e., TSR).
2. A start command instructed the participant to start shooting baskets.

3. Once the epoch started, the participant was given voice feedback every 30 seconds regarding the number of baskets remaining and the time available.
4. A stop command instructed the participant to stop shooting baskets, which ended the epoch.
5. At the end of each epoch, the participant's performance was rated and relayed to him/her as excellent, above average, or below average.

Each epoch was followed by subjective reports that took 30-60 seconds to collect. The subjective assessment procedure was the same as the protocol used in the affective modeling tasks in Chapter II. After the subjective report was complete, the next epoch would begin. To prevent habituation, a time interval of at least seven days between any two RBB sessions was enforced.

Affect-sensitive Behavior Adaptation in Closed-Loop HRI

We defined the state, action, state transition, and reward functions so that the affect-sensitive robot behavior adaptation problem could be solved using the QV-learning algorithm as described by Wiering (2005) and Appendix B. The set of states consisted of three robot behaviors as described in Table 3.1. In every state, the robot has three possible actions (1/2/3) that correspond to choosing behavior 1, 2, or 3, respectively, for the next time step (i.e., next epoch). Each robot behavior persists for one full epoch and the state/behavior transition occurs only at the end of an epoch. The detection of consequent affective cues (i.e., the real-time prediction of the liking level for the next epoch) was used to evaluate the desirability of a certain action. To have the robot adapt to a child's individual preference, a reward function was defined based on the predicted liking level. If the consequent liking level was recognized as high, the contributing action was interpreted as positive and a reward was granted (reward = 1); otherwise the robot received a punishment (reward = -1). QV-learning uses this reward function to have the robot learn how to select the behavior that was expected to result in a high liking level and therefore positively influenced the actual affective (e.g., liking) experience of the child.

RBB1 enables state and action exploration where the behavior-switching actions are chosen randomly, with the number of visits to each state evenly distributed. The V-function and Q-function are updated using Eqn. B.1 and Eqn. B.2 from Appendix B. After RBB1, the subjective reports are analyzed to examine the impacts of different behaviors on each participant's preference. In RBB2 the robot starts from a non-preferred behavior/state and continues the learning process by using Eqn. B.1 and Eqn. B.2. A greedy action selection mechanism is used to choose the behavior-switching action with the highest Q-value.

Because of the limited number of states and actions in this proof-of-concept experiment, tabular representation is used for the V-function and the Q-function. To prevent a certain action and/or state from being overly dominant and to counteract the habituation effect, the values of $Q(s, a)$ and $V(s)$ are bounded by using the reward or punishment encountered in the interaction. The parameters in Eqn. B1 and Eqn. B.2 are chosen as $\alpha = 0.8$ and $\gamma = 0.9$. Before RBB1 begins, the initial values in the V-table and the Q-table are set to 0.

Results

The six children with ASD who completed the HCI experiment in Chapter II also took part in the robot basketball task. The results described here are based on the RBB1 (non-affect based) task and the affect-sensitive closed-loop interaction between the children with ASD and the robot in the RBB2 (liking-based) task.

First, results are presented to validate that different behaviors of the robot had distinguishable impacts on the liking level of the children with ASD. To reduce the bias of validation, in RBB1 the robot selects behaviors randomly and the occurrence of each behavior is evenly distributed. Figure 3.4 shows the average labeled liking level for each behavior as reported by the therapist in RBB1. The difference of the impact is significant for five children (participants A, B, D, E, and F) and moderate for participant C. By performing two-way ANOVA analysis on the behavior (i.e., most-preferred, moderately-preferred, and least-preferred behavior) and participant, it was found that the differences

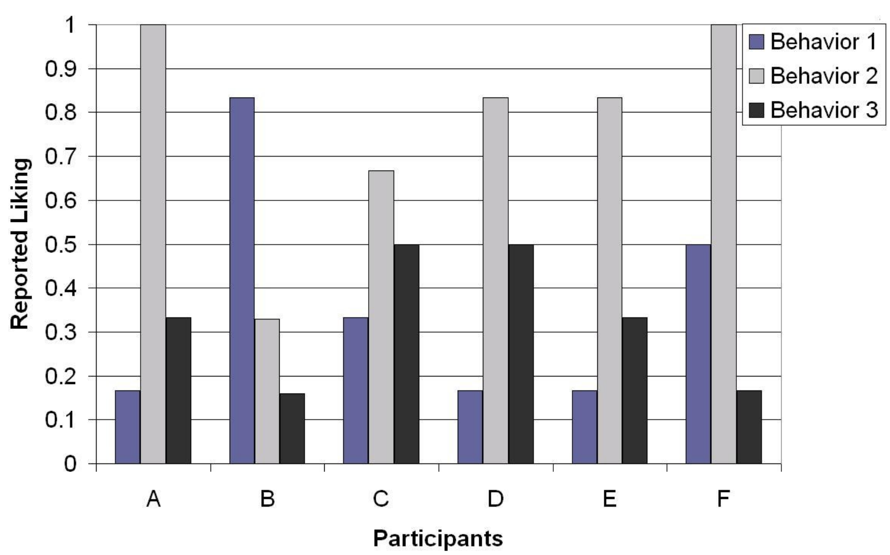


Figure 3.4 Mean liking level for different behaviors in RBB1

of reported liking for different behaviors are statistically significant ($p < 0.05$), whereas no significant effect due to different participants was observed. Furthermore, it was also observed that different children with ASD may have different preferences for the robot's behaviors. These results demonstrated that it is important to have a robot learn the individual's preference and adapt to it automatically, which may allow a more tailored and affect-sensitive interaction between children with ASD and the robot. For example, when a robot learns that a certain behavior is liked more by a particular child, it can choose that behavior as his/her "social feedback" or "reinforcer" in robot-assisted autism intervention. Playful interaction will be more likely to emerge by addressing a child's preference.

Second, the predictive accuracy of how closely the real-time physiology-based quantitative measures of liking, as obtained from affective models developed in Chapter II, matched with that of the subjective rating of liking made by the therapist during the RBB1 and RBB2 tasks (36 epochs total per child) is presented in Table 3.2. The average predictive accuracy across all the participants was approximately 81.0%. The highest was 86.1% for Child D, and the lowest was 77.8% for Child B and Child E. Note that the affective model was evaluated based on physiological data obtained on-line from a real-time application for children with ASD. However, this prediction accuracy is comparable to the results achieved through off-line analysis for typical individuals (Nasoz et al., 2004; Rani et al., 2006).

Table 3.2 Real-time classification accuracy of liking

	Child ID					
	A	B	C	D	E	F
Prediction Accuracy (%)	83.3	77.8	80.6	86.1	77.8	80.6

Third, we present results about robot behavior adaptation and investigate its impact on the interaction between the children with ASD and the robot. Table 3.3 shows the percentages of different behaviors that were chosen in the RBB2 session for each participant. The robot learned the individual's preference and selected the most-preferred behavior with high probability for all the participants. Averaged across all participants, the most-preferred, moderately-preferred, and least-preferred behaviors were chosen 72.5%, 16.7%, and 10.8% of the time, respectively. The preference of a behavior was defined by the reported liking level in RBB1 as shown in Figure 3.4. To understand these results more clearly we describe an individual case. Figure 3.5 shows the affect-sensitive behavior adaptation in RBB2 for Child A, who prefers behavior 2 most (refer to Figure 3.4). The real-time predicted liking level (i.e., high/low) is denoted by "0" for low and "1" for high. The robot starts in a non-preferred behavior (behavior 1) and then explores other behaviors before settling on the most-preferred behavior (behavior 2) where the liking level is high (as confirmed by the affective model prediction as well as the therapist's subjective report). After a considerable time interacting with behavior 2 (e.g., epoch 7), the participant appears to not enjoy this behavior as much as before. The affective model detects this change and returns negative rewards. The QV-learning

Table 3.3 Proportion of Different Behaviors Performed in RBB2

Child ID	Most-Liked Behavior		Moderate-Liked Behavior		Least-Liked Behavior	
	ID	Proportion	ID	Proportion	ID	Proportion
A	2	82.4%	3	11.8%	1	5.8%
B	1	70.6%	2	17.7%	3	11.7%
C	2	58.8%	3	23.5%	1	17.7%
D	2	76.5%	3	11.8%	1	11.7%
E	2	76.5%	3	17.6%	1	5.9%
F	2	70.6%	1	17.7%	3	11.7%

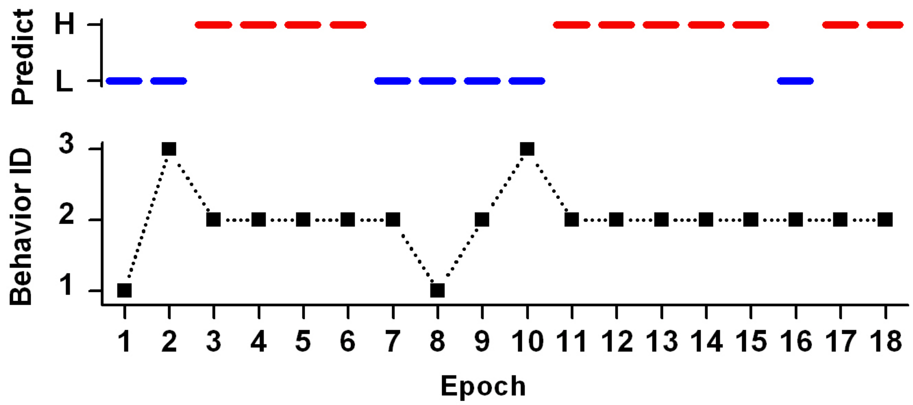


Figure 3.5 Behavior selected by affect-sensitive robot in RBB2 for Child A

algorithm updates its state/action space and directs the robot to switch behaviors. However, after exploring other behaviors, the robot eventually finds that behavior 2 is the most-preferred by Child A (e.g., epoch 11) and continues the interaction using this behavior. At epoch 16, even though the predicted liking level is low, due to high frequent positive rewards received for behavior 2, the robot checks the updated Q-function and remains at this behavior. There could be several reasons why less-preferred behaviors were chosen in RBB2. The learned behavior selection policy might not have been optimal after the exploration in RBB1, and the QV-learning algorithm continued the learning process in RBB2. Another reason could be that the affective model is not 100% accurate and may return false reward/punishment, which may have given the robot imperfect instruction for behavior switches. Habituation to the most-preferred behavior during RBB2 could also be a factor that might have contributed to temporary changes in preference which led the robot to choose other behaviors.

In Table 3.3, the robot chose a less-preferred behavior more often for Child C than for other participants. As can be seen in Figure 3.4, Child C does not show differences of liking among the three behaviors as significantly as the other children. This instance of less-distinct preference could result in more inconsistent rewards/punishments and the robot switching behaviors more frequently. However, despite the above possible reasons for choosing less-preferred behaviors, Table 3.3 and Figure 3.5 show that the robot is capable of identifying and selecting the preferred behavior autonomously in most of the epochs for all participants and thus positively influencing the subjective liking level of the children with ASD (as shown in Figure 3.6).

In Figure 3.6, we present results to demonstrate that active monitoring of participants' liking and autonomously selecting the preferred behavior allowed children with ASD to maintain high liking levels. The average labeled liking levels of the participants as reported by the therapist during the two sessions were compared, and the small bars represent the standard error of the mean. The agreement between the therapist and parent on the subjective liking level is substantial for both RBB sessions and has a larger *K* statistic value (0.71) than that of the other two possible reporter pairs (0.43 for

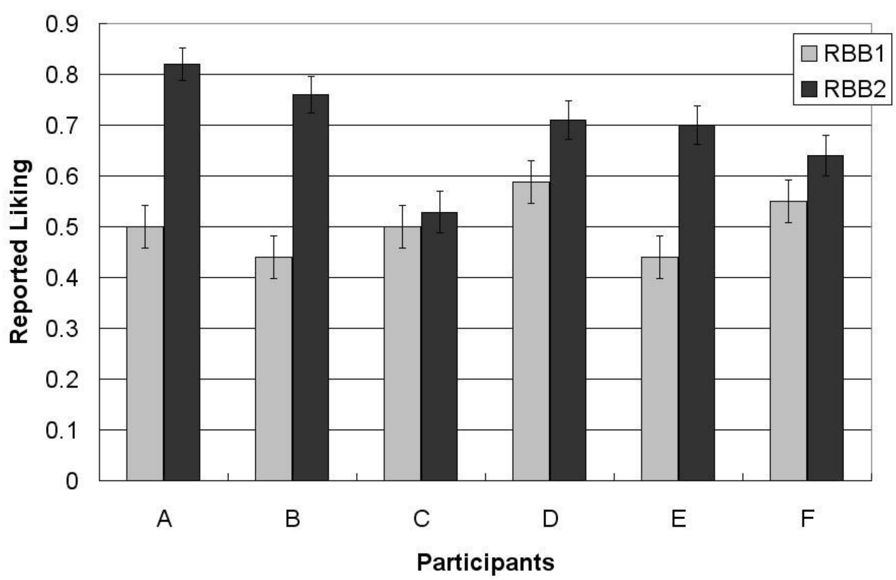


Figure 3.6 Subjective liking as reported by therapist

the parent and children and 0.39 for the therapist and children). The lighter bars in Figure 3.6 indicate the liking level during the RBB1 session (i.e., when the robot selected behaviors randomly), and the darker bars show the liking level during the RBB2 session (i.e., when the robot learned the individual preference and chose the appropriate behavior accordingly). It can be seen that for all participants liking level was maintained, and for five of the six children liking level increased. There was no significant increase for Child C during the liking-based session as compared to the non-affect based session. As mentioned earlier, the impact of the different robot behaviors on the liking level of Child C is not as significant as that of the others, which may impede the robot in finding the preferred behavior and hence impede the robot in effectively influencing the subjective liking level positively. Note that RBB1 presents a typically balanced interaction with equal numbers of most-preferred, moderately-preferred, and least-preferred epochs, and the comparisons in Figure 3.6 are not between liking-based sessions and sessions of least-preferred epochs. To determine the effects of the session type and participant on the reported liking, a two-way ANOVA test was performed. The null hypothesis that there is no change in liking level between liking-based sessions and non-affect based sessions could be rejected at the 99.5% confidence level. Additionally, no significant impact due to different participants was observed. This was an important result as the robot continued learning and utilizing the information regarding the probable liking level of children to adjust its behaviors. This ability enables the robot to adapt its behavior selection policy in real time and hence keep the participant in a higher liking level.

Discussion

In this chapter, we have proposed a novel framework for affect-sensitive HRI where the robot can implicitly detect the affective states of the children with ASD as discerned by the therapist and respond to it accordingly. To account for the phenomenon of person-stereotypy and the diverse affective characteristics of the children with ASD, we employed an individual-specific approach for affective modeling. An intensive study was preformed based on a large sample size of observations (approximately 85 epochs over 6 hours) for each of the six children with ASD. The time spent collecting the training data for affective modeling can be justified by the current ASD intervention practices (Tarkan, 2002). However, note that the methodology for inducing, gathering, and modeling the experimental data is not dependent on the participants. The consistently reliable prediction accuracy for each participant demonstrated that it was feasible to model the affective states of children with ASD via psychophysiological analysis.

To investigate the affect-sensitive closed-loop interaction between the children with ASD and the robot, we designed a proof-of-concept task, robot-based basketball, and developed an experimental system for its real-time implementation and verification. The real-time prediction of liking level of the children with ASD was accomplished with an average accuracy of 81.0%. The robot learned individual preferences of the children with ASD over time based on the interaction experience and the predicted liking level and hence autonomously selected the most-preferred behavior, on average, 72.6% of the time. We have observed that such affect-sensitive robot behavior adaptation has led to an increase in reported liking level of the children with ASD. This is the first time, to our knowledge, that the affective states of children with ASD have been detected via a physiology-based affect recognition technique in real time. This is also the first time that the impact of affect-sensitive closed-loop interaction between a robot and children with ASD has been demonstrated experimentally.

CHAPTER IV

VIRTUAL ENVIRONMENT SYSTEM FOR SOCIAL INTERACTION

Introduction

While research indicates the usefulness of technology-assisted intervention (e.g., computers, robots, and VR) for children with ASD, researchers have not yet fully utilized the potential of these technologies in the realm of social communication. Research in this chapter is designed to promote the development of an innovative technological approach for examining social communication difficulty for individuals with ASD with the aim of applying this technology as a future intervention paradigm. Furthermore, there are no design guidelines as to how to develop socially acceptable virtual peers (e.g., embodied humanoid robots or computerized humanoid characters) to be used for social skill intervention for children with ASD (Dautenhahn and Werry, 2004; Parsons, et al., 2005; Pioggia, et al., 2005; Strickland et al., 1996). In particular, it is important to know how these virtual peers should display intentions (e.g., facial expressions) and how they should interact (e.g., amount of eye contact, proximity to the child, etc.) to relay the intended social skill teaching to this population. Additionally, it is important to develop an evaluation method that is not solely dependent on self-report because of the known difficulty of self expression exhibited by children with ASD (Hill et al., 2004). Research that systematically develops social peers and studies social interaction with children with ASD in a step-by-step manner is thus necessary. Similar to how sophisticated machines are designed in virtual environments prior to fabrication, this chapter covers the design of virtual peers in VR environments that would enable a systematic evaluation and manipulation of different components of social interaction through virtual social peers.

If virtual peers are to be used to impart social skills, a primary deficit for the ASD population, considerable attention needs to be paid to aspects of social acceptability of such intervention aids. As a first step, this work investigates social design of humanoid characters (i.e., avatars) in a VR environment for children with ASD. In this chapter we describe the design of a virtual environment system for social interaction. Initial results are presented, and a more extensive evaluation of the system is investigated in Chapter V. The design is evaluated through an experiment plan that combines subjective ratings from a clinical observer/therapist with physiological responses indicative of affective states from the participant, both collected when a participant engages in social tasks with the avatars in a VR environment. Two social parameters of importance, namely eye gaze and social distance, are systematically varied to analyze the response of the participant.

This chapter describes the design and development of a virtual environment system for social interaction that is capable of systematic manipulation of various design parameters that are important for the development of virtual social peers. We describe the socialization and expressivity of the VR characters. The VR system is formulated to present realistic social communication tasks to children with ASD, and the children's affective response during the tasks are monitored through physiological signals and observations from a clinician. This system is capable of systematically manipulating specific aspects of social communication to more fully understand how to design virtual peers for use in technology-assisted intervention for children with ASD.

Task Design

For ASD intervention, VR is often effectively experienced on a desktop system using standard computer input devices (Parsons and Mitchell, 2002). The focus of this work is on desktop VR applications, chosen over more immersive technologies because it is more accessible, affordable, and less susceptible to "cybersickness" problems (nausea, headaches, dizziness) potentially associated with head-mounted devices (Cobb et al., 1999). Therefore, users view the VR environment on a computer monitor from the first-person perspective. Realistic VR scenarios are created for interaction with virtual social peers (i.e., expressive humanoid avatars). Vizard (worldviz.com), a commercially available VR design package, is employed to develop the environments. Within the controllable VR environment, components of the interaction are systematically manipulated to allow users to explore different social compositions. The avatars can make direct and averted eye contact. They can converse by matching their mouth movements to recorded sound files. Open-ended interaction between a participant actively speaking and an avatar interpreting and responding to the speech is not feasible in the current system design as it is a computationally intractable problem. Instead the participant responds to the avatars using a keypad to select from transparent text boxes superimposed in the corner of the VR scene.

Social Parameters

Eye gaze and social distance, the social parameters of interest, are organized in a 4x2 experimental design, which makes possible eight distinct situations. These parameters were chosen because they play significant roles in social communication and interaction (Bancroft, 1995), and manipulation of these factors may elicit variations in affective reactions (Argyle and Dean, 1965) and physiological responses (Groden et al., 2005). Each situation is represented three times, which creates 24 trials in the experiment, following a Latin Square design to balance for sequencing and order effects (Keppel, 1991). Each trial of an experiment session includes one avatar for one-on-one interaction with the participant. Participants are asked to engage in an interactive social task in the virtual environment. The specific task is modified such that the social communication parameters can be repeatedly explored while sustaining engagement. In each trial, participants are instructed to watch and listen as the virtual peer tells a 2-min story. The stories are written in first-person. Thus, the task can be likened to having different people introduce themselves to the user, which is comparable to research on social anxiety and social conventions (Argyle and Dean, 1965; Schneiderman and Ewens, 1971; Sommer, 1962). Other social parameters, such as facial expression and vocal tone are kept as neutral as possible. However, we also attempt to make the task interesting enough so that participants do not become excessively detached based on habituation or dull content.

The eye gaze parameter dictates the percentage of time an avatar looks at the participant (i.e., staring straight out of the computer monitor). Four types of eye gaze are examined. These are defined as "straight," "averted," "normal," and "flip of normal." Straight gaze means looking straight ahead for the duration of the story (i.e., for the entire trial). Averted gaze means the avatar never attempts to make direct eye contact with the participant, but instead alternates between looking to the left, right, and up. Based on social psychology literature from experimental observations of typical humans (Argyle and Cook, 1976) and algorithms adopted by the artificial intelligence community to create realistic virtual characters (Colburn et al., 2000; Garau et al., 2001), normal eye gaze is defined as a mix of straight and averted gaze. A person displays varying mixes of direct and averted eye contact depending on if the person is speaking or listening during face-to-face conversations. Since the virtual peer in the VR environment is speaking, we use the gaze definitions for a person speaking, which is approximately 30% straight gaze and 70% averted gaze (Argyle and Cook, 1976; Colburn et al., 2000). Research represents averted gaze as looking more than 10° away from center in evenly-distributed, randomly-selected directions (Garau et al., 2001; Jenkins et al., 2006). Therefore, our averted gaze is an even distribution (33.3% each) of gazing left, right, and up more than 10° from center. Flip gaze is defined as the flip of normal, which means looking straight approximately 70% of the time and averted 30% of the time, which is indicative of a person's gaze while listening.

The social distance parameter is characterized by the distance between the avatar and the user. Two types of social distance, termed "invasive" and "decorum," are examined. In the VR environment, distance is simulated but can be appropriately represented to the view of the participant. For invasive distance, the virtual peer stands approximately 1.5 ft. from the main view of the scene. This social distance has been characterized as intimate space not used for meeting people for the first time or for having casual conversations with friends (Hall, 1955). A distance of 1.5 ft. apart has been investigated by several research groups in experiments with similar experimental setups to ours in which two people are specifically positioned while one introduces himself/herself to the other and discusses a personal topic for approximately 2 min (Argyle and Dean, 1965; Schneiderman and Ewens, 1971; Sommer, 1962), and this invasive distance is characterized by eliciting uncomfortable feelings and attempts to increase the distance to achieve a social equilibrium consistent with comfortable social interaction (Argyle and Dean, 1965). Decorum distance means the avatar stands approximately 4.5 ft. from the main view of the scene. This social distance is consistent with conversations when meeting a new person or a casual friend (Hall, 1966), and research indicates this distance results in a more comfortable conversation experience than the invasive distance (Argyle and Dean, 1965). Using Vizard software we project virtual social peers who display different eye gaze patterns at different distances; two examples are shown in Figure 4.1.

Humanoid Avatars

The virtual social peers have a fixed male or female body, but Dr. Jeremy Bailenson, director of the Virtual Human Interaction Lab at Stanford University, provided a set of distinct humanoid avatar heads for use in this work. The set of 26 heads was created from front and side 2D photographs of college-age students. Using 3DMeNow software (biovirtual.com), the photos were formed into 3D heads that can be used in Vizard. Even though Bailenson's avatar heads are slightly older than the participants recruited for this study, they are used because of the following advantages: (i) open accessibility, (ii) age range close to our participant pool's peers, (iii) and the authentic facial features (e.g., variations in skin complexion, brow line, nose dimensions, etc.) allow the interaction to be interpreted as realistically as possible.

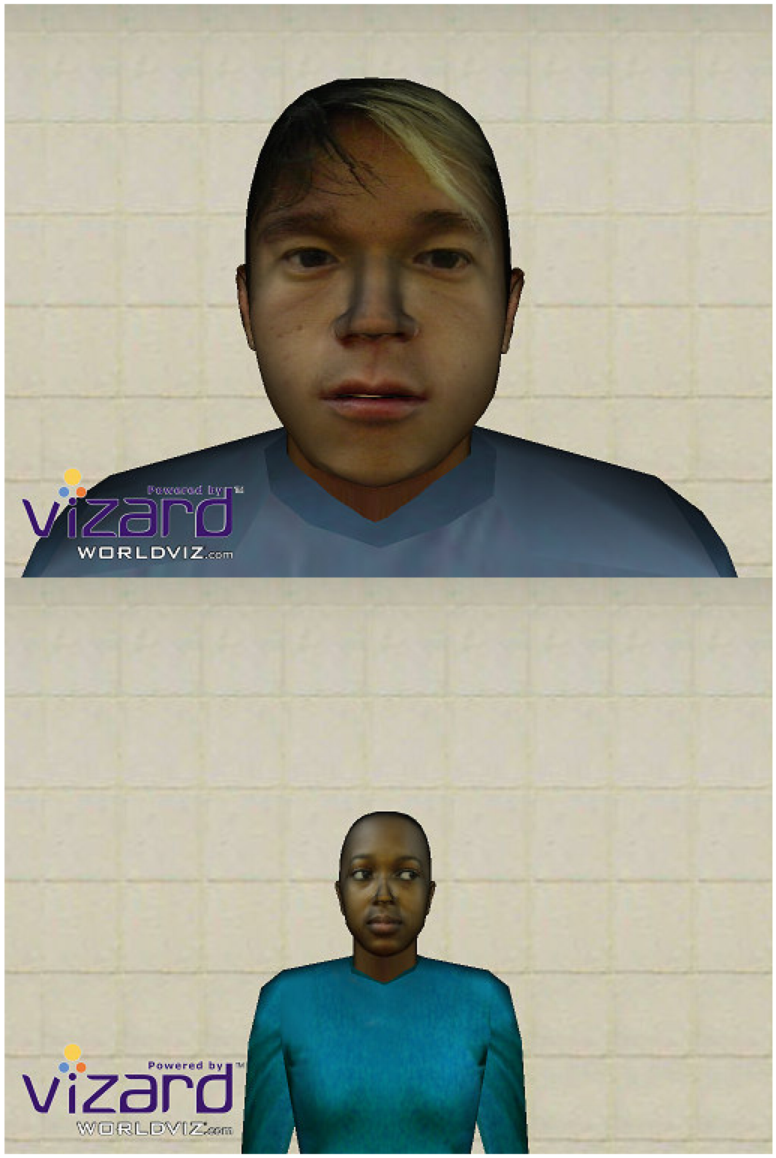


Figure 4.1 At *top* an avatar displays straight gaze at an invasive distance, while on *bottom* an avatar stands at a decorum distance and looks to her left in an averted gaze.

The stories the virtual peers share are adapted from DIBELS (Dynamic Indicators of Basic Early Literacy Skills; dibels.uoregon.edu/measures/) reading assessments. These short readings are offered nationwide for internal educational use but may not be resold or distributed on a for-profit basis. The University of Oregon Center on Teaching and Learning encourages the use of these materials as long as they are not used to coach elementary school children before a reading assessment. The assessments are written on topics such as geographical locations, weather phenomena, and intriguing occupations. The readings from fifth grade were chosen based on length and because this grade level corresponds to vocabulary tests used in the Chapter II experiments in the design of an easy-to-medium level of the Anagrams game used with a similar recruitment pool. In each trial of the experiment, an avatar narrates one of these first-person stories to the user. The voices were gathered from teenagers and college-age students from the regional area. Their ages (range = 13-22 years, mean = 18.5 yrs, SD = 2.3 yrs) are similar to the age of people used for the avatar heads and our participant pool.

Social Interaction

The interaction involves the virtual social peer telling a story while a participant listens. At the end of the story, the avatar asks the participant a question about the story. The questions are designed to facilitate interaction and to serve as a possible objective measure of engagement. The participant is not aware of the exact question before the story begins so that he/she engages in the task and is not focused on listening to one specific part of the discourse. The questions are intended to be easy to answer correctly if the participant listened to the story. Near the beginning of the first experiment session, the participant takes part in two demonstrations of the process of the VR task; therefore, any difficulty over correctly answering the questions that could be related to not understanding the process of the task is dealt with prior to starting the experiment and collecting data. Each question is accompanied by three possible answer choices (see Figure 4.2). The correct choice is spoken at least five times during the story, which is sufficient for the information to be relayed (Jonides et al., 2008), and the incorrect

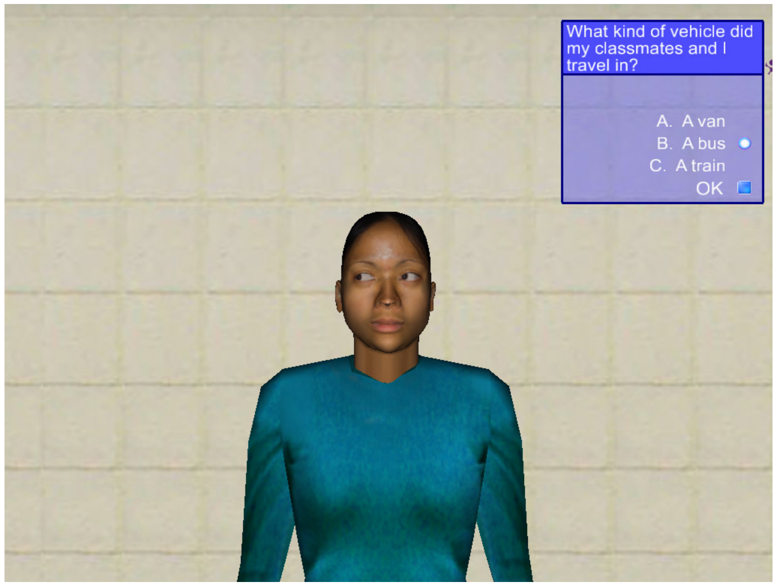


Figure 4.2 Example question at the end of a story

choices are never spoken in the story. For example, the story preceding the question shown in Figure 4.2 is about a bus breaking down on the way to a school picnic. At the end of the story the avatar asks, "What kind of vehicle did my classmates and I travel in?" The story includes the line "...a car appeared at the top of a hill...;" therefore, the offered choices to the question (A. A van, B. A bus, C. A train) do not include "car" as an option. We expect that a participant who engages in the task would achieve near to or complete 100% accuracy on the questions; and consequently, a severely low percentage of correct answers would indicate a lack of engagement with the task.

System Refinement

Efforts were made to minimize reactions due solely to viewing a virtual peer by choosing the 10 most-neutral avatar heads based on a survey of 20 students. Therefore, reactions during the experiment could be reasonably expected to be related to change in eye gaze and/or social distance and not due to viewing the avatar alone. Participants for the avatar head survey were recruited from undergraduate engineering and psychology courses at Vanderbilt University. Twenty (10 male) students completed the survey. These students were instructed to visit a webpage to complete a survey on their impressions of the 26 avatars (see Figure 4.3).

Participants were asked to rate each avatar on four questions. Two questions were designed to measure elicited reactions from viewing the avatars (Q1 and Q2), and two were designed to determine participants' perceptions of each avatar's display of emotion (Q3 and Q4). Following descriptive terms from Lang et al. (1999), participants were asked to rate how they felt when looking at each avatar on a 5-point scale of valence and arousal. To gauge affective reactions to images, valence and arousal are important measures to consider (Bradley, 1994). Participants were also asked two questions about what emotion the avatar was conveying (Ekman, 1992; Frijda, 1986).

Q1, Valence:
When looking at this avatar, I feel...

unhappy,		happy,
annoyed,	neutral	pleased,
despaired		hopeful

○ ○ ○ ○ ○

Q2, Arousal:
When looking at this avatar, I feel...

calm,		excited,
sleepy,	neutral	wide-awake,
unaroused		aroused

○ ○ ○ ○ ○

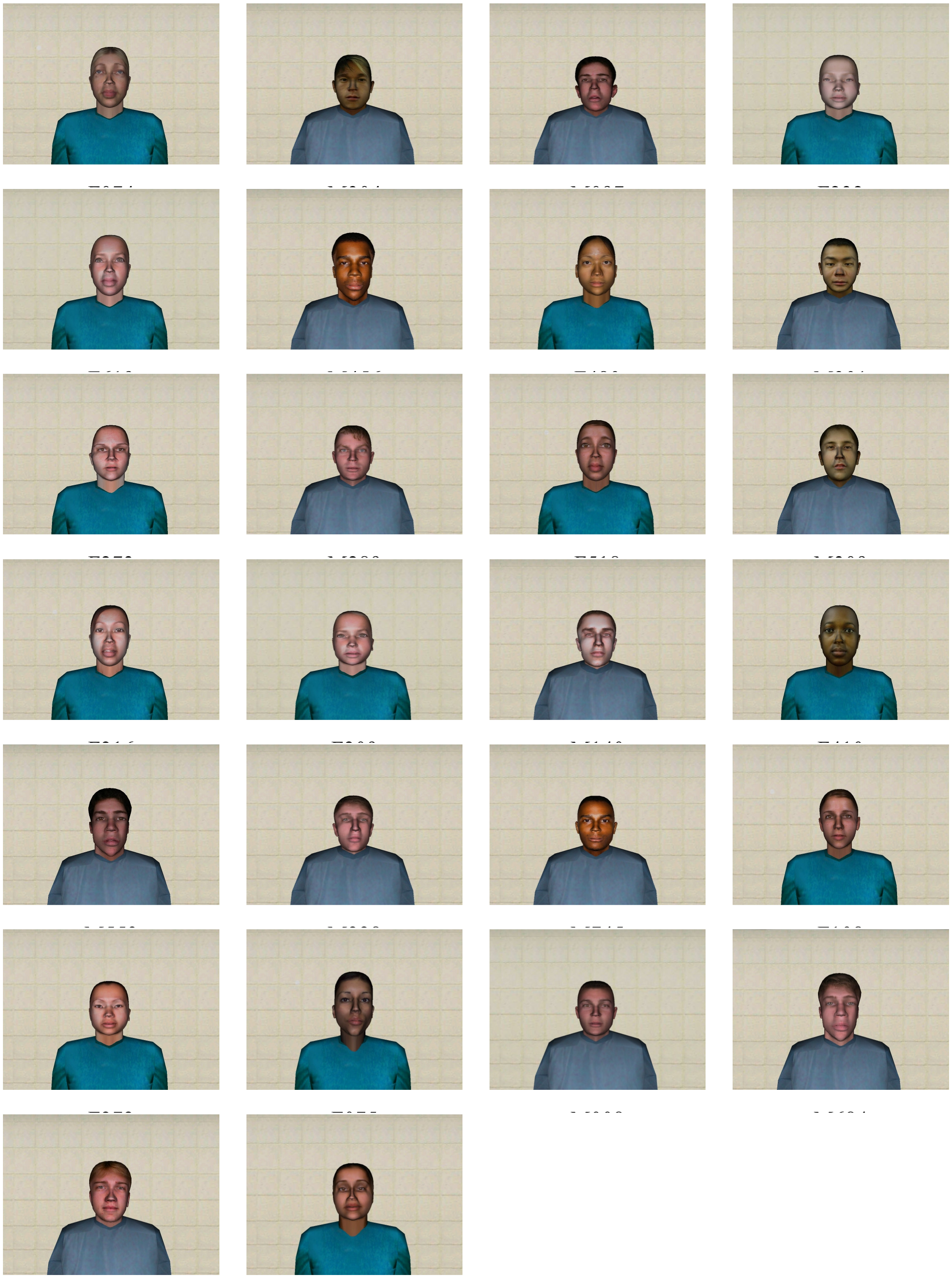


Figure 4.3 Screenshots and IDs of all 26 avatar heads atop the fixed male or female bodies are shown in the random order viewed by participant 1 of the survey.

Q3, Emotion:
Do you feel that this avatar is expressing a specific emotion or no emotion/neutral?
○ a specific emotion
○ no emotion/neutral

Q4, Degree of Chosen Emotion:
If you had to choose, which emotion would you say this avatar is expressing? Please indicate to what degree you would say this avatar is expressing your chosen emotion. (Make only one mark.)

	Low		Medium		High
anger	○	○	○	○	○
disgust	○	○	○	○	○
fear	○	○	○	○	○
happiness	○	○	○	○	○
sadness	○	○	○	○	○
surprise	○	○	○	○	○

Each participant was presented with all 26 avatar heads (see Figure 4.3) in a randomized order. After participants answered evaluation questions on one avatar head, they were presented with the next avatar head. This process continued until they evaluated all 26 avatar heads.

Survey Analysis

The survey ratings were collected from 20 students to determine the most-neutral avatar heads. The average value of the valence and arousal ratings, on a [-2,2] scale, were calculated from Q1 and Q2, respectively. The scale had a neutral point at 0, which was our desired point of reference. Therefore, it was desirable to identify the avatar heads with ratings closest to 0,0 on the valence vs. arousal affective space. The mean value of the valence and arousal ratings were used to determine the Euclidean distance from the 0,0 origin in the valence-arousal affective space. The Euclidean distance measurement was divided by $2\sqrt{2}$ to normalize the values to a [0,1] scale.

$$E_{norm,i} = \frac{\sqrt{(V_{mean,i} - 0)^2 + (A_{mean,i} - 0)^2}}{2\sqrt{2}} \tag{4.1}$$

Eqn. 4.1 shows the normalized Euclidean distance calculation, where i represents each individual avatar head (26 total), $V_{mean,i}$ is the average valence rating for each avatar head from Q1, $A_{mean,i}$ is the average valence rating for each avatar head from Q2, and $E_{norm,i}$ is the normalized Euclidean distance from 0,0 in the valence-arousal affective space. Ratings from Q3 and Q4 were also considered in the overall rating of the avatar heads. For Q3, the portion of respondents answering "a specific emotion" was calculated for each avatar head. When analyzing results for Q4, we did not discriminate on which emotion was chosen (e.g., happiness, anger, etc.), because we were most concerned with the emotion being as minimally expressed as possible. Ratings from Q4 on a [0,4] scale were divided by 4 to achieve a [0,1] scale.

For all three measurements, a lower score reflects our desired outcome. Therefore, the most-neutral avatar heads were considered as having (i) the shortest distance away from 0,0 in the valence-arousal space, (ii) the least portion of respondents answering "a specific emotion" to Q3, and (iii) the lowest degree of emotion expression to Q4. After normalization, the three measurements were all on a [0,1] scale. Thus, we combined these measurements into a weighted equation for an overall rating of the avatar heads. A strong emphasis was placed on the valence and arousal ratings, because these have shown to be a reliable assessment of affect and proven to sufficiently cover the affective space (Bradley, 1994). In this survey we also took into account, but to a lesser extent, whether the avatar heads were perceived as expressing little to no emotion/neutral emotion. The combined equation used to rate the avatar heads is as follows,

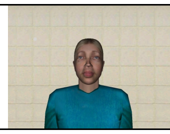
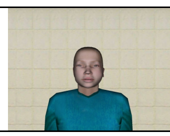
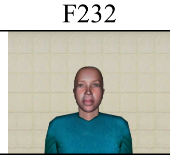
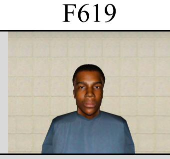
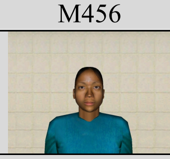
$$R_i = 0.8\left(E_{norm,i}\right) + 0.1\left(Q_{3,i}\right) + 0.1\left(Q_{4,i}\right) \tag{4.2}$$

where R_i is the overall rating of avatar head i, $E_{norm,i}$ is the normalized Euclidean distance from 0,0 in the valence-arousal affective space, $Q_{3,i}$ is the average rating for responses to Q3 on the survey, and $Q_{4,i}$ is the average rating for responses to Q4 on the survey (see Table 4.1).

Survey Results

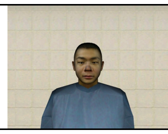
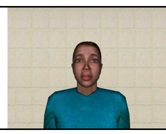
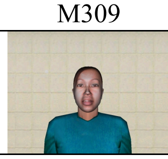
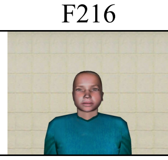
This is the first time these avatar heads and expressions have been tested to determine if Vizard's facial morph expressions are interpreted by viewers as intended. The "neutral" morph was used in the survey and for the virtual social peers in the VR system. Vizard supplies other emotion morphs (e.g., "happy," "surprise," etc.) for use with the avatar heads. Although Vizard uses common methods for defining the morphs (i.e., furrowed brow for "angry" morph, elevated brow and slightly open mouth for "surprise" morph, etc.)(Ekman, 1993; Frijda, 1986), Vizard has not tested user perception of the morphs to establish if the morphs convey the designed emotion to the viewer.

Table 4.1 Ratings from the survey are listed for measurements of valence, arousal, emotion, and degree of chosen emotion. The most-neutral avatar heads (shaded in light gray) have the lowest combined weighted scores for columns 4-6 of the table (see Eqn. 4.2).

Screenshot	Mean Survey Measurement across Respondents (N=20)					
	Q_1[a]	Q_2[b]	Eqn. 4.1[c]	Q_3[d]	Q_4[e]	Eqn. 4.2[f]
 F074	-0.80	-0.65	0.36	0.45	0.34	0.37
 M304	-0.25	-0.05	0.09	0.20	0.10	0.10
 M097	-0.50	-0.15	0.18	0.60	0.28	0.24
 F232	-0.40	-0.60	0.25	0.25	0.10	0.24
 F619	-0.75	0.20	0.27	0.60	0.25	0.30
 M456	-0.20	-0.20	0.10	0.15	0.10	0.11
 F480	-0.15	-0.10	0.06	0.30	0.18	0.10

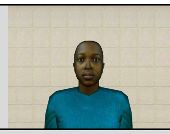
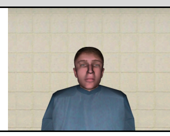
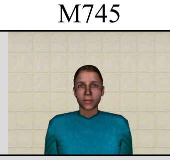
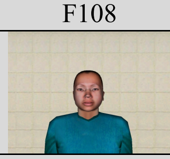
[a]Q_1: Mean valance rating on a [-2,2] scale
[b]Q_2: Mean arousal rating on a [-2,2] scale
[c]Eqn. 4.1: Euclidean distance from neutral origin 0,0 on valence vs. arousal affective space, divided by $2\sqrt{2}$ to achieve a [0,1] scale
[d]Q_3: Portion of respondents that answered "a specific emotion" to Q3
[e]Q_4: Mean degree of expression of chosen emotion, original [0,4] scale divided by 4 to achieve a [0,1] scale
[f]Eqn. 4.2: Weighted sum of Eqn. 4.1, Q_3, and Q_4

Table 4.1 (continued) Ratings from the survey are listed for measurements of valence, arousal, emotion, and degree of chosen emotion. The most-neutral avatar heads (shaded in light gray) have the lowest combined weighted scores for columns 4-6 of the table (see Eqn. 4.2).

Screenshot	**Mean Survey Measurement across Respondents (N=20)**					
	Q_1[a]	Q_2[b]	Eqn. 4.1[c]	Q_3[d]	Q_4[e]	Eqn. 4.2[f]
 M301	-0.50	0.00	0.18	0.20	0.16	0.18
 F273	-0.35	0.05	0.13	0.60	0.25	0.19
 M289	-0.75	0.40	0.30	0.80	0.38	0.36
 F518	-0.70	0.05	0.25	0.80	0.38	0.32
 M309	-0.35	-0.50	0.22	0.35	0.11	0.22
 F216	-0.45	0.20	0.17	0.65	0.35	0.24
 F209	-0.55	-0.05	0.20	0.55	0.23	0.23

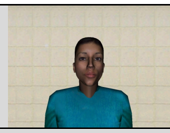
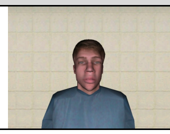
[a]Q_1: Mean valance rating on a [-2,2] scale
[b]Q_2: Mean arousal rating on a [-2,2] scale
[c]Eqn. 4.1: Euclidean distance from neutral origin 0,0 on valence vs. arousal affective space, divided by $2\sqrt{2}$ to achieve a [0,1] scale
[d]Q_3: Portion of respondents that answered "a specific emotion" to Q3
[e]Q_4: Mean degree of expression of chosen emotion, original [0,4] scale divided by 4 to achieve a [0,1] scale
[f]Eqn. 4.2: Weighted sum of Eqn. 4.1, Q_3, and Q_4

Table 4.1 (continued) Ratings from the survey are listed for measurements of valence, arousal, emotion, and degree of chosen emotion. The most-neutral avatar heads (shaded in light gray) have the lowest combined weighted scores for columns 4-6 of the table (see Eqn. 4.2).

Screenshot	**Mean Survey Measurement across Respondents (N=20)**					
	Q_1[a]	Q_2[b]	Eqn. 4.1[c]	Q_3[d]	Q_4[e]	Eqn. 4.2[f]
 M140	-0.40	-0.30	0.18	0.35	0.15	0.19
 F410	0.10	0.10	0.05	0.40	0.24	0.10
 M553	-0.30	-0.10	0.11	0.50	0.16	0.16
 M228	-0.85	-0.40	0.33	0.45	0.29	0.34
 M745	0.20	0.25	0.11	0.65	0.38	0.19
 F108	0.20	0.10	0.08	0.25	0.10	0.10
 F272	-0.30	-0.20	0.13	0.25	0.18	0.14

[a]Q_1: Mean valance rating on a [-2,2] scale
[b]Q_2: Mean arousal rating on a [-2,2] scale
[c]Eqn. 4.1: Euclidean distance from neutral origin 0,0 on valence vs. arousal affective space, divided by $2\sqrt{2}$ to achieve a [0,1] scale
[d]Q_3: Portion of respondents that answered "a specific emotion" to Q3
[e]Q_4: Mean degree of expression of chosen emotion, original [0,4] scale divided by 4 to achieve a [0,1] scale
[f]Eqn. 4.2: Weighted sum of Eqn. 4.1, Q_3, and Q_4

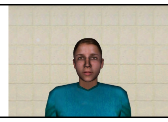
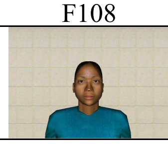
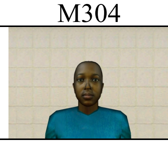
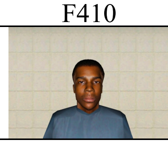
Table 4.1 (continued) Ratings from the survey are listed for measurements of valence, arousal, emotion, and degree of chosen emotion. The most-neutral avatar heads (shaded in light gray) have the lowest combined weighted scores for columns 4-6 of the table (see Eqn. 4.2).

Screenshot	**Mean Survey Measurement across Respondents (N=20)**					
	Q_1[a]	Q_2[b]	Eqn. 4.1[c]	Q_3[d]	Q_4[e]	Eqn. 4.2[f]
 F075	-0.15	0.00	0.05	0.55	0.16	0.11
 M008	0.10	0.15	0.06	0.60	0.28	0.14
 M684	-0.75	-0.65	0.35	0.45	0.33	0.36
 M327	-0.15	-0.20	0.09	0.60	0.28	0.16
 F183	-0.50	-0.20	0.19	0.70	0.23	0.24

[a]Q_1: Mean valance rating on a [-2,2] scale
[b]Q_2: Mean arousal rating on a [-2,2] scale
[c]Eqn. 4.1: Euclidean distance from neutral origin 0,0 on valence vs. arousal affective space, divided by $2\sqrt{2}$ to achieve a [0,1] scale
[d]Q_3: Portion of respondents that answered "a specific emotion" to Q3
[e]Q_4: Mean degree of expression of chosen emotion, original [0,4] scale divided by 4 to achieve a [0,1] scale
[f]Eqn. 4.2: Weighted sum of Eqn. 4.1, Q_3, and Q_4

The undergraduate students who completed the avatar head survey ranged in age from 18-21 yrs with a mean = 19.2 and SD = 0.9. The results, shown in Table 4.2, are sorted in ascending order from lowest to highest neutral rating. The 10 avatar heads with the lowest scores are used for the virtual social peers in the VR experiments. The four lowest female avatar heads and four lowest male avatar heads were used in the eight

Table 4.2 Shown are the 10 avatar heads with the most-neutral ratings from the survey. These avatar heads are used for the virtual social peers in the VR experiments. Also listed are their assigned experiment conditions (EC). The EC's use the following abbreviations: for social distance, Invasive (I) and Decorum (D); for eye gaze, Straight (S), Averted (A), Normal (N), and Flip (F) of normal.

EC	Screenshot	Mean Survey Measurement across Respondents (N=20)					
		Q_1[a]	Q_2[b]	Eqn. 4.1[c]	Q_3[d]	Q_4[e]	Eqn. 4.2[f]
IN	 F108	0.20	0.10	0.08	0.25	0.10	0.10
DA	 F480	-0.15	-0.10	0.06	0.30	0.18	0.10
DS	 M304	-0.25	-0.05	0.09	0.20	0.10	0.10
DF	 F410	0.10	0.10	0.05	0.40	0.24	0.10
IF	 M456	-0.20	-0.20	0.10	0.15	0.10	0.11
IS	 F075	-0.15	0.00	0.05	0.55	0.16	0.11
IA	 M008	0.10	0.15	0.06	0.60	0.28	0.14
IA *Demo*	 F272	-0.30	-0.20	0.13	0.25	0.18	0.14
DN	 M553	-0.30	-0.10	0.11	0.50	0.16	0.16
DN *Demo*	 M327	-0.15	-0.20	0.09	0.60	0.28	0.16

experiment conditions. The female and male avatar heads with the highest neutral rating of the bottom 10 are used for the demonstration of the VR interaction during session one of the experiment. Each of the eight experiment conditions are shown three times, creating 24 trials in the experiment, which are divided over two sessions on two different days for each participant.

Experiment Protocol

Participant Recruitment

Participants are recruited through existing clinical and research programs of the Vanderbilt Kennedy Center's Treatment and Research Institute for Autism Spectrum Disorders and Vanderbilt University Medical Center. Our protocol calls for enlisting children with ASD age 13-18 years old and an age- and verbal-ability-matched control group of typically-developing (TD) children. ASD participants must have documentation of their diagnosis on the autism spectrum, either Autism Spectrum Disorder, Autistic Disorder, or Asperger's Syndrome, according to their medical records. For all participants, the Social Responsiveness Scale (SRS; Constantino, 2002) profile sheet and Social Communication Questionnaire (SCQ; Rutter et al., 2003a) are completed by a participant's parent/caregiver before the first session to provide an index of current functioning and ASD symptom profiles. Selection is also based on a receptive vocabulary standard score of 80 or above on the PPVT-III. The PPVT-III (Dunn and Dunn, 1997) was used as an inclusion criterion in the research covered in Chapters II and III. The chosen age range and intelligence testing cutoff represents a method of partial control for the reading skill requirements of the task and ensures that participants are able to perform the interaction tasks. All written components of the current design are accompanied by audio readings, thus alleviating some of the language requirements, which could open the prospect of including younger participants or those with less language and/or reading skills in future studies.

Procedure

The commitment required of participants is a total of two sessions (i.e., approximately 2.5 hrs). The first session runs approximately 1.5 hrs, due to gathering consent and assent, administering the PPVT-III, and running demonstrations of the social task. The second session lasts about 1 hr. For each completed session, a participant receives compensation in the form of gift cards.

The equipment setup includes a computer dedicated to the social interaction tasks where the participants interacts with the VR environment, Biopac biological feedback equipment that collects physiological signals of the participant, and another PC dedicated to acquiring signals from the Biopac system (see Figure 4.4). The Vizard Virtual Reality Toolkit ran on a computer (C1) connected to the Biopac system via a parallel port to transmit task-related event-markers (e.g., start and stop of a trial). The physiological signals along with the event markers were acquired by the Biopac system and sent over an Ethernet link to the Biopac computer (C2). We also video recorded the sessions to cross-reference observations made during the experiment. The clinical observer/therapist and a participant's parent/caregiver watched the participant from the view of the video camera, whose signal was routed to a television hidden from the view of the participant.

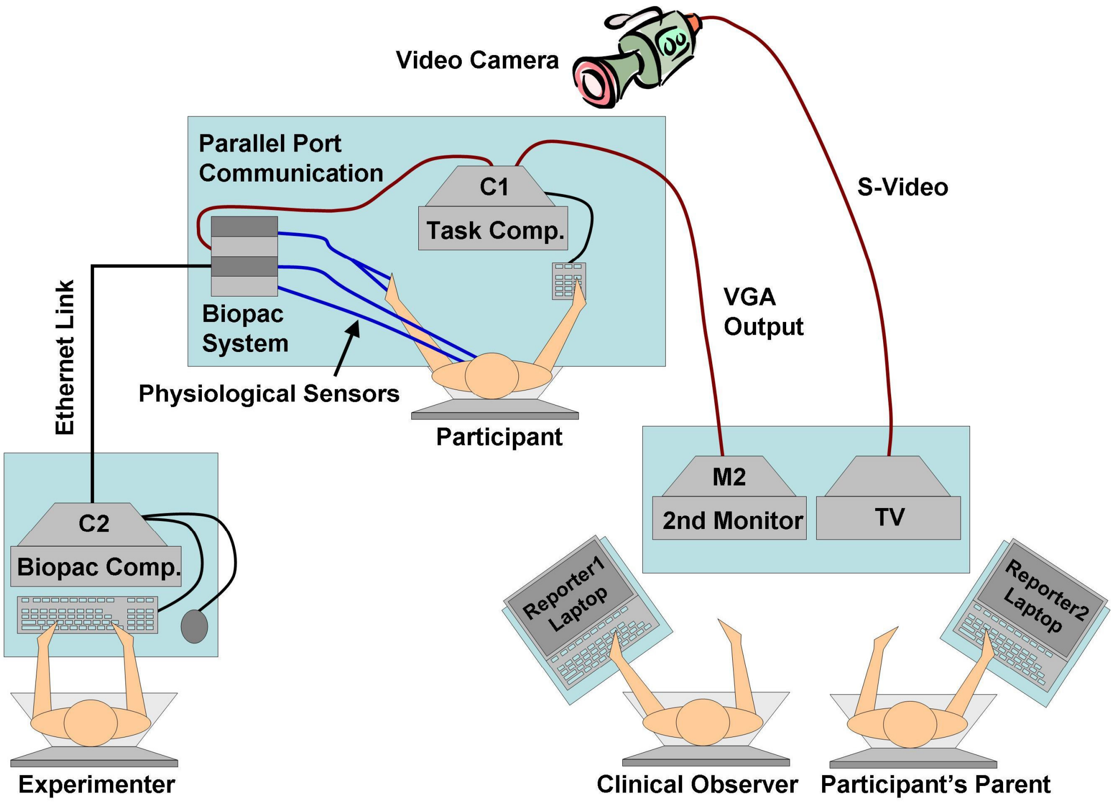


Figure 4.4 Equipment setup for collecting physiological data and subjective reports in the VR social tasks

The signal from the participant's computer screen where the task was presented was routed to a separate computer monitor (M2) so that the clinical observer and parent could view how the task progressed.

Each participant engages in two VR-based social interaction sessions on two different days. During the first session, the participants are told about the experiment purpose, the sensors, and the VR tasks. After the physiological sensors are placed, the participants are asked to relax quietly for three minutes while a resting/baseline recording of physiological signals is taken. The first session includes two demonstrations of the VR task, the resting/baseline physiological measurement, and a set of eight 2-min trials with different virtual social peers. The second session consists of the resting/baseline physiological measurement and the remaining 16 trials of social interaction tasks. After each trial, the participant answers the story question and self report questions on affective states. The clinical observer and parent/caregiver also rate what they think the level (i.e., low or high) of the affective states of anxiety, engagement, and enjoyment/liking was for the participant during the finished trial.

Results

After completing sessions with a few participants, one pair was pulled out for initial data analysis. Results from two male children, one with ASD (A1) and one TD (T1), are presented in this section. Their characteristics are shown in Table 4.3. A1 had a confirmed diagnosis using DSM-IV criteria (American Psychiatric Association, 2000) as well as scores from ADOS-G, ADI-R, SRS, and SCQ assessments. The total score, combining Communication and Social Interaction scores, on the ADOS-G (Autism Diagnostic Observation Schedule-Generic; Lord et al., 2000) has a cutoff for autism spectrum of 7 and a cutoff for autism of 10. The total score, combining Social Interaction, Communication, Repetitive Activities, and Stereotyped Early Development scores, on the ADI-R (Autism Diagnostic Interview-Revised; Rutter et al., 2003b) has a cutoff for autism spectrum of 22. A1 had a total score on the ADOS-G and ADI-R of 9 and 63, respectively, and scored above ASD cutoffs on the SRS and SCQ. T1's scores on the SRS

Table 4.3 Participant characteristics for initial analysis of system. The participants were matched by gender, age, and PPVT standard score.

Participant ID	Age (years)	PPVT[g]	SRS[h]	SCQ[i]
A1 (male)	17.333	119	78	30
T1 (male)	17.667	118	37	3

[g]Peabody Picture Vocabulary Test-3rd edition Standard score (Dunn and Dunn, 1997)
[h]Social Responsiveness Scale Total *T*-score (Constantino, 2002)
[i]Social Communication Questionnaire Total score (Rutter et al., 2003a)

and SCQ of 37 and 3, respectively, did not meet cutoffs for ASD. Both participants completed two sessions with the VR system. Percent accuracy on correctly answering the story questions revealed that the participants attended to the task; A1 and T1 achieved 100% accuracy on the story questions.

Evidence of overt behaviors as well as more subtle reactions to the different experiment conditions was demonstrated. Participant A1 showed considerable reactions to the virtual social peers standing at the invasive distance or using increased amounts of eye contact by temporarily leaning far back from or looking away from the monitor when they appeared on screen. In post-interview, his mother was surprised to observe such a stark reaction to the change in stimuli. Although accustomed to withdrawing behavior in complex, overwhelming social situations, the mother agreed that the story content and avatars' facial expressions were neutral and was therefore perplexed to see such a reaction from her child to the change in distance or eye gaze alone. These reactions and reflections highlight an advantage that systems like the virtual environment system for social interaction described here can provide to autism intervention. Because such technology can focus on each element of an interaction, minimizing distractions, and can do so with realistic representations of real-world settings; the VR system can systematically manipulate each element of an interaction and observe the effect. Therefore, the VR system can go beyond identifying a broad scope of situations that are affect-inducing. This system can pinpoint what components of a situation bring about an affective reaction to identify which specific component could be a vulnerability during social interaction.

The results of the values of Kappa (*K*) statistics, averaged across the three affective states showed that among the three possible reporter pairs for A1 (Therapist-Parent (T/P), Therapist-Child (T/C), and Parent-Child (P/C)) the agreement between the therapist and A1's parent (T/P) had the largest value (mean *K*-statistic for A1, T/P = 0.34, T/C = 0.22, P/C = 0.27). T1's parent was unable to participant in the experiment sessions, but the T/C *K*-statistic for T1 was also small (mean *K*-statistic for T1, T/C = 0.18). The same therapist/clinical observer was involved in all of the experiment sessions, which aided in establishing a consistent reporter. Based on previous work from Chapters II and III on reporter reliability and consistency, results from *K* analysis of current results, previous observations of participants reporting constant ratings even as tasks varied, the common use of an observer or parent for rating emotions of children (Eisenberg et al., 1995), and the possible unreliability of self-reports on emotions from adolescents with an without ASD (Barkley, 1998; Hill et al., 2004), as a experiment design methodology reports from the therapist were used whenever relating the objective physiological data to the children's affective states.

The overt reactions reflected ratings on affective states from the clinical observer and the subtle variations in physiological signals during the experiment trials. The correlation between the physiological features and subjective ratings for A1 and T1 are listed in Table 4.4. A1 and T1 showed significant correlations in their variations of physiological reactions versus variations in affective states. Therefore, the VR system shows it can elicit variations in both affective ratings and physiological signals to changes in social experimental stimuli. The findings are similar to observations in social anxiety research of typical adults in real-world settings (Argyle and Dean, 1965;

Schneiderman and Ewens, 1971; Sommer, 1962) but have now been examined with observations and physiological signals for children in a virtual interaction. This is the first step towards examining how children react to and accept the virtual social peers as realistic to real-world settings. Establishing realistic interactions builds a basis for creating more complex settings for social communication intervention. Further examination of reactions to the social stimuli from an expanded group of participants is presented in Chapter V.

Discussion

Social communication and social information processing are thought to represent core domains of impairment in children with ASD. This research may enhance our ability to understand the specific vulnerabilities in social communication of children with ASD as well as provide comparisons to a TD population. This system can identify specific social communication deficits for individual children and may determine patterns of physiological response across participants. Investigation using virtual social peers is not only cost and time efficient, it is also necessary to understand the complexity of social tasks. Systematic manipulation of facial expressions, eye gaze, social distance, vocal tone, and gestures need to be studied with virtual social peers where such manipulation is easy to perform, repeatable, and highly controllable. Questions on social communication can be exhaustively explored using controlled studies in virtual environments with virtual social peers interacting with the target population. Studies like the one presented here will provide insight on how such virtual peers display intentions, how they should interact, and how their interactions with children with ASD should be regulated. In that sense, this

Table 4.4 Individual analysis of significant correlations between affective state ratings and physiological features for each participant (A1 and T1). Reports on affective states across all 24 experiment trials were correlated with the 51 extracted physiological features from the raw signals. The physiological features are the same as listed in Table 2.1. Only features with significant correlation ($p < 0.05$) are shown. The sign of the correlation defines a direct/positive or inverse/negative relationship.

Participant	Affective State	Physiological Feature	Pearson Correlation, r	Significance, *p*-value
A1	Anxiety	IBI_ECGstd	0.5893	0.0024
		PTTstd	0.4377	0.0325
		D3_HSstd	0.6273	0.0010
		D4_HSstd	0.5707	0.0036
		Tonicmean	-0.4387	0.0320
		Tonicslope	0.4289	0.0365
		Corstd	0.5835	0.0028
		Blink_Peakmean	0.6156	0.0014
		Blinkstd	0.4819	0.0171
		Trapstd	0.4659	0.0218
	Enjoyment	IBI_ECGstd	-0.6589	0.0005
		IBI_PPGstd	-0.5715	0.0035
		PTTstd	-0.5798	0.0030
		D3_HSstd	-0.6798	0.0003
		D4_HSstd	-0.7154	8.5223e-5
		Tonicmean	0.5741	0.0034
		Tonicslope	-0.4996	0.0129
		Corstd	-0.5692	0.0037
		Blink_Peakmean	-0.6938	0.0002
		Blinkstd	-0.4730	0.0196
		Trapstd	-0.4828	0.0169
		Tempmean	0.5204	0.0091
	Engagement	Para	-0.4421	0.0306
		IBI_ECGstd	-0.6803	0.0003
		IBI_PPGstd	-0.4845	0.0164
		PTTstd	-0.4335	0.0343
		D3_HSstd	-0.5952	0.0022
		D4_HSstd	-0.5995	0.0020
		Tonicmean	0.5003	0.0128
		Tonicslope	-0.4256	0.0381
		Corstd	-0.6586	0.0005
		Blink_Peakmean	-0.6000	0.0019
		Blinkstd	-0.5044	0.0120
		Zygstd	-0.4111	0.0460
		Trapstd	-0.4589	0.0241
		Tempmean	0.4313	0.0354
T1	Anxiety	Para	0.4938	0.0142
		IBI_ECGstd	0.6111	0.0015
		PPG_Peakmean	-0.4541	0.0258
		IBI_PPGmean	-0.5547	0.0049
		PTTstd	0.4136	0.0445
		PEPstd	0.4672	0.0213
		IBI_ICGstd	0.6072	0.0017
		Trapstd	0.4768	0.0185
		Tempmean	-0.4344	0.0339
		Tempstd	0.4492	0.0277

Table 4.4 (continued) Individual analysis of significant correlations between affective state ratings and physiological features for each participant (A1 and T1). Reports on affective states across all 24 experiment trials were correlated with the 51 extracted physiological features from the raw signals. The physiological features are the same as listed in Table 2.1. Only features with significant correlation ($p < 0.05$) are shown. The sign of the correlation defines a direct/positive or inverse/negative relationship.

Participant	Affective State	Physiological Feature	Pearson Correlation, r	Significance, *p*-value
T1	Enjoyment	IBI_ECGstd	-0.4889	0.0153
		IBI_ICGstd	-0.4917	0.0147
		Trapstd	-0.4900	0.0151
		Tempstd	-0.4238	0.0390
	Engagement	IBI_ECGstd	-0.6419	0.0007
		PEPstd	-0.5314	0.0075
		IBI_ICGstd	-0.6055	0.0017
		Corstd	-0.4390	0.0319
		Zygstd	-0.4146	0.0440
		Trapstd	-0.6060	0.0017

work is one of the first that presents a design platform for virtual peers for specific applications that is analogous to well-adopted practices in the manufacturing industry where computer-aided design inevitably precedes any manufacturing. The design of integrating VR social interaction tasks and biofeedback sensor technology is novel yet relevant to the current priorities of technology-assisted ASD intervention. With further development to equip the system with real-time affect recognition capabilities, an autism intervention paradigm could use the system in the future for adaptively responding to the effects of elements of social interaction that lead to struggles in social communication for children with ASD.

CHAPTER V

PHYSIOLOGICAL ANALYSIS OF AFFECTIVE REACTIONS DURING SOCIAL INTERACTION

Introduction

A growing number of studies have been exploring the capability VR technologies to address the social communication deficits of children with ASD. Initial results indicate that VR holds promise as a potential alternative intervention approach with broad accessibility (Parsons and Mitchell, 2002). VR-based systems have demonstrated the advantages of this technology to facilitate children with ASD to learn social skills in designed tasks (Strickland et al., 1996). However, to date the capability of the VR technology has not been fully explored to examine the factors that lead to difficulties in social communication, which could be critical in designing an efficient intervention plan. VR paradigms may offer a way to combine the strengths from cognitive behavioral interventions as well as skills-based approaches (Parsons and Mitchell, 2002), while utilizing potential strengths for individuals with ASD (i.e., visual and auditory learning paradigms). A virtual world that allows for controllable complexity and minimized distractions may be less intimidating or confusing for children with ASD to interact with and hence models simplified but embodied social interaction (Moore et al., 2000; Standen and Brown, 2005). Since VR mimics real environments in terms of imagery and contexts, it may allow for efficient generalization of skills from the VR environment to the real world (Cromby et al., 1996). Furthermore, while changing and controlling environments is challenging during real-world interventions, VR possesses the advantage of being a robust but flexible system that can reliably repeat as well as adaptively modify environments across and within contexts (Sherman and Craig, 2003).

The primary objective of this chapter is to evaluate a VR-based social interaction system capable of objectively identifying specific communication aspects that induce an affective response in ASD and TD individuals by using a physiology-based approach. This research may identify significant differences in responses to elements of social interaction between the groups as well as within-group analysis and may enhance our ability to understand and tailor interventions to the specific vulnerabilities in social communication of children with ASD. The results could provide valuable information to caregivers and clinicians about the specific affect-eliciting aspects of social communication such that this feedback could drive behavioral interventions that scaffold skills from basic levels of comfort. In the future, the system could be developed into an intervention tool for detecting and adaptively responding to the effects of components of social interaction that lead to struggles in social communication in children with ASD. The ability to detect the physiological processes that are a part of impairments in social communication may also prove important for understanding the physiological mechanisms that underlie the presumed core impairments associated with ASD themselves.

This chapter describes an investigation into socially-driven VR interactions to guide future intervention of children with ASD. The VR environment monitors affective changes during social contexts for children with ASD and TD children. In particular, we study how the affective states of anxiety, engagement, and enjoyment/liking; measured by ratings from a clinical observer and a participant's physiological signals; vary with respect to the variation of specific communication factors (e.g., social distance and eye contact) presented in the virtual environment. Finally, we discuss how our findings, regarding the interaction with the VR social avatars, define similarities and differences in responses between and within the two groups.

Experimental Investigation

Participants

A group of 13 (10 male) children with ASD and a matched group of TD children age 13-18 years old participated in the VR experiment. Their characteristics are shown in Tables 5.1 and 5.2. The majority of male participants is reflective of the autism community, which has been found to have a male to female ratio of 4:1 (Ehlers and Gillberg, 1993). All ASD participants had a confirmed diagnosis from evaluations by a licensed clinical psychologist using DSM-IV criteria according to their medical records. Additionally, all but three participants in the ASD group met cutoffs for ASD according to ADOS and ADI-R assessments. P9, P10, and P12 did not have ADOS or ADI-R records; however, their scores on the SRS and SCQ questionnaires meet ASD cutoffs. None of the participants in the TD group met ASD cutoffs on the SRS or SCQ questionnaires.

All 26 participants underwent the PPVT-III (Peabody Picture Vocabulary Test-3rd edition) to assess cognitive function (Dunn and Dunn, 1997). The PPVT-III is a measure of single-word receptive vocabulary that is often used as a proxy for IQ testing because of its high correlations with standardized tests such as the Wechsler Intelligence Scale for

Table 5.1 Participant Characteristics. The participants were matched by age and PPVT standard score as well as gender and handedness (see Table 5.2). No significant group difference was found for gender, age, handedness, and standard score on the PPVT.

Participant (Gender)	**Age (years)**	**PPVT[a] Standard score**	**SRS[b] Total *T*-score**	**SCQ[c] Total score**	**ADOS-G[d] Total score (cutoff = 7)**	**ADI-R[e] Total score (cutoff = 22)**
ASD						
P1 (Male)	16.500	97	63	17	9	49
P2 (Male)	17.250	112	90	23	11	51
P3 (Male)	13.000	133	73	17	7	31
P4 (Male)	13.833	126	69	23	11	56
P5 (Male)	17.333	119	78	30	9	63
P6 (Male)	15.500	110	73	13	7	33
P7 (Male)	16.167	98	65	18	10	44
P8 (Female)	15.167	83	90	28	10	62
P9 (Male)	15.083	92	87	20	–	–
P10 (Female)	18.583	102	87	14	–	–
P11 (Female)	18.083	95	90	31	13	68
P12 (Male)	14.333	107	85	20	–	–
P13 (Male)	17.500	103	83	31	20	51
Group Mean (SD)	16.0 (1.7)	105.9 (14.0)	79.5 (9.9)	21.9 (6.3)	10.7 (3.7)	50.8 (12.3)
TD						
P21 (Male)	15.833	124	42	4	–	–
P22 (Male)	17.333	110	36	2	–	–
P23 (Male)	13.167	119	40	6	–	–
P24 (Male)	13.333	103	53	2	–	–
P25 (Male)	17.667	118	37	3	–	–
P26 (Male)	13.917	124	40	1	–	–
P27 (Male)	14.833	128	47	1	–	–
P28 (Female)	15.417	101	46	8	–	–
P29 (Male)	15.583	117	38	0	–	–
P30 (Female)	18.167	100	38	1	–	–
P31 (Female)	16.167	102	52	1	–	–
P32 (Male)	14.083	135	36	5	–	–
P33 (Male)	17.167	97	40	9	–	–
Group Mean (SD)	15.6 (1.7)	113.7 (12.3)	41.9 (5.8)	3.3 (2.9)	–	–
t-value	0.66	1.50	11.84	9.62		
p-value	ns	ns	<0.001	<0.001		
Exact *p*-value	0.5175	0.1468	1.6500e-11*	1.0341e-9*		

[a]Peabody Picture Vocabulary Test-3rd edition (Dunn and Dunn, 1997)
[b]Social Responsiveness Scale (Constantino, 2002)
[c]Social Communication Questionnaire (Rutter et al., 2003a)
[d]Autism Diagnostic Observation Scale-Generic: Module 3 or 4 depending upon subject's developmental level (Lord et al., 2000)
[e]Autism Diagnostic Interview-Revised (Rutter et al., 2003b)
Significant group differences, *p<0.001.
No significant group differences were found for the age or PPVT standard score variables (p>0.05 for all).

Table 5.2 Group frequencies. The participant groups were matched by gender and handedness as well as age and PPVT standard score (see Table 5.1).

Group	Gender Frequency Male/Female	Handedness Frequency Right/Left
ASD	10/3	12/1
TD	10/3	12/1

Children (Bee and Boyd, 2004). It provides standard scores with a mean of 100 and a standard deviation of 15, and the DSM-IV classifies full scale IQ's above 70 as non-retarded (American Psychiatric Association, 2000). Participants in this study obtained a standard score of 80 or above on the PPVT-III measure.

The SRS (Social Responsiveness Scale) is a 65-item, 15-min. parent-report questionnaire designed to quantitatively measure the severity of autism-related symptoms. This measure provides an index of ASD-related social competence with questions related to social awareness, social information processing, capacity for reciprocal social communication, social anxiety/avoidance, and autistic preoccupations and traits. The SRS has been shown to correlate on the order of 0.7 with the ADI-R (Constantino et al., 2003). Behaviors and characteristics are rated on a 4-point scale that ranges from "Not True" to "Almost Always True." The SRS generates a total *T*-score reflecting severity of social deficits in the autism spectrum, as well as five Treatment Subscales: Receptive, Cognitive, Expressive, and Motivational aspects of social behavior, and Autistic Preoccupations. The *T*-score categorizes measurements in the Normal Range (≤59T), Mild to Moderate ASD Range (60*T*-75*T*), or Severe Range (≥76*T*) (Constantino, 2002). All of our TD participants scored within the Normal Range. Five ASD participants ranked within the Mild or Moderate Range (P1, P3, P4, P6, and P7) with the remaining eight falling into the Severe Range (P2, P5, P8, P9, P10, P11, P12, and P13).

The SCQ (Social Communication Questionnaire) is a brief instrument for the valid screening or verification of ASD symptoms in children that has been developed from the critical items of the ADI (Autism Diagnostic Interview) and compiled into a parent report questionnaire (Rutter et al., 2003a). As in the ADI, these questions tap the three critical autism diagnostic domains of qualitative impairments in reciprocal social interaction, communication, and repetitive and stereotyped patterns of behavior. Among 200 children and adolescents, domain scale scores of the SCQ were significantly correlated with corresponding scores derived from the full ADI ($r = 0.55$ to 0.71, $p < 0.005$) (Berument et al., 1999). Analysis indicated that the SCQ was comparable to the ADI in discriminating ASD from non-ASD, autism vs. mental retardation, and autism vs. other aspects of ASD. A cutoff score of 13 is recommended to maximize valid ascertainment of cases of ASD (specificity) while minimizing errors of omission (sensitivity). The SCQ was designed for use with children over the age of four years with a mental age of at least two years. All ASD participants met ASD cutoffs on the SCQ while no participant in the TD group did.

The ADOS-G (Autism Diagnostic Observation Schedule-Generic) is a 45-min. semi-structured standardized observational assessment of play, social interaction, and communicative skills that was designed as a diagnostic tool for identifying the presence of autism (Lord et al., 2000). It is organized into four modules, which are distinguished by their appropriateness for use with individuals functioning at different developmental levels, ranging from nonverbal children to highly fluent adults. Each module provides a set of behavioral ratings in five domains: Language and Communication, Reciprocal Social Interaction, Play or Imagination/Creativity, Stereotyped Behaviors and Restricted

Interests, and Other Abnormal Behaviors. The scoring algorithm provides cutoffs that can be used to discriminate between a diagnosis of autism, autism spectrum, or non-spectrum. Across all modules, inter-observer agreement for the algorithm score was 0.92, and the test-retest correlation was 0.82. Agreement about diagnostic classification (autism vs. autism spectrum vs. non-spectrum) ranged from 81%-93% (Lord et al., 2000). After coding ratings on the five domains, a total score on the two main components of Communication and Reciprocal Social Interaction equal to or above 7 would indicate autism spectrum, and a score of 10 or more would indicate autistic disorder.

The ADI-R (Autism Diagnostic Interview-Revised) is a semi-structured, investigator-based interview for parents/caregivers that was developed for the purpose of diagnostic classification of individuals who may have autism or other pervasive developmental disorders (Rutter et al., 2003b). This interview covers areas of background and history, early development, acquisition and loss of skills, language and communication, social development and play, favorite activities/toys, interests and behaviors, and general behaviors. The ADI-R provides explicit scoring criteria that yield cutoff scores in the domains of social reciprocity, language and communication, and restricted and repetitive activities. The scores from a subset of critical items of the ADI-R are summed to yield scores for each domain; cutoffs are used to determine whether the individual's diagnostic classification is consistent with an autism spectrum disorder. This measure possesses strong psychometric properties in terms of inter-observer agreement, internal consistency, and test-retest reliability. The ADI-R has been found to discriminate autism from non-autism in individuals with mental ages of at least 18 months (Lord et al., 1997). A total score on the four domains: Reciprocal Social Interaction, Communication, Restricted and Repetitive Patterns of Behavior, and Evidence of Abnormal Development before 36 months of age, of the ADI-R equal to or above 22 would indicate autistic disorder (Rutter et al., 2003b).

Procedure

Our hypothesis was that manipulation of the social parameters may elicit variations in affective reactions (Argyle and Dean, 1965; Bancroft, 1995) and physiological responses (Farroni et al., 2002; Groden et al., 2005). A participant is likely to experience a range of short-lived affective states (i.e., emotions such as anxiety, interest, etc.) as he/she interacts with the VR system. However, these feelings should not be more intense than the levels of these emotions that are commonly experienced in daily life and should not carry over when the participant leaves the laboratory. Physiological signals from the participant and ratings of affective states from a clinical observer, a participant's parent, and self-reports from the participant were recorded during the 2-min. experiment trials. The interval between the stop and start of a trial to collect reports and allow the participant to choose when to begin the next trial ranged from approximately 1-2.5 min. The equipment setup for the VR experiment is illustrated in Figure 4.4. A photo of the therapist and parent in front of the second computer monitor, TV, and their reporter laptops is shown in Figure 5.1. We were unable to attain parent participation for the TD children, but a parent of each ASD child participated in the two experiment sessions for his/her child. The participant, clinical observer, and parent (ASD group only) reported on levels of anxiety, enjoyment/liking, and engagement after each experiment trial. The participant answered on a 3-point scale (e.g., low-Disengaged, medium-Neutral,

Figure 5.1 Photo of clinical observer, on left, and a participant's parent, on right, sitting in front of the second computer monitor, TV, and their reporter laptops; printed with permission.

high-Engaged) for simplicity and consistency with the story question answer choices. The clinical observer and parent rated what they thought the level of the affective state was for the participant during the finished trial on a 9-point scale (e.g., 1-Very Disengaged, 5-Neutral, 9-Very Engaged) and a binary scale (e.g., Low Engagement or High Engagement). The 9-point scale was used to compare against the participant's reports for calculating *K*-statistics. The binary scale was used to label trials as "high" or "low" for data analysis.

A participant with the Biopac sensors attached and positioned in front of the VR task computer is pictured in Figure 5.2. The physiological signals were processed to

Figure 5.2 Photo of a participant sitting in front of the VR task computer; printed with permission. The Biopac MP150 equipment and amplifier modules for each physiological sensor sit on top of a black cart in the foreground.

extract features, which are the individual measurable properties of the physiological signals that could be correlated to affective states. Extracted features from the signals were compared to subjective reports to relate the participant's affective reactions and physiological responses with respect to the various social stimuli. The physiological signals recorded in this work are the same as those described in Chapter II with the features listed in Table 2.1. These signals were collected using a Biopac MP150 system (biopac.com) and small wearable sensors placed on a participant's left eyebrow (Corrugator Supercilii EMG), left cheek (Zygomaticus Major), upper back/lower neck muscle on right (Upper Trapezius EMG), chest (ECG and Heart Sound), neck and torso

(ICG), ring and pointer finger of left hand (GSR), middle finger of left hand (PPG), and thumb on the participant's left hand (Skin Temperature). Participants used their right hand to press a keypad for interactions with the VR system. The sensors have been successfully used to collect physiological data of typical individuals (Rani et al., 2006) and children with ASD in the Chapter II and Chapter III experiments.

Results

Each child in the two groups completed two sessions with the virtual environment system for social interaction. Results from the accuracy of correctly answering the story questions revealed that the participants attended to the task. Percent accuracy for the ASD and TD group was 97% and 99%, respectively, and no group difference was found (p>0.05, p=0.2601). The results from *K*-statistic again revealed that the Therapist-Parent (T/P) reporter pair had the largest mean when compared to the Therapist-Child (T/C) and Parent-Child (P/C) pairs (mean *K*-statistic for ASD group: T/P = 0.21, $p < 0.05$) for the ASD group. The means of the *K*-statistic values between the children and either the therapist or parent were smaller (T/C = 0.11, P/C = 0.06). The mean *K*-statistic for the TD group when comparing reports between the therapist to the TD children was also small (mean *K*-statistic for TD group: T/C = 0.04). The same therapist was involved in all of the experiment sessions for the ASD and TD groups, and the subjective report from the therapist was used as the reference points linking the objective physiological data to the participant's affective states.

The ASD group's physiological signals showed significant changes to trials rated as eliciting "low anxiety" (LA) versus "high anxiety" (HA). The TD children also showed significant physiological reactions to the experimental stimuli for trials rated as LA or HA in similar and different ways than their ASD counterparts. Reactions occurred for changes in social distance and eye gaze. The features of interest include GSR phasic response rate (Phasicrate), PPG peak maximum (PPG_Peakmax), and mean time of pre-ejection period (PEPmean). Phasicrate is measured in peaks per minute (ppm) and represents a rapid increase in skin conductance similar to a peak. As shown in Table 5.3, both the ASD and TD group had a significant increase in Phasicrate, between trials rated as LA and HA for trials in which the social distance parameter was set to Invasive for all variations of the eye gaze parameter. As anxiety increased in these conditions, Phasicrate significantly increased within the groups, and between the groups the ASD children shown more phasic peaks per minute during LA and HA.

Table 5.3 Listed are results of Phasicrate from the GSR signal compared between trials labeled as LA and HA. The trials considered were ones in which the social distance parameter was set to Invasive for all variations of the eye gaze parameter.

Clinical Observer label	**ASD group Phasicrate (ppm) *M (SD)***	**TD group Phasicrate (ppm) *M (SD)***	**Group Differences**	
LA	4.43 (2.75)	3.23 (2.78)	Exact *p*-value	0.0227*
HA	5.80 (3.55)	4.46 (3.57)	Exact *p*-value	0.0286*
t-value	-2.18	-2.33		
Exact *p*-value	0.0311*	0.0211*		

Significant difference, *p<0.05

Other conditions showed contrasting results for TD and ASD groups. Between trials labeled LA and HA for experiment condition of Straight eye gaze with distance varying, the TD group had a significant increase in PPG_Peakmax, but the ASD group did not (see Table 5.4). The difference between groups for this feature was significant for LA but not HA. PPG_Peakmax is the maximum amplitude of detected PPG peaks, measured in μV, which showed significant differences in this study. PEP is calculated as the difference in onset of ICG time-derivative peak to onset of ECG R peak and is measured in ms. The experiment condition of Averted eye gaze with distance varying elicited a significant increase in the mean value of PEP within the ASD group but not the TD group (see Table 5.5) and showed higher values for the ASD group compared to the TD group for LA but not HA.

Table 5.4 Listed are results of PPG_Peakmax compared between trials labeled as LA and HA. The trials considered were ones in which the eye gaze parameter was set to Straight for all variations of the social distance parameter.

Clinical Observer label	**ASD group PPG_Peakmax (μV) *M (SD)***	**TD group PPG_Peakmax (μV) *M (SD)***	**Group Differences**	
LA	2.93 (3.12)	1.21 (1.18)	Exact *p*-value	0.0002*
HA	4.24 (4.39)	3.42 (4.67)	Exact *p*-value	0.5245
t-value	-1.52	-3.37		
Exact *p*-value	0.1329	0.0012*		

Significant difference, *p<0.05

Table 5.5 Listed are results of PEPmean compared between trials labeled as LA and HA. The trials considered were ones in which the eye gaze parameter was set to Averted for all variations of the social distance parameter.

Clinical Observer Label	**ASD group PEPmean (ms) *M (SD)***	**TD group PEPmean (ms) *M (SD)***	**Group Differences**	
LA	156.61 (23.49)	145.54 (20.91)	Exact *p*-value	0.0110*
HA	143.85 (26.57)	146.79 (19.63)	Exact *p*-value	0.6645
t-value	2.13	-0.25		
Exact *p*-value	0.0367*	0.8044		

Significant difference, *p<0.05

When the eye gaze was 100% direct (Straight), it overpowered the distance parameter for the ASD group. For these conditions the ASD group found the Invasive and Decorum distance similarly anxiety-inducing (i.e., the Straight gaze caused too much anxiety for the distance to cause degradation of anxiety), but TD children were able to discern differences in these conditions. For 100% indirect eye gaze (Averted), the ASD group showed a significant difference in physiology for ratings of LA versus HA while the distance parameter varied from Invasive to Decorum, but TD children reacted equally to the different settings. Distance did not cause TD children to become more anxious when the eye gaze was minimal, but the ASD children showed a significant change to these experiment conditions.

As reports on enjoyment/liking varied from "low liking" (LL) to "high liking" (HL), physiological signals also varied significantly. The features of interest include the SD of the IBI of the ECG signal (IBI_ECGstd), SD of pulse transit time (PTTstd), and mean value of skin temperature (Tempmean). IBI_ECGstd is the SD of the interval of time between R peaks of the ECG signal; SD has the same unit of the IBI of ECG which is measured in ms. As shown in Table 5.6, the ASD group showed a significant decrease in IBI_ECGstd between LL and HL, but the TD group did not. The difference between groups was not significant for either level of liking/enjoyment. This result is for trials in which eye gaze was set to Straight for variations in the social distance parameter.

Table 5.6 Listed are results of IBI_ECGstd compared between trials labeled as LL and HL. The trials considered were ones in which the eye gaze parameter was set to Straight for all variations of the social distance parameter.

Clinical Observer Label	**ASD group IBI_ECGstd (ms) *M (SD)***	**TD group IBI_ECGstd (ms) *M (SD)***	**Group Differences**	
LL	88.30 (27.08)	73.89 (37.89)	Exact *p*-value	0.0524
HL	71.62 (22.27)	72.56 (35.99)	Exact *p*-value	0.8981
t-value	2.83	0.17		
Exact *p*-value	0.0060*	0.8758		

Significant difference, *$p<0.05$

PTT is the time in ms for a pulse to travel from the heart to the periphery and is thus calculated from the ECG and PPG signals. For trials in which the distance was set to Invasive for all variations in eye gaze, the SD of PTT decreased significantly for the TD group, and PTTstd was smaller for the TD group than the ASD group for HL (Table 5.7). Also for these conditions of distance and eye gaze, Tempmean did not vary significantly within the groups but did for LL and HL between the groups (Table 5.8). Tempmean is measured in °F in this experiment.

Table 5.7 Listed are results of PTTstd compared between trials labeled as LL and HL. The trials considered were ones in which the social distance parameter was set to Invasive for all variations of the eye gaze parameter.

Clinical Observer label	**ASD group PTTstd (ms) *M (SD)***	**TD group PTTstd (ms) *M (SD)***	**Group Differences**	
LL	145.57 (79.02)	126.76 (56.51)	Exact *p*-value	0.0542
HL	157.59 (45.60)	103.57 (51.51)	Exact *p*-value	0.0001*
t-value	-0.82	2.62		
Exact *p*-value	0.4111	0.0097*		

Significant difference, *p<0.05

Table 5.8 Listed are results of Tempmean compared between trials labeled as LL and HL. The trials considered were ones in which the social distance parameter was set to Invasive for all variations of the eye gaze parameter.

Clinical Observer label	**ASD group Tempmean (°F) *M (SD)***	**TD group Tempmean (°F) *M (SD)***	**Group Differences**	
LL	87.61 (6.77)	89.71 (5.03)	Exact *p*-value	0.0134*
HL	86.66 (6.66)	90.53 (4.18)	Exact *p*-value	0.0007*
t-value	0.71	-1.07		
Exact *p*-value	0.4780	0.2842		

Significant difference, *p<0.05

The results for significant changes in physiological signals to trials rated as eliciting "low engagement" (LE) versus "high engagement" (HE) within and between the ASD and TD groups are shown in Tables 5.9-5.11. The features of interest include

Tempmean measured in °F, SD of PTT, and mean values of the tonic signal from GSR (Tonicmean). For Straight eye gaze with varying distance, Tempmean increased significantly for the ASD group between LE and HE, but no difference was shown for the TD groups or between groups (see Table 5.9).

Table 5.9 Listed are results of Tempmean compared between trials labeled as LE and HE. The trials considered were ones in which the eye gaze parameter was set to Straight for all variations of the social distance parameter.

Clinical Observer Label	**ASD group Tempmean (°F)** ***M (SD)***	**TD group Tempmean (°F)** ***M (SD)***	**Group Differences**	
LE	84.98 (7.15)	88.66 (6.03)	Exact *p*-value	0.0663
HE	88.94 (5.83)	90.36 (4.56)	Exact *p*-value	0.1687
t-value	-2.65	-1.30		
Exact *p*-value	0.0098*	0.1981		

Significant difference, *p<0.05

For trials under the condition that eye gaze varied while social distance was fixed to the Invasive distance, PTTstd and Tonicmean varied for LE and HE. The SD of PTT decreased significantly for only the TD group and between the groups showed a significantly larger value for the ASD group for HE only (Table 5.10). The response of tonic GSR is the ongoing skin conductance in the absence of any particular peaks

Table 5.10 Listed are results of PTTstd compared between trials labeled as LE and HE. The trials considered were ones in which the social distance parameter was set to Invasive for all variations of the eye gaze parameter.

Clinical Observer label	**ASD group PTTstd (ms)** ***M (SD)***	**TD group PTTstd (ms)** ***M (SD)***	**Group Differences**	
LE	146.59 (88.08)	128.20 (57.73)	Exact *p*-value	0.1475
HE	150.11 (45.38)	109.58 (52.94)	Exact *p*-value	1.5925e-6*
t-value	-0.29	2.08		
Exact *p*-value	0.7695	0.0393*		

Significant difference, *p<0.05

(i.e., absence of phasic response) and is measured in μS. Tonicmean decreased in these experiment conditions between LE and HE for the ASD group only, and no difference is shown for changes in Tonicmean between the groups for LE and HE (Table 5.11).

Table 5.11 Listed are results of Tonicmean compared between trials labeled as LE and HE. The trials considered were ones in which the social distance parameter was set to Invasive for all variations of the eye gaze parameter.

Clinical Observer Label	**ASD group Tonicmean (μS) *M (SD)***	**TD group Tonicmean (μS) *M (SD)***	**Group Differences**	
LE	9.11 (3.81)	8.41 (5.09)	Exact *p*-value	0.3306
HE	7.88 (3.46)	8.71 (4.93)	Exact *p*-value	0.2450
t-value	2.06	-0.37		
Exact *p*-value	0.0410*	0.7135		

Significant difference, **p*<0.05

Discussion

This work is the first time a large set of physiological features have been examined for a sizeable group of ASD and TD children during interaction with social stimuli presented on a VR platform for elicitation of multiple affective states. The results show the VR system provokes variations in both affective ratings and physiological signals to changes in social experimental stimuli for children with ASD and TD children. This work used virtual social peers to systematically manipulate specific aspects of social communication and provides a vital step towards development of future social interventions using technologies such as VR for the ASD population. Since physiological signals have been shown to differentiate during social interaction with a virtual environment, the signals could be a useful measure in real-time VR-assisted social skill intervention, an important therapeutic instrument for addressing the core deficits in the ASD population, for adaptation of the interaction. Future work will involve developing an expanded set of VR social interaction scenarios and investigation of fast and robust learning mechanisms that would permit a virtual peer's adaptive response within complex social interaction tasks. Individual affective models could be made in the future with a sufficiently large number of examples of trails per participant to relate physiological features to affective states. Group models from a small number trials from each participant have been used to evaluate affective states of typical adults (Zhai and Barreto, 2006). If group models were to be created from the data collected in this experiment, a separate ASD and TD model would be necessary considering the significant physiological differences detected between the groups.

CHAPTER VI

CONTRIBUTIONS AND FUTURE WORK

Contributions

The contributions of this dissertation are in the area of psychophysiological analysis of affective computing. The research strategy utilizes a physiological approach to achieve the primary objective of developing technology-based assessment tools capable of identifying specific aspects of interaction that induce an affective response in individuals with ASD. The main contributions of this dissertation are:

1. Individual affective models are constructed from physiological signals and subjective reports collected during HCI involving children with ASD. Two computer-based cognitive tasks are designed to elicit the affective states of liking, anxiety, and engagement that are considered important in autism intervention. A large set of physiological indices are investigated that may correlate with the target affective states of children with ASD. Subjective reports on the affective states from a clinical observer, a parent, and the child are collected and analyzed. This work allows for the formation of affect-sensitive computer-based ASD intervention tools. Generally, an experienced therapist continuously monitors the affective cues of the children with ASD and adjusts the course of the intervention accordingly. In this work, we address the problem of how to make technology-based ASD intervention tools affect-sensitive by designing therapist-like affective models of the children with ASD based on their physiological responses. A SVM-based affective model yields reliable prediction with approximately 82.9% success

when using the therapist's reports. This is the first time, to our knowledge, that the affective states of children with ASD have been experimentally detected via a physiology-based affect recognition technique.

2. The impact of applying the individual affective models during closed-loop HRI is evaluated in Chapter III. This method presents a physiology-based affect-inference mechanism for robot-assisted intervention where the robot can detect the affective states of a child with ASD as discerned by a clinical observer and adapt its behaviors accordingly. Chapter III is the first step toward developing "understanding" robots for use in future ASD intervention. This is the first time, to our knowledge, that the affective states of children with ASD are detected via a physiology-based affect recognition technique in real time. This is also the first time that the impact of affect-sensitive closed-loop interaction between a robot and a child with ASD is demonstrated experimentally.

3. A VR-based social interaction system for exploring physiological responses to social communication is designed and developed. Social interaction modules for affect elicitation are created using virtual environments. In Chapter IV, the VR system is formulated to present realistic social communication tasks to the children with ASD and can monitor their affective response using physiological signals. Examination of responses of TD children is also made for comparison. This system is capable of systematically manipulating specific aspects of social communication to more fully understand its salient components.

4. The physiological response from interacting with the VR platform is evaluated to explore affective response characteristics during social interaction for the ASD and TD groups. Chapter V measures the discriminating capability of the physiological features

and identifies ones that have significant influence during social communication in the VR environment for the children with ASD and TD children. The affective states of anxiety, engagement, and enjoyment are measured by physiological signals and examined for how they vary with respect to the variation of specific communication factors (e.g., social distance and eye contact). This work is the first time a large set of physiological features have been examined for a sizeable group of ASD and TD children during interaction with social stimuli presented on a VR platform for elicitation of multiple affective states. The results show the VR system provokes variations in both affective ratings and physiological signals to changes in social experimental stimuli for children with ASD and TD children.

There is increasing consensus that development of assistive therapeutic tools can make application of intensive intervention for children with ASD more readily accessible. In recent years, various applications of advanced interactive technologies have been investigated to facilitate and/or partially automate the existing behavioral intervention that addresses specific deficits associated with autism. However, the current technology-assisted therapeutic tools for children with ASD do not possess the ability of deciphering the affective cues of the children, which could be critical given that the affective factors of children with ASD have significant impacts on the intervention practice. In Chapter II, a physiology-based affect modeling framework for children with ASD was presented. The developed model could allow the recognition of affective states of the child with ASD from the physiological signals in real time and provide the basis for computer-based affect-sensitive interactive autism intervention. In Chapter III, how to augment the interactive autism intervention was investigated by having a robot respond appropriately to the inferred level of a target affective state based on the affective model described in Chapter II. VR-based intervention tools that address the social communication deficits of children with ASD were developed in Chapter IV and evaluated in Chapter V.

Future Work

The physiology-based affect-sensitive technology described here could be employed to develop new intervention paradigms, which could promote interventions for individuals with ASD that are practical, widely available, and specific to the unique strengths and vulnerabilities of individuals with ASD. With further integration, a VR and physiological profiling system could be effective for use in developing and adapting controlled environments that help individuals explore social interaction dynamics gradually but automatically (i.e., introducing the aspects of social communication that are more challenging based on physiological data). Future work may include a reduction of the verbal components in the cognitive tasks which would allow application to the broader ASD population. Also, the research could benefit from exploring and merging other types of signals and features proven useful in affective computing, such as pupil diameter from eye-tracking data, with the current set of physiological signals. These ideas are currently being explored by researchers in our laboratory.

Note that the presented work requires physiological sensing that has its own limitations. For example, one needs to wear physiological sensors, and use of such sensors could be restrictive under certain circumstances. However, none of the participants in our studies had any objection in wearing the physiological sensors. Similar observations were achieved by Conati et al. (2003) that suggested concerns for intrusiveness of physiological sensors could be lessened for children in a game-like environment. Given the rapid progress in wearable computing with small, non-invasive sensors and wireless communication, physiological sensors can be worn in a wireless manner such as in physiological sensing clothing and accessories (Picard, 1997; Wijesiriwardana et al., 2004), which could alleviate possible constraints on experimental design. Physiology-based affect recognition can be appropriate and useful for the application of interactive autism intervention and could be used conjunctively with other modalities (e.g., facial expression, vocal intonation, etc.) to allow flexible and robust affective modeling for children with ASD.

Future work may also involve designing socially-directed interaction experiments with embodied robots interacting with children with ASD. For example, the real-time affect recognition and response system described here could be integrated with a life-like android face developed by Hanson Robotics (hansonrobotics.com), which can produce accurate examples of common facial expressions that convey affective states. This affective information could be used as feedback for empathy exercises to help children with ASD recognize their own emotions.

APPENDIX

A. PATTERN RECOGNITION USING SVM

SVM, pioneered by Vapnik (1998), is an excellent tool for classification (Burges, 1998). Its appeal lies in its strong association with statistical learning theory as it approximates the structural risk minimization principle. Good generalization performance can be achieved by maximizing the margin, where margin is defined as the sum of the distances of the hyperplane from the nearest data points of each of the two classes. SVM is a linear machine working in a high k-dimensional feature space formed by an implicit embedding of n-dimensional input data X into a k-dimensional feature space ($k > n$) through the use of a nonlinear mapping $\phi(X)$. This allows for the use of linear algebra and geometry to separate the data, which is normally only separable with nonlinear rules in the input space. The problem of finding a linear classifier for given data points with known class labels can be described as finding a separating hyperplane $W^T\varphi(X)$ that satisfies

$$y_i\left(W^T\varphi(X_i)\right) = y_i\left(\sum_{j=1}^{k} w_j\phi_j(X_i) + w_0\right) \geq 1 - \xi_i \qquad \text{(A.1)}$$

where N represents the number of training data pairs (X_i, y_i) indexed by $i = 1,2,\ldots, N$, $y_i \in \{+1, -1\}$ represents the class label, $\varphi(X) = [\phi_0(X), \phi_1(X),\ldots, \phi_k(X)]^T$ is the mapped feature vector [$\phi_0(X) = 1$], and $W = [w_0, w_1,\ldots, w_k]$ is the weight vector of the network.

The nonnegative slack variable ξ_i generalizes the linear classifier with soft margin to deal with nonlinearly separable problems.

All operations in learning and testing modes are done in SVM using a so-called kernel function defined as $K(X_i, X) = \varphi^T(X_i)\varphi(X)$ (Vapnik, 1998). The kernel function allows for efficient computation of inner products directly in the feature space and circumvents the difficulty of specifying the nonlinear mapping explicitly. One distinctive fact about SVM is that the learning task is reduced to a dual quadratic programming problem by introducing the Lagrange multipliers α_i (Vapnik, 1998). Maximize

$$Q(\alpha) = \sum_{i=1}^{N} \alpha_i - \frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j K\left(X_i, X_j\right) \tag{A.2}$$

subject to $\sum_{i=1}^{N} \alpha_i y_i = 0$ and $0 \leq \alpha_i \leq C$

where C is a user-defined regularization parameter that determines the balance between the complexity of the network characterized by the weight vector W and the error of classification of data. The corresponding α_i multipliers are nonzero only for the support vectors (i.e., the training points nearest to the hyperplane). The SVM approach is able to deal with noisy data and over-fitting by allowing for some misclassifications on the training set (Burges, 1998). This characteristic makes it particularly suitable for affect recognition because the physiology data is noisy and the training set size is often small. Another important feature of SVM is that the quadratic programming leads in all cases to the global minimum of the cost function. With the kernel representation, SVM provides an efficient technique that can tackle the difficult, high-dimensional affect recognition problem.

B. BEHAVIOR ADAPTATION USING *QV*-LEARNING

QV-learning (Wiering, 2005), a variant of the standard reinforcement learning algorithm *Q*-learning (Watkins and Dayan, 1992), was applied to achieve the affect-sensitive behavior adaptation in the Chapter III experiments. *QV*-learning keeps track of both a *Q*-function and a *V*-function. The *Q*-function represents the utility value $Q(s, a)$ for every possible pair of state s and action a. The *V*-function indicates the utility value $V(s)$ for each state s. The state value $V(s_t)$ and *Q*-value $Q(s_t, a_t)$ at step t are updated after each experience (s_t, a_t, r_t, s_{t+1}) by:

$$V(s_t) := V(s_t) + \alpha\left(r_t + \gamma V(s_{t+1}) - V(s_t)\right) \tag{B.1}$$

$$Q(s_t, a_t) := Q(s_t, a_t) + \alpha\left(r_t + \gamma V(s_{t+1}) - Q(s_t, a_t)\right) \tag{B.2}$$

where r_t is the received reward that measures the desirability of the action a_t when it is applied on state s_t and causes the system to evolve to state s_{t+1}. The difference between (B.2) and the conventional *Q*-learning rule is that *QV*-learning uses *V*-values learned in (B.1) and is not defined solely in terms of *Q*-values. Since $V(s)$ is updated more often than $Q(s, a)$, *QV*-learning may permit a fast learning process (Wiering, 2005) and enable the robot to efficiently find a behavior selection policy during HRI.

REFERENCES

Agrawal, P., Liu, C., and Sarkar, N. (2008). "Interaction between human and robot: An affect-inspired approach," *Interaction Studies*, vol. 9, no. 2, pp. 230–257.

American Psychiatric Association. (2000). *Diagnostic and Statistical Manual of Mental Disorders: DSM-IV-TR*, 4th ed. Washington, DC: American Psychiatric Association.

Argyle, M. and Cook, M. (1976). *Gaze and Mutual Gaze*. Cambridge, MA: Cambridge University Press.

Argyle, M. and Dean, J. (1965). "Eye-contact, distance and Affiliation," S*ociometry*, vol. 28, no. 3, pp. 289-304.

Asha, K., Ajay, K., Naznin, V., George, T., and Peter, D. (2005). "Gesture-based affective computing on motion capture data," In *Proceedings of the International Conference on Affective Computing and Intelligent Interaction*.

Bancroft, W. J. (1995). "Research in Nonverbal Communication and Its Relationship to Pedagogy and Suggestopedia," *ERIC*.

Barkley, R. A. (1998). *Attention Deficit Hyperactivity Disorder: A Handbook for Diagnosis and Treatment*, 2nd ed. New York: Guilford Press.

Bartlett, M. S., Littlewort, G., Fasel, I., and Movellan, J. R. (2003). "Real Time Face Detection and Facial Expression Recognition: Development and Applications to Human Computer Interaction," *Conference on Computer Vision and Pattern Recognition Workshop*.

Barnhart, C. L. and Barnhart, R. K. (Eds.). (1984). *The World Book Dictionary: Vocabulary inventories*. Chicago: Field Enterprises Educational Corporation.

Bee, H. and Boyd, D. (2004). *The Developing Child*, 10th ed. Boston: Pearson.

Ben Shalom, D., Mostofsky, S. H., Hazlett, R. L., Goldberg, M. C., Landa, R. J., Faran, Y., McLeod, D. R., and Hoehn-Saric, R. (2006). "Normal physiological emotions but differences in expression of conscious feelings in children with high-functioning autism," *Journal of Autism and Developmental Disorders*, vol. 36, no. 3, pp. 395-400.

Bernard-Opitz, V., Sriram, N., and Nakhoda-Sapuan, S. (2001). "Enhancing social problem solving in children with autism and normal children through computer-assisted instruction," *Journal of Autism and Developmental Disorders*, vol. 31, no. 4, pp. 377-384.

Berument, S. K., Rutter, M., Lord, C., Pickles, A., and Bailey, A. (1999). "Autism Screening Questionnaire: Diagnostic Validity," *British Journal of Psychiatry*, vol. 175, pp. 444-451.

Bethel, C., Salomon, K., Murphy, R., and Burke, J. (2007). "Survey of psychophysiology measurements applied to human-robot interaction," in *Proceedings of the IEEE International Symposium on Robot and Human Interactive Communication*, Jeju, Korea.

Bradley, M. M. (1994). "Emotional memory: a dimensional analysis," in *Emotions: Essays on emotion theory*, S. Van Goozen, N. E. Van de Poll, and I. A. Sergeant, Eds. Hillsdale, NJ: Erlbaum, pp. 97-134.

Bradley, M. M. (2000). "Emotion and motivation," in *Handbook of Psychophysiology*, J. T. Cacioppo, L. G. Tassinary, and G. Berntson, Eds. New York: Cambridge University Press, pp. 602-642.

Brown, R. M., Hall, L. R., Holtzer, R., Brown, S. L., and Brown, N. L. (1997). "Gender and video game performance," *Sex Roles*, vol. 36, pp. 793-812.

Burges, C. J. C. (1998). "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, pp. 121-167.

Cacioppo, J. T., Berntson, G. G., Larsen, J. T., Poehlmann, K. M., and Ito, T. A. (2000). "The psychophysiology of emotion," in *Handbook of Emotions*, M. Lewis and J. M. Haviland-Jones, Eds. New York: Guilford, pp. 173-191.

Cantwell, D. P., Lewinsohn, P. M., Rohde, P., and Seeley, J. R. (1997). "Correspondence between adolescent report and parent report of psychiatric diagnostic data," *Journal of the American Academy of Child and Adolescent Psychiatry*, vol. 36, pp. 610-619.

CDC. (2007). "Prevalence of autism spectrum disorders – Autism and developmental disabilities monitoring network, 14 sites, United States, 2002," *Morbidity and Mortality Weekly Report (MMWR) Surveillance Summaries*, vol. 56, pp. 12-28.

Chen, S. H. and Bernard-Opitz, V. (1993). "Comparison of personal and computer-assisted instruction for children with autism," *Mental Retardation*, vol. 31, no. 6, pp. 368-376.

Cobb, S. V. G., Nichols, S., Ramsey, A., and Wilson, J. R. (1999). "Virtual reality-induced symptoms and effects," *Presence*, vol. 8, no. 2, pp. 169-186.

Colburn, A., Drucker, S., and Cohen, M. (2000). "The role of eye-gaze in avatar-mediated conversational interfaces," in *SIGGRAPH Sketches and Applications*, New Orleans, Louisiana, USA.

Conati, C., Chabbal, R., and Maclaren, H. (2003). "A study on using biometric sensors for detecting user emotions in educational games," *Workshop on Assessing and Adapting to User Attitude and Affects: Why, When and How*, Pittsburgh, Pennsylvania, USA.

Constantino, J. N. (2002). *The Social Responsiveness Scale*. Los Angeles, California: Western Psychological Services.

Constantino, J. N., Davis, S. A., Todd, R. D., Schindler, M. K., Gross, M. M., Brophy, S. L., Metzger, L. M., Shoushtari, C. S., Splinter, R., and Reich, W. (2003). "Validation of a brief Quantitative Measure of Autistic Traits: Comparison of the Social Responsiveness Scale with the Autism Diagnostic Interview-Revised," *Journal of Autism and Developmental Disorders*, vol. 33, no. 4, pp. 427-433.

Cowie, R., Douglas-Cowie, E., Tsapatsoulis, N., Votsis, G., Kollias, S., Fellenz, W., and Taylor, J. G. (2001). "Emotion recognition in human-computer interaction," *IEEE Signal Processing Magazine*, vol. 18, no. 1, pp. 32-80.

Cromby, J. J., Standen, P. J., and Brown, D. J. (1996). "The potentials of virtual environments in the education and training of people with learning disabilities," *Journal Intellectual Disability Research*, vol. 40, pp. 489-501.

Dautenhahn, K. and Werry, I. (2004). "Towards interactive robots in autism therapy: Background, motivation and challenges," *Pragmatics and Cognition*, vol. 12, pp. 1-35.

Dautenhahn, K., Werry, I., Salter, T., and te Boekhorst, R. (2003). "Towards adaptive autonomous robots in autism therapy: Varieties of interactions," *IEEE International Symposium on Computational Intelligence in Robotics and Automation*, pp. 577-582, Kobe, Japan.

Dolch, E. W., (1948). *Problems in Reading*. The Garrard Press.

Dunn, L. M. and Dunn, L. M. (1997). PPVT-III: Peabody Picture Vocabulary Test, 3rd ed. Circle Pines, Minnesota: American Guidance Service.

Eisenberg, N., Fabes, R., Murphy, B., Maszk, P., Smith, M., and Karbon, M. (1995). "The role of emotionality and regulation in children's social functioning: a longitudinal study," *Child Development*, vol. 66, pp. 1360-1384.

Ekman, P. (1992). "An argument for basic emotions," *Cognition & Emotion*, vol. 6, pp. 169-200.

Ekman, P. (1993). "Facial expression and emotion," *American Psychologist*, vol. 48, pp. 384-392.

Ehlers, S. and Gillberg, C. (1993). "The Epidemiology of Asperger syndrome: A total population study," *Journal of Child Psychology and Psychiatry*, vol. 34, no. 8, pp. 1327-1350.

Ernsperger, L. (2003). *Keys to Success for Teaching Students with Autism*. Arlington, Texas: Future Horizons.

Farroni, T., Csibra, G., Simion, F., and Johnson, M. H. (2002). "Eye contact detection in humans from birth," in *Proceedings of the National Academy of Sciences*, vol. 99 no. 14.

Fong, T., Nourbakhsh, I., and Dautenhahn, K. (2003). "A survey of socially interactive robots," *Robotics and Autonomous Systems*, vol. 42, no. 3/4, pp. 143-166.

Frijda, N. H. (1986). *The emotions*. New York: Cambridge University.

Garau, M., Slater, M., Bee, S., and Sasse, M. A. (2001). "The impact of eye gaze on communication using humanoid avatars," in *Proceedings of the SIGCHI conference on human factors in computing systems*, pp. 309-316.

Gaylord-Ross, R. J., Haring, T. G., Breen, C., Pitts-Conway, V., (1984). "The training and generalization of social interaction skills with autistic youth," *Journal of Applied Behavior Analysis*, vol. 17, pp. 229–247.

Gilleade, K., Dix, A., and Allanson, J. (2005). "Affective videogames and modes of affective gaming: Assist me, challenge me, emote me," in *Proceedings of the Digital Games Research Association Conference*.

Gillott, A., Furniss, F., and Walter, A. (2001). "Anxiety in high-functioning children with autism," *Autism*, vol. 5, no. 3, pp. 277-286.

Green, D., Baird, G., Barnett, A. L., Henderson, L., Huber, J., Henderson, S. E. (2002). "The severity and nature of motor impairment in Asperger's syndrome: a comparison with specific developmental disorder of motor function," *Journal of Child Psychology and Psychiatry and Allied Disciplines*, vol. 43, no. 5, pp. 655-668.

Groden, J., Goodwin, M. S., Baron, M. G., Groden, G., Velicer, W. F., Lipsitt, L. P., Hofmann, S. G., and Plummer, B. (2005). "Assessing cardiovascular responses to stressors in individuals with autism spectrum disorders," *Focus on Autism and Other Developmental Disabilities*, vol. 20, no. 4, pp. 244-252.

Hall, E. T. (1955). "The Anthropology of Manners," *Scientific American*, vol. 192, pp. 84-90.

Hall, E. T. (1966). *The Hidden Dimension*. New York: Doubleday & Company, Inc.

Hill, E., Berthoz, S., and Frith, U. (2004). "Brief report: cognitive processing of own emotions in individuals with autistic spectrum disorder and in their relatives," *Journal of Autism and Developmental Disabilities*, vol. 34, no. 2.

IRB (Institutional Review Board). (February 9, 2004). *IRB Policy IV.B.1: Procedure for Documentation of Informed Consent for Human Subjects Research*. Nashville, Tennessee: Vanderbilt University Medical Center, /http://mcapps01.mc.vanderbilt.edu/ IRB/policy&procedures.nsf.

Jacobson, J. W., Mulick, J. A., and Green, G. (1998). "Cost-benefit estimates for early intensive behavioral intervention for young children with autism – General model and single state case," *Behavioral Interventions*, vol. 13, no. 4.

Jenkins, R., Beaver, J. D., and Calder, A. J. (2006). "I thought you were looking at me: Direction-specific aftereffects in gaze perception," *Psychological Science*, vol. 17, pp. 506-513.

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., and Moore, K. S. (2008). "The mind and brain of short-term memory," *Annual Review of Psychology*, vol. 59, pp. 193-224.

Kapoor, A., Mota, S., and Picard, R. W. (2001). "Towards a learning companion that recognizes affect," in *Proceedings of Emotional and Intelligent II: The Tangled Knot of Social Cognition AAAI Fall Symposium*.

Kataoka, H., Kano, H., Yoshida, H., Saijo, A., Yasuda, M., and Osumi, M. (1998). "Development of a skin temperature measuring system for non-contact stress evaluation," in *Proceedings of the IEEE Annual Conference on Engineering Medicine Biology Society*, pp. 940-943.

Keppel, G. (1991). *Design and Analysis: A Researcher's Handbook*, 3rd ed. Prentice Hall.

Kerr, S. and Durkin, K. (2004). "Understanding of thought bubbles as mental representations in children with autism: Implications for theory of mind," *Journal of Autism and Developmental Disorders*, vol. 34, no. 6.

Kleinsmith, A., Ravindra De Silva, P. and Bianchi-Berthouze, N. (2005). "Recognizing emotion from postures: cross-cultural differences in user modeling," In *Proceedings of User Modeling*, pp. 50-59.

Kozima, H., Michalowski, M. P., and Nakagawa, C. (2009). "Keepon: A playful robot for research, therapy, and entertainment," *International Journal of Social Robotics*, vol. 1, no. 1, pp. 3-18.

Kramer, A., Sirevaag, E., and Braune, R. (1987). "A Psychophysiological Assessment of Operator Workload during Simulated Flight Missions," *Human Factors*, vol. 29, no. 2, pp. 145-160.

Kulic, D. and Croft, E. (2007). "Physiological and subjective responses to articulated robot motion," *Robotica*, vol. 25, pp. 13-27.

Lang, P. J., Bradley M. M., and Cuthbert B. N. (1999). *International affective pictures system (IAPS): Technical manual and affective ratings*. Gainesville, FL: University of Florida, Center for Research in Psychophysiology.

Lacey, J. I. and Lacey, B. C. (1958). "Verification and extension of the principle of autonomic response-stereotypy," *American Journal of Psychology*, vol. 71, no. 1, pp. 50-73.

Lee, C. M. and Narayanan, S. S. (2005). "Toward detecting emotions in spoken dialogs," *IEEE Transactions on Speech and Audio Processing*, vol., 13, no. 2, pp. 293-303.

Lord, C., Pickles, A., McLennan, J., Rutter, M., Bregman, J., Folstein, S., Fombonne, E., Leboyer, M., and Minshew, N. (1997). "Diagnosing autism: analyses of data from the Autism diagnostic interview," *Journal of Autism and Developmental Disorders*, vol. 27, pp. 501–517.

Lord, C., Risi, S., Lambrecht, L., Cook, E. H., Leventhal, B. L., DiLavore, P. C., Pickles, A., and Rutter, M. (2000). "The Autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism," *Journal of Autism and Developmental Disorders*, vol. 30, no. 3, pp. 205-223.

Mandryk, R. L. and Atkins, M. S. (2007). "A fuzzy physiological approach for continuously modeling emotion during interaction with play technologies," *International Journal of Human-Computer Studies*, vol. 65, no. 4, pp. 329-347.

Mandryk, R. L., Inkpen, K. M., and Calvert, T. W. (2006). "Using physiological techniques to measure user experience with entertainment technologies," *International Journal of Human-Computer Studies*, vol. 25, no. 2, pp. 141–158.

Meehan, M., Razzaque, S., Insko, B., Whitton, M., and Brooks, Jr., F. P. (2005). "Review of four studies on the use of physiological reaction as a measure of presence in stressful virtual environments," *Applied Psychophysiology and Biofeedback*, vol. 30, no. 3, pp. 239-258.

Michaud, F. and Theberge-Turmel, C. (2002). "Mobile robotic toys and autism," in *Socially Intelligent Agents: Creating Relationships With Computers and Robots*, K. Dautenhahn, A. H. Bond, L. Canamero, and B. Edmonds, Eds. Norwell, MA: Kluwer, pp. 125-132.

Mitchell, P., Parsons, S., and Leonard, A. (2007). "Using virtual environments for teaching social understanding to adolescents with autistic spectrum disorders," *Journal of Autism and Developmental Disorders*, vol. 37, pp. 589-600.

Moore, D. J., McGrath, P., and Thorpe, J. (2000). "Computer aided learning for people with autism - A framework for research and development," *Innovations in Education and*

Training International, vol. 37, no. 3, pp. 218-228.

Myers, S. M., Johnson, C. P., American Academy of Pediatrics, and Council on Children with Disabilities. (2007). "Management of children with autism spectrum disorders," *Pediatrics*, vol. 120, no. 5, pp. 1162-1182.

Nasoz, F., Alvarez, K., Lisetti, C., and Finkelstein, N. (2004). "Emotion recognition from physiological signals using wireless sensors for presence technologies," *International Journal of Cognition, Technology, and Work – Special Issue on Presence*, vol. 6, no. 1, pp. 4-14.

Neumann, S. A. and Waldstein, S. R. (2001). "Similar patterns of cardiovascular response during emotion activation as a function of affective valence and arousal and gender," Journal of Psychosomatic Research, vol. 50, pp. 245-253.

NRC (National Research Council). (2001). *Educating Children with Autism*. Washington, DC: National Academy Press.

Pantic M. and Rothkrantz, L. J. M. (2003). "Toward an affect-sensitive multimodal human–computer interaction," *Proceedings of the IEEE*, vol. 91, no. 9, pp. 1370-1390.

Papert, S. (1993). *Mindstorms: Children, Computers, and Powerful Ideas*, 2nd ed. New York: Basic Books.

Papillo, J. F. and Shapiro, D. (1990). "The cardiovascular system," in *Principles of Psychophysiology: Physical, Social, and Inferential Elements*, J. T. Cacioppo and L. G. Tassinary, Eds. New York: Cambridge University Press.

Parsons, S. and Mitchell, P. (2002). "The potential of virtual reality in social skills training for people with autistic spectrum disorders," *Journal of Intellectual Disability Research*, vol. 46, pp. 430-443.

Parsons, S., Mitchell, P., and Leonard, A. (2004). "The use and understanding of virtual environments by adolescents with autistic spectrum disorders," *Journal of Autism and Developmental Disorders*, vol. 34, no. 4, pp. 449-466.

Parsons, S., Mitchell, P., and Leonard, A. (2005). "Do adolescents with autistic spectrum disorders adhere to social conventions in virtual environments?," *Autism*, vol. 9, pp. 95-117.

Pecchinenda, A. and Smith, C. A. (1996). "The affective significance of skin conductance activity during a difficult problem-solving task," *Cognition & Emotion*, vol. 10, no. 5, pp. 481-504.

Picard, R. W. (1997). *Affective Computing*. Cambridge, MA: MIT Press.

Picard, R. W., Vyzas, E., and Healey, J. (2001). "Toward machine emotional intelligence: analysis of affective physiological state," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 10, pp. 1175-1191.

Pioggia, G., Igliozzi, R., Ferro, M., Ahluwalia, A., Muratori, F., and De Rossi, D. (2005). "An android for enhancing social skills and emotion recognition in people with autism," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 13, no. 4, pp. 507-515.

Prendinger, H., Mori, J., and Ishizuka, M., (2005). "Using human physiology to evaluate subtle expressivity of a virtual quizmaster in a mathematical game," *International Journal of Human-Computer Studies*, vol. 62, no. 2, pp. 231-245.

Rani, P. (2005). *Psychophysiology-based Affective Communication for Implicit Human-robot Interaction*. Doctoral Dissertation, Vanderbilt University.

Rani, P., Liu, C., and Sarkar, N. (2006). "Affective feedback in closed loop human-robot interaction," in *Proceeding of the 1st ACM SIGCHI/SIGART Conference on Human-Robot interaction,* Salt Lake City, Utah.

Rani, P., Liu, C., Sarkar, N., and Vanman, E. (2006). "An empirical study of machine learning techniques for affect recognition in human-robot interaction," *Pattern Analysis and Applications*, vol. 9, no. 1, pp. 58-69.

Rani, P., Sarkar, N., Smith, C. A., and Kirby, L. D. (2004). "Anxiety detecting robotic system – towards implicit human-robot collaboration," *Robotica*, vol. 22, pp. 85-95.

Robins, B., Dickerson, P., and Dautenhahn, K. (2005). "Robots as embodied beings – Interactionally sensitive body movements in interactions among autistic children and a robot," in *Proceedings of the 14th IEEE International Workshop on Robot and Human Interactive Communication*, Nashville, Tennessee.

Robins, B., Dickerson, P., Stribling, P., and Dautenhahn, K. (2004). "Robot-mediated joint attention in children with autism: A case study in robot–human interaction," *Interaction Studies*, vol. 5, no. 2, pp. 161-198.

Rogers, S. J. (1998). "Empirically supported comprehensive treatments for young children with autism," *Journal of Clinical Child Psychology*, vol. 27, pp. 168-179.

Rogers, S. J. (2000). "Interventions that facilitate socialization in children with autism," *Journal of Autism and Developmental Disorders*, vol. 30, no. 5, pp. 399-409.

Ruble, L. A. and Robson, D. M. (2006). "Individual and environmental determinants of engagement in autism," *Journal of Autism and Developmental Disorders*, vol. 37, no. 8, pp. 1457-1468.

Rutter, M. (2006). "Autism: its recognition, early diagnosis, and service implications," *Journal of Developmental & Behavioral Pediatrics*, vol. 27, no. 2, Supplement 2, pp. S54-S58.

Rutter, M., Bailey, A., Berument, S., Lord, C., and Pickles, A. (2003a). *Social Communication Questionnaire*. Los Angeles, California: Western Psychological Services.

Rutter, M., Le Couteur, A., and Lord, C. (2003b). *Autism diagnostic interview revised WPS edition manual*. Los Angeles, California: Western Psychological Services.

Scassellati, B. (2005). "Quantitative metrics of social response for autism diagnosis," in *Proceedings of the IEEE International Workshop on Robot and Human Interactive Communication*, pp. 585-590, Nashville, Tennessee.

Schneiderman, M. H. and Ewens, W. L. (1971). "The Cognitive Effects of Spatial Invasion," *The Pacific Sociological Review*, vol. 14, no. 4, pp. 469-486.

Schultz, R. T. (2005). "Developmental deficits in social perception in autism: the role of the amygdala and fusiform face area," *International Journal of Developmental Neuroscience*, vol. 23, pp. 125-141.

Seip, J. (1996). *Teaching the Autistic and Developmentally Delayed: A Guide for Staff Training and Development*. Delta, BC: Author.

Sharpe, D. L. and Baker, D. L. (2007). "Financial issues associated with having a child with autism," *Journal of Family and Economic Issues*, vol. 28, pp. 247-264.

Sherer, M. R. and Schreibman, L. (2005). "Individual behavioral profiles and predictors of treatment effectiveness for children with autism," *Journal of Consulting and Clinical Psychology*, vol. 73, no. 3, pp. 525–538.

Sherman, W. R. and Craig, A. B. (2003). *Understanding virtual reality: interface, application, and design*. Boston: Morgan Kaufmann Publishers.

Siegel, S. and Castellan, J. N. J. (1988). *Nonparametric Statistics for the Behavioral Sciences*. New York: McGraw-Hill.

Silva, P. R. D., Osano, M., Marasinghe, A., and Madurapperuma, A. P. (2006). "A computational model for recognizing emotion with intensity for machine vision applications," *IEICE Transactions on Information and Systems*, vol. E89-D, no. 7, pp. 2171-2179.

Smith, C. A. (1989). "Dimensions of appraisal and physiological response in emotion," *Journal of Personality and Social Psychology*, vol. 56, no. 3, pp. 339-353.

Sommer, R. (1962). "The Distance for Comfortable Conversations: A Further Study," *Sociometry*, vol. 25, no. 1, pp. 111-116.

Standen, P. J. and Brown, D. J. (2005). "Virtual reality in the rehabilitation of people with intellectual disabilities: review," *CyberPsychology and Behavior*, vol. 8, no. 3, pp. 272-282; discussion pp. 283-288.

Strickland, D. (1997). "Virtual reality for the treatment of autism," in *Virtual reality in neuropsycho-physiology*, G. Riva, Ed., pp. 81-86, Amsterdam: IOS Press.

Strickland, D., Marcus, L. M., Mesibov, G. B., and Hogan, K. (1996). "Brief report: two case studies using virtual reality as a learning tool for autistic children," *Journal of Autism and Developmental Disorders*, vol. 26, no. 6.

Stokes, S. (2000). "Assistive technology for children with autism," published under a *CESA 7 contract funded by the Wisconsin Department of Public Instruction*.

Swettenham, J. (1996). "Can children with autism be taught to understand false belief using computers?," *Journal of Child Psychology and Psychiatry*, vol. 37, no. 2, pp. 157-165.

Tarkan, L. (2002). "Autism therapy is called effective, but rare," *New York Times*.

Tartaro, A. and Cassell, J. (2007). "Using virtual peer technology as an intervention for children with autism," in *Towards Universal Usability: Designing Computer Interfaces for Diverse User Populations*, J. Lazar, Ed., Chichester, UK: John Wiley and Sons.

Toichi, M. and Kamio, Y. (2003). "Paradoxical autonomic response to mental tasks in autism," *Journal of Autism and Developmental Disorders*, vol. 33, no. 4, pp. 417-426.

Trepagnier, C. Y., Sebrechts, M. M., Finkelmeyer, A., Stewart, W., Woodford, J., and Coleman, M. (2006). "Simulating social interaction to address deficits of autistic spectrum disorder in children," *CyberPsychology and Behavior*, vol. 9, no. 2, pp. 213-217.

Vansteelandt, K., Van Mechelen, I., and Nezlek, J. (2005). "The co-occurrence of emotions in daily life: a multilevel approach," *Journal of Research in Personality*, vol. 39, no. 3, pp. 325-335.

Vapnik, V. N. (1998). *Statistical Learning Theory*. New York: Wiley-Interscience.

Vicente, K., Thornton, D., and Moray, N. (1987). "Spectral-Analysis of Sinus Arrhythmia - a Measure of Mental Effort," *Human Factors*, vol. 29, no. 2, pp. 171-182.

Watkins, C. J. C. H. and Dayan, P. (1992). "Q-learning," *Machine Learning*, vol. 8, pp. 279-292.

Werry, I., Dautenhahn, K., and Harwin, W. (2001). "Investigating a robot as a therapy partner for children with autism," in *Proceedings of the 6th European conference for the advancement of assistive technology*, Ljubljana, Slovenia.

Wieder, S. and Greenspan, S. (2005). "Can children with autism master the core deficits and become empathetic, creative, and reflective?," *The Journal of Developmental and Learning Disorders*, vol. 9, pp. 1-29.

Wiering, M. A. (2005). "*QV* (λ)-learning: A new on-policy reinforcement learning algorithm," in *Proceedings of the 7th European Workshop on Reinforcement Learning*, pp. 17-18, Napoli, Italy.

Wijesiriwardana, R., Mitcham, K., and Dias, T. (2004). "Fibre-meshed transducers based real time wearable physiological information monitoring system," In *Proceedings of the International Symposium on Wearable Computers*, Washington, DC.

Zhai J., and Barreto A. (2006). "Stress Detection in Computer Users through Noninvasive Monitoring of Physiological Signals," *Biomedical Science Instrumentation*, vol. 42, pp. 495-500.

Zhai, J., Barreto, A., Chin, C., and Li, C. (2005). "User Stress Detection in Human-Computer Interactions," *Biomedical Sciences Instrumentation*, vol. 41, pp. 277-286.