

# Design of a Gaze-Sensitive Virtual Social Interactive System for Children With Autism

Uttama Lahiri, *Student Member, IEEE*, Zachary Warren, and Nilanjan Sarkar, *Senior Member, IEEE*

**Abstract**—Impairments in social communication skills are thought to be core deficits in children with autism spectrum disorder (ASD). In recent years, several assistive technologies, particularly Virtual Reality (VR), have been investigated to promote social interactions in this population. It is well known that children with ASD demonstrate atypical viewing patterns during social interactions and thus monitoring eye-gaze can be valuable to design intervention strategies. While several studies have used eye-tracking technology to monitor eye-gaze for offline analysis, there exists no real-time system that can monitor eye-gaze dynamically and provide individualized feedback. Given the promise of VR-based social interaction and the usefulness of monitoring eye-gaze in real-time, a novel VR-based dynamic eye-tracking system is developed in this work. This system, called Virtual Interactive system with Gaze-sensitive Adaptive Response Technology (VIGART), is capable of delivering individualized feedback based on a child's dynamic gaze patterns during VR-based interaction. Results from a usability study with six adolescents with ASD are presented that examines the acceptability and usefulness of VIGART. The results in terms of improvement in behavioral viewing and changes in relevant eye physiological indexes of participants while interacting with VIGART indicate the potential of this novel technology.

**Index Terms**—Autism spectrum disorder (ASD), blink rate, eye-tracking, fixation duration, pupil diameter, virtual reality (VR).

## I. INTRODUCTION

Autism is a neurodevelopmental disorder (prevalence rates in the United States recorded as approximately 1 in 110 for the broad autism spectrum [1]) characterized by core deficits in social interaction and communication accompanied with infrequent engagement in social interactions [2], and impaired understanding of mental states of others [3]. To understand the social communication vulnerabilities of individuals with Autism Spectrum Disorder (ASD), research has examined how they process salient social cues, specifically from faces [4]. The ability to derive socially relevant information from faces is a

fundamental skill for facilitating reciprocal social interactions [5]. However, individuals with ASD exhibit poor eye contact during social communication in which they tend to fixate less towards human faces and more towards other objects within visual stimulus [6].

There is growing consensus that intensive behavioral intervention programs can significantly improve long term outcomes for these individuals [7]. Faced with limited trained professional resources, researchers are employing technology e.g., computer technology [8], virtual reality (VR) environments [9], [10], and robotic systems [11] to develop more accessible individualized intervention [12].

In this work, we develop a novel VR-based system for social interaction to be used with adolescents with ASD. In our usability study, we present VR-based tasks on a computer screen as VR is often effectively experienced using standard desktop computer input devices [13] for ASD intervention. We chose VR because of its malleability, controllability, replicability, modifiable sensory stimulation, and an ability to pragmatically individualize intervention approaches and reinforcement strategies [14]. VR can illustrate scenarios which can be changed to accommodate various situations that may not be feasible in a given therapeutic setting with space limitations and resource deficits [9].

Despite potential advantages, current VR environments as applied to assistive intervention for children with ASD are designed in an open-loop fashion [9], [10]. These VR systems may be able to chain learning via aspects of performance; however, they are not capable of a high degree of individualization. A recent work using VR has demonstrated the feasibility of linking the gaze behavior of a virtual character with a human observer's gaze position during joint-attention tasks [15]. Specifically, these systems though may automatically detect and respond based on one's viewing pattern, cannot objectively identify and predict social engagement, understand viewing patterns, and psychophysiological effect of the specific child based on attentive indexes. Thus, the development of interactive systems based on dynamic gaze patterns of these children to address some of their core deficits in communication and social domains is still at its infancy.

In the present work, we developed a novel Virtual Interactive system with Gaze-sensitive Adaptive Response Technology (VIGART) using VR-based social situation as a platform for delivering individualized feedback based on one's dynamic gaze patterns. Additionally, we report the results from a usability study with VIGART to investigate the impact of this system on the engagement level, behavioral viewing, and eye physiological indexes of six adolescents with ASD.

Manuscript received July 21, 2010; revised January 26, 2011; accepted April 02, 2011. Date of publication May 23, 2011; date of current version August 10, 2011. This work was supported in part by the National Science Foundation under Grant 0967170 and in part by the National Institutes of Health under Grant 1R01MH091102-01A1.

U. Lahiri and N. Sarkar are with the Mechanical Engineering Department, Vanderbilt University, Nashville, TN 37212 USA (e-mail: uttama.lahiri@vanderbilt.edu; nilanjan.sarkar@vanderbilt.edu).

Z. Warren is with the Psychiatry Department, Vanderbilt University, Nashville, TN 37212 USA (e-mail: zachary.e.warren@vanderbilt.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TNSRE.2011.2153874

This paper is organized as follows. In Section II, we present the objectives and scope of our work. Section III discusses the design specifications of the VIGART along with the design rationale of the gaze-based closed-loop interaction providing individualized feedback. The design details of the usability study are presented in Section IV. The results are presented in Section V. Section VI summarizes the contributions of this work and outlines the scope of future research.

## II. OBJECTIVES AND SCOPE

The construct of behavioral engagement of children with ASD is the ground basis for many evidence based intervention paradigms focused on bolstering core skill areas [16]. Given the importance of engagement, viewing patterns [5], [17], and psychophysiology [18], developing a gaze-sensitive VR-based system for ASD intervention that can provide individualized feedback based on a child's real-time gaze pattern and can potentially elicit changes in engagement level, behavioral viewing, and eye physiological indexes of the child can be critical. This can be a step towards achieving realistic social interaction to challenge, and scaffold skill development in particular areas of vulnerability for these children. In the current study, our objectives were two-fold: 1) to design a novel VIGART capable of delivering individualized feedback based on a child's dynamic gaze patterns; and 2) to conduct a usability study to examine the acceptability of VIGART and investigate how adolescents with ASD interact with such a system and respond in terms of behavioral viewing and eye physiological indexes while participating in the virtual social tasks. In addition, we also investigated whether VIGART can measure behavioral viewing indexes and eye physiological indexes dynamically that have the potential to be used as objective measures of one's engagement as well as emotion recognition capability; though our usability study was not designed to improve the participants' ability to correctly recognize an emotional expression. The behavioral viewing indexes that we studied were the fixation duration and fixation frequency because of their role in predicting attention on a specific region of visual stimulus [19]–[21]. Similarly, eye physiological indexes such as the blink rate and pupil diameter were also measured because of their sensitivity to emotion recognition [22], [23].

## III. DESIGN OF VIGART

The dynamic closed-loop interaction provided by VIGART has three main subsystems: 1) a VR-platform that can present social tasks; 2) a real-time eye-gaze monitoring mechanism; and 3) an integration module that establishes communication between the first two subsystems.

### A. VR-Based Task Presentation

VR-based tasks are created using Vizard design package from Worldviz.<sup>1</sup> This software comes with a limited number of avatars, virtual objects, and scenes that can be used to create a story in VR. However, a number of enhancements were made on the VR-platform. We develop more extensive social situations with appropriate contexts, and avatars whose age and



Fig. 1. Screenshots of avatars demonstrating neutral (top), happy (middle), and angry (bottom) facial expression.

appearance resemble those of the participants' peers without trying to achieve exact similarities.

Thus new avatar heads are created from 2D photographs of teenagers, which are then converted to 3D heads by "3DmeNow" software for compatibility with Vizard. Facial expressions (e.g., "neutral," "happy," and "angry") (Fig. 1) are morphed by "PeopleMaker" software. The avatar's eyes are made to blink randomly with an interval between 1 and 2 s to render automatic animation of a virtual face similar to the work of Itti *et al.* [24]. Different cultures have varying rules for social distance. For example, the overcrowded nature of Asian countries causes people to be accustomed with very close distances, whereas western culture considers very close distances as uncomfortable [25]. Though our system is capable of simulating variations in personal distance, in our presented work, the avatar stands at 3 ft from the origin of main scene view in VR to simulate the close phase of social distance suitable for western culture [25]. The first-person stories shared by avatars are adapted from dynamic indicators of basic early literacy skills [26] reading assessments and includes content thought to be related to potential topics of school presentations (e.g., reports on experiences, trips, etc.). Audio files are developed first by using the text-to-speech "NaturalReader" converter and then recorded using "Audacity" software. For the avatars to speak the content of the story, these audio files are lip-synched using a Vizard-based speak module. Additionally, the VR-based visual stimuli (e.g., avatar's face, objects of interest, etc.) is characterized by a set of regions of interest (ROIs). These ROIs are used by the dynamic eye-tracking algorithm we present in this paper to keep track of the eye-gaze of the participants as they interact with the VR-based tasks.

### B. Real-Time Gaze Monitoring Mechanism

The VIGART captures eye data of a participant interacting with an avatar using Eye-Tracker goggles [27]. This eye-tracker comes with some basic feature acquiring capabilities (e.g., acquiring raw normalized pupil diameter, pupil aspect ratio (PAR), etc.) for offline analysis.

In this study, we correlate the extracted eye-gaze features reflecting the behavioral viewing patterns of a participant with ASD with his/her engagement level because engagement, which is defined as "sustained attention to an activity or person," [7] is one of the key factors for these children to make substantial gains in communication and social domains [28]. In addition, we correlate the extracted features reflecting the behavioral viewing patterns and eye physiological indexes, with emotion recognition capability of the participants because emotion recognition is characterized as one of the core deficits indicating ASD [3] as well as its importance in social communication [21].

<sup>1</sup><http://www.worldviz.com/>

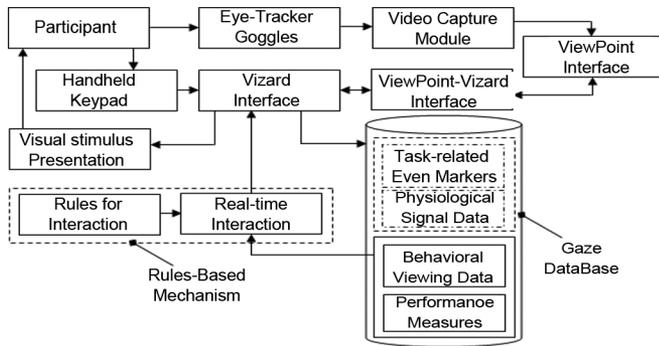


Fig. 2. Schematic of data acquisition and the control mechanism used.

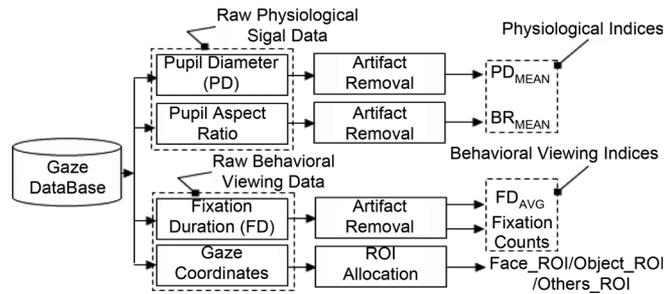


Fig. 3. Schematic of feature extraction.

1) *Data Acquisition*: The Eye-Tracker that we use comes with a Video Capture Module with refresh rates of 60 Hz (low precision) and 30 Hz (high precision) to acquire a participant's gaze data using a software called Viewpoint. We have designed Viewpoint-Vizard Handshake module (Fig. 2), and a new gaze database that captures the task-related event markers (e.g., trial start/stop, amount of one's viewing of different ROIs, etc.), raw physiological signal data (pupil diameter and pupil aspect ratio), behavioral viewing data (fixation duration and 2D gaze coordinates), and performance measures (e.g., a participant's responses to questions asked by this system) with a refresh rate of 30 Hz in a time-synchronized manner. Signal processing techniques such as windowing and thresholding are used to filter these data to eliminate noise and subsequently extract the relevant features.

2) *Feature Extraction*: The Gaze Database (Fig. 2) is processed to extract four features, which are mean pupil diameter ( $PD_{MEAN}$ ), mean blink rate ( $BR_{MEAN}$ ), sum of fixation counts (SFC), and average fixation duration ( $FD_{AVG}$ ), for each ROI from each segment of the signals monitored (Fig. 3).

*Computation of Mean Pupil Diameter ( $PD_{MEAN}$ )*: The raw normalized pupil diameter is recorded by the Viewpoint software in terms of normalized value (0-1) with respect to the EyeCamera window of the eye-tracker. Literature indicates the importance of measuring actual pupil diameter of typically developing (TD) [29] and ASD [30] participants. For example, Anderson *et al.* [30] report larger tonic actual pupil size in children with ASD than their TD counterparts. In the present work, in order to extract the actual pupil diameter at each instant, we use the recorded data on pupil aspect ratio (i.e., the ratio of the major and minor axes of the pupil image) defining the eye image (with 1 indicating a perfect circle) to compute optimal

pupil diameter. Artifact removal incorporates elimination of the discontinuities in the raw pupil diameter data due to blinking effects and other minor artifacts. Then using the actual Eye-Camera window dimensions [27], the true pupil diameter (in millimeters) is computed. Using the record of ROIs visited by the eye at each instant, the mean pupil diameter corresponding to each ROI is computed.

*Computation of Mean Blink Rate ( $BR_{MEAN}$ )*: The blink rate is determined using the pupil aspect ratio data recorded by the Viewpoint software as mentioned by Arrington [31]. In the present work, we computed the blink rate by considering the number of times the pupil aspect ratio value falls below the lower threshold of 0.5 for a 1 min window, which allowed us to detect blink rate with an accuracy of  $\pm 0.5\%$ . Subsequently, the mean blink rate corresponding to each ROI is computed.

*Computation of Fixation Counts and Average Fixation Duration ( $FD_{AVG}$ )*: The recorded data on fixation duration corresponding to each ROI is first filtered to remove the artifacts due to blinking and noise spikes are eliminated by thresholding. Literature indicates that fixations typically last between 200–600 ms [32]. In the present study, we compute the fixation duration by using a thresholding window of 200 ms as the lower limit to eliminate the blinking effects and 450 ms as the upper limit (i.e., up to 1.5 standard deviations from the lower threshold). These limits provided a reliable data range restricted by noise due to glare effects of cameras of the eye-tracker that we used. Subsequently, the sum of fixation counts (SFC) and average fixation duration are computed for each ROI.

*Detection of ROIs Viewed*: Prior to our study, we carried out a systematic validation of the overall system performance in terms of accuracy and repeatability of the eye-gaze data with lab members. In this validation study, participants were asked to fixate on each of 15 predefined target points appearing randomly on the visual stimulus each for 5 s. The viewing area was normalized on a 0–1 scale, where the left-hand bottom corner and the right-hand top corner of the screen were assigned the (0, 0) and (1, 1) coordinates, respectively. The average mean square errors for the  $x$  and  $y$  gaze coordinates for the 15 predefined points were found to be 0.0053 and 0.0077 (i.e., eight pixels along  $x$  and seven pixels along  $y$  for the entire presented visual stimulus), respectively.

Afterwards, we designed a computational algorithm, the real-time gaze-based feedback algorithm (RGFA) that determines whether the gaze coordinates correspond to our task-specific segmented regions of the visual stimulus presented to participants. Subsequently, RGFA assigns numeric tags (e.g., 1, 2, etc.) for each ROI. In the present study, we segmented the VR-based visual stimulus into three ROIs: avatar's face (Face\_ROI), a context-relevant object (Object\_ROI), and rest of the VR environment (Others\_ROI) (Fig. 4). Face\_ROI captures the forehead, eye brows, eyes and surrounding muscles, nose, cheeks, mouth and surrounding muscles. Object\_ROI captures a context-relevant object (e.g., for a story on outdoor games, the context-relevant object is a picture displaying collage of snapshots of narrated games).

The task-related event markers along with the ROI tags are then used by Real-time Gaze-based Feedback Algorithm

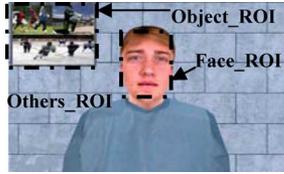


Fig. 4. Allocation of ROIs (Face\_ROI, Object\_ROI, and Others\_ROI).

to compute the physiological and behavioral indexes during viewing of different ROIs by a participant.

### C. Communication Between VR-Based Task Presentation and Real-Time Gaze Monitoring Modules

1) *Rationale Behind Gaze-Based Individualized Feedback Mechanism:* The VIGART is capable of providing a participant with individualized feedback based on the behavioral viewing patterns so as to capture his/her attention to a task. In dyadic communication, gaze information underlying one's expressive behavior plays a vital role in regulating conversation flow. Looking behavior is a function of the cultural upbringing of an individual [33]. Specifically, some cultures may give a negative connotation to looking directly at others' eyes, while others may consider this as appropriate. Our presented system has the flexibility of changing the gaze parameter to suit different cultural requirements. In our present work we use gaze definition suitable for western culture with a listener looking at the speaker 70% of the time during an interaction identified as "normal while listening" [34], [35].

In the present study, a participant can serve as an audience while the avatars narrate personal stories and display context-relevant facial expressions (Fig. 1) to capture the mood inherent in the content of the story. Subsequently, the participant's fixation duration for Face\_ROI viewing as a percentage of total fixation duration, extracted from the behavioral viewing data and the performance measure (the participant's response to question asked by VIGART) initiate a rule-based mechanism (Fig. 2) to trigger the VIGART to provide feedback (Table I) to the participant using the individualized real-time gaze-based feedback algorithm (Section III-C2). The VIGART can generate complex individualized feedback based on one's response to several questions, any performance measures defined for a given task, and his/her actual viewing pattern. However, for the usability study presented here, a simple rule-based system is designed to demonstrate the potential of VIGART. A feedback is generated based on whether the participant could recognize the avatar's emotion and how much he/she looked at the avatar's face (Table I).

2) *Overview of the Individualized Real-Time Gaze-Based Feedback Algorithm (RGFA):* The data flow diagram for the RGFA (Fig. 5) presents a brief overview of the logic used by VIGART. Real-time gaze coordinates of a participant interacting with an avatar are acquired using the Viewpoint software and converted to VR (Vizard) compatible format using a Viewpoint-Vizard handshake module (Fig. 2). A computer, where the VR-based tasks are presented, runs Viewpoint Software at the background and Vizard software at the foreground. The RGFA triggers a 33 ms timer to acquire the gaze coordinates.

TABLE I  
RATIONALE BEHIND ATTENTION-BASED REAL-TIME  
MOTIVATIONAL FEEDBACK

Response to Q1	$t \geq 70\%$	VIGART Response [Label]
Right	Yes	Your classmate really enjoyed having you in the audience. You have paid attention to her and made her feel comfortable. Keep it up! [S1]
Right	No	Your classmate did not know if you were interested in the presentation. If you pay more attention to her, she will feel more comfortable. [S2]
Wrong	Yes	Your classmate felt comfortable in having you in the audience. However, you may try to pay some more attention to her as she makes the presentation so that you can correctly understand how she is feeling. [S3]
Wrong	No	Your classmate would have felt more comfortable if you had paid more attention to her. If you pay more attention, then you will correctly know how she is feeling and make her feel comfortable. [S4]

Q1 : Question asked by the system; t : Duration of participant's looking towards the Face\_ROI of visual stimulus.

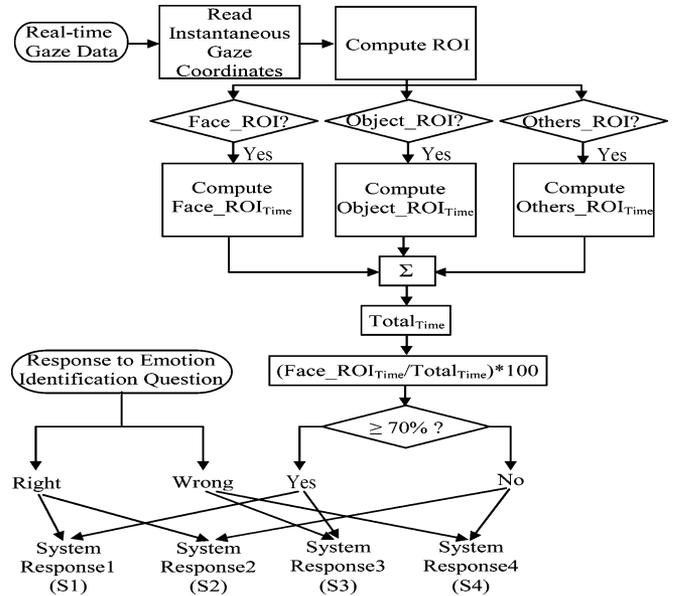


Fig. 5. Individualized RGFA.

Based on the participant's 2D gaze-coordinates, the RGFA then computes the specific ROI looked at by the participant. Times spent by the participant looking at different ROIs are stored in the respective buffers that are added up at each instant during participant-avatar interaction. This determines the participant's looking time towards the presented stimuli in terms of Face\_ROI<sub>Time</sub>, Object\_ROI<sub>Time</sub>, and Others\_ROI<sub>Time</sub>. Various measures can be computed from these data based on the requirement of a specific intervention protocol. For example, for the usability study presented here, these times are summed up to get the Total<sub>Time</sub> from which the percentage of time spent by a participant in looking at Face\_ROI is computed. In this specific usability study, four different responses (S1-S4) (Table I) are generated based on how much a participant looked at the face of the avatar and whether he/she correctly recognized the feeling of the avatar.

3) *Real-Time Interaction Module:* The gaze metrics as obtained from the Real-time Gaze-based Feedback Algorithm and any predefined interaction rules can be combined using the real-

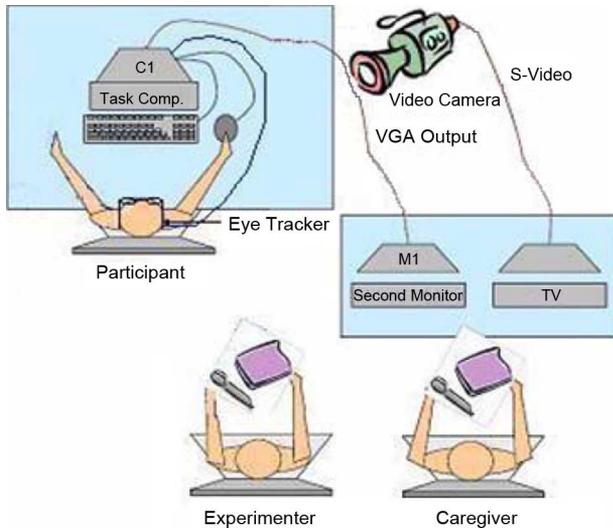


Fig. 6. Experimental setup.

time interaction module (Fig. 2) to provide individualized feedback in an attempt to improve the participant’s engagement to the social tasks. While not utilized in the usability study presented in this paper, the system has the capability of communicating with the VR-based platform using our Viewpoint-Vizard interface to modify the avatar’s states to create an interactive interaction. We have augmented the avatar’s capabilities in several ways. 1) The avatar can change his/her facial expressions during communication; e.g., the avatar may start with a neutral face and then based on the content of the story gradually change to happy/angry facial expression. The system triggers a token that activates the vizard-based “morph” module to change the percentage of smiling, or frowning, etc. to create such a change. 2) The avatar can walk towards, point to, and turn and look at an object of interest either as a part of interaction or to bring back the participant’s attention based on his/her current looking pattern. The system initiates the “walk to” or “turn and look” or “point” modules that we have developed and modified in vizard. 3) The avatar can also provide verbal feedback during the trial interrupting its original conversation/storytelling by initiating a separate conversation thread within vizard, if the protocol demands such a feedback. These capabilities can be accessed at every 33 ms if required. However, in the present usability study we have not fully exploited these extended capabilities of the system since they were not needed for the basic study. It is conceivable that for a sophisticated intervention study, there will be intervention rules as to when and how to change the states of the avatar and such capabilities would be beneficial.

#### IV. DESIGN OF THE USABILITY STUDY

##### A. Experimental Set-Up

The experiment was created using the VR environment described in Section III-A. The participant’s eye movements were tracked by the Eye-Tracker discussed in Section III-B. Stimuli were presented on a 17” computer monitor (C1) (Fig. 6). A chin rest with a height-adjustable telescopic shaft was designed

TABLE II  
INDIVIDUAL PARTICIPANT CHARACTERISTICS

Participant (Gender)	Age (years)	PPVT	SRS (cutoff=60)	SCQ (cutoff=13)	ADOS-G (cutoff=7)
ASD1 (m)	13.83	126	69	23	11
ASD2 (m)	15.5	110	73	13	7
ASD3 (f)	15.17	83	90	28	10
ASD4 (m)	16.5	97	63	17	9
ASD5 (m)	15.08	92	87	20	NA
ASD6 (m)	17.5	103	83	31	20
Mean (SD)	15.60(1.27)	102 (15)	78 (11)	22 (7)	11 (5)

NA: Not Available; m : Male; f : Female

and used to stabilize the participant’s head and maintain participant-monitor distance of 50 cm, considered as an appropriate distance in social gaze-based experiments [36]. Uniform room illumination was maintained throughout the experiment. The task computer (C1) presented the VR-based social tasks in the foreground and computed dynamic gaze information in the background using the eye-tracking data. Gaze data along with task-related event markers (e.g., starting and stopping of trials, participant feedback, etc.) were logged in a time-synchronized manner. The participant’s caregiver watched the participant from a video camera view that was routed to a television, hidden from the participant’s view. The signal from C1 was routed to a separate monitor (M1) for the caregiver to view how the task progressed. Based on these two observations, the observer (caregiver) rated the participant’s engagement level.

##### B. Participant Recruitment and Inclusion Criteria

Six adolescents (Male:  $n = 5$ , Female:  $n = 1$ ) with high-functioning ASD, ages 13–17 y ( $M = 15.60$  y,  $SD = 1.27$  y) participated in the usability study of the designed system. All participants were recruited through existing clinical research programs at Vanderbilt University and had established clinical diagnoses of ASD. Participants were also required to score  $\geq 80$  on Peabody Picture Vocabulary Test-3rd Edition (PPVT-III) [37] to ensure that language skills were adequate to participate in the current protocol. Data on core ASD related symptoms and functioning was obtained through parent report on Social Responsiveness Scale (SRS) [38] profile sheet and Social Communication Questionnaire (SCQ) [39] with all participants falling above the clinical thresholds (Table II). The SRS (cutoff T-score  $\geq 60$ ) generates a total score reflecting the severity of social deficits in the autism spectrum. From the SRS scores, we find that ASD1, ASD2, and ASD4 can be considered to be in mild to moderate ASD range, and ASD3, ASD5, and ASD6 were in the severe range of ASD (Table II). The SCQ (cutoff of 13) is a parent report questionnaire aimed at evaluating critical autism diagnostic domains of qualitative impairments. Autism Diagnostic Observation Schedule-Generic (ADOS-G) scores (cutoff of 7) [40] were available for five of the six participants from prior evaluations (Table II).

##### C. Tasks and Procedures

In the present usability study, we constructed five VR-based social interaction scenarios (trial1-trial5) in which the avatars (and henceforth referred to as the “virtual classmates” of the participants) narrated personal stories on diverse topics such as, outdoor sports, travel, favorite food, etc.

Each participant took part in an approximately 50 min laboratory visit. First, a brief adaptation session was carried out. In the first phase of adaptation, before the participants and their caregivers were asked to sign the assent and the consent forms respectively, the experimenter briefed the participant about the experiment, and showed the experimental setup. This phase took approximately 10 min. In the second phase of adaptation, the participant sat comfortably on a height-adjustable chair and was asked to wear the eye-tracker goggles. The chair was adjusted so that his/her eyes were collinear with the center of the task computer, C1 (Fig. 6). The experimenter told him/her that he/she could choose to withdraw anytime from the experiments for any reason, especially if he/she was not comfortable interacting with VIGART. The participant was then asked to rest for 3 min to acclimate him/her to the experimental setup. This second phase of adaptation took approximately another 10 min. Then the eye-tracker was calibrated. The average calibration time was approximately 15 s in which the participant sequentially fixated on a grid of 16 points displayed randomly on the task computer. The participants viewed an initial instruction screen followed by their virtual classmate narrating a personal story. Each storytelling trial was approximately 3 min long. The participants were asked to imagine that the avatars were his/her classmates at school giving presentations on several different topics. They were informed that after the presentations they would be required to answer a few questions about the presentation. They were also asked to try and make their classmate feel as comfortable as possible while listening to the presentation. However, it was not explicitly stated that in a presentation a speaker feels good when the audience pay attention to him/her (by looking towards the speaker). The idea here was to give feedback to the participants about their viewing patterns and thereby study how that affects the participants as the task proceeded. The experiment began with trial1 with the virtual classmate exhibiting a “neutral” facial expression [Fig. 1(a)] and narrating a personal story. This trial was followed by four other trials that were similar to the trial1 except that in these subsequent trials the virtual classmate displayed “happy” [Fig. 1(b)] or “angry” [Fig. 1(c)] facial expressions to capture the mood inherent in the content of the story. In the present study we used three female, two male avatar heads with one neutral, three happy, and one angry facial expressions. We randomized the trial2-trial5 among the participants and the avatars displayed context-relevant facial expressions. After each trial, the participant was asked an emotion-identification question (Q1) and a story-related question (Q2). The Q1 was about the virtual classmate’s emotion with three answer choices (A: Happy; B: Angry; C: Not Sure). The Q2 was about some basic facts as narrated in the story. It also had three answer choices. The participant responded with a keypad. Q2 was asked to encourage a participant to pay attention to the story content. Depending on the participant’s response to Q1 and how much attention he/she paid to the virtual classmate, as measured by the real-time computation of the percentage of time spent in looking at the classmate’s face, VIGART encouraged the participant to either pay more or keep the same attention towards the presenter (Table I). After each trial, the observer rated about what he/she thought about how engaged the participant was during the trial using a 1–9 scale (1: least engaged;

9: most engaged). Each participant was compensated with a gift card for completing a session.

## V. RESULTS AND DISCUSSION

Here we present the results of our usability study with six adolescents with ASD to examine the acceptability of VIGART by the target population, and investigate the potential of VIGART to elicit variation in participant’s engagement level. Subsequently, we also analyze whether VIGART can be used to define objective measures of engagement and emotion recognition capabilities so that such objective measures can be utilized in future intervention strategies using VIGART. All the results reported here are descriptive only. More definitive results supported by inferential statistics would require a larger sample size and comparison to a control group.

### A. System Acceptability

In the current usability study, we wanted to investigate whether VIGART was acceptable to adolescents with ASD. We tested our system with a small sample of six adolescents with ASD. In spite of being given the option of withdrawing from the experiment at any time during their interaction with VIGART, all the participants completed the session. An exit survey revealed that all participants liked interacting with VIGART, and had no problem in wearing the eye-tracker goggles, understanding the stories narrated by their virtual classmates, or responding to questions asked by VIGART. In fact five of them inquired whether there would be any future participation possibilities with VIGART. Thus it is reasonable to infer from this small usability study that the VIGART has a potential to be accepted by the target population.

### B. Potential of VIGART to Elicit Variation in Participants’ Engagement Level

We wanted to assess whether the VIGART can be used in virtual social interaction to create different engagement levels among the participants so that engagement manipulation using individualized feedback could be potentially feasible in the future as a part of intervention. In our usability study with the VIGART, the participants’ caregivers rated as to what they felt regarding the participants’ engagement level while interacting socially with their virtual classmates. As engagement can be represented in terms of “sustained attention to an activity or person” [7], we asked the caregivers to rate the participants’ engagement by observing their attention to the social task by using a 1–9 scale. Table III indicates that interaction with the VIGART was able to elicit a range of different engagement levels among the participants causing an increase of 64% from the lowest to the highest reported engagement levels, considering two out of five trials for each participant. Note that the lower rating on engagement given by the caregivers for five of the six participants (except ASD2) occurred at the initial trials and they received higher individual ratings as the session progressed. Further analysis revealed that ASD2 was incorrect in responding to story-related question in trial5 which may be due to his lower engagement. The caregiver of ASD2 reported that he liked the story in trial1 the most and the trial5 the least. Only ASD2 among all the

TABLE III  
COMPARATIVE ANALYSIS OF VARIATION IN THE BEHAVIORAL VIEWING INDEXES WITH REPORTED ENGAGEMENT LEVEL OF PARTICIPANTS

	Reported Engagement Rating (Trial No.)		SFC for Face <sub>ROI</sub> as % of total SFC (%)		SFC for NonFace <sub>ROI</sub> as % of total SFC (%)		FD <sub>AVG</sub> for Face <sub>ROI</sub> (ms)		FD <sub>AVG</sub> for NonFace <sub>ROI</sub> (ms)	
	LE	HE	LE	HE	LE	HE	LE	HE	LE	HE
ASD1	2 (1)	5 (4)	56	80.5	44	19.5	315	363	312	313
ASD2	6 (5)	7 (1)	80	88.2	20	11.8	319	339	254	297
ASD3	4 (1)	7 (4)	77.7	98.5	22.3	1.5	298	379	300	322
ASD4	6 (1)	9 (2)	89.1	98.8	10.9	1.2	343	354	342	297
ASD5	3 (3)	6 (4)	79.3	84.9	20.7	15.1	272	296	269	288
ASD6	4 (1)	7 (4)	69.7	84.7	30.3	15.3	295	307	301	305

LE : Lowest Rating on Engagement; HE : Highest Rating on Engagement  
SFC: Sum of Fixation Counts; FD<sub>AVG</sub>: Average Fixation Duration

participants did not inquire about future participation possibilities with VIGART. Thus, it is reasonable to infer that the individualized feedback provided by VIGART using dynamic gaze measurement as the session progressed contributed towards the higher engagement ratings.

#### C. Objective Measures of Engagement Through Behavioral Viewing Indexes [Sum of Fixation Counts (SFC) and Fixation Duration (FD)] Using VIGART

In the usability study we investigated the potential of VIGART to influence the behavioral viewing indexes of the participants. While caregivers' ratings showed an increase in engagement (Table III), we also wanted to investigate whether the objective measurement of behavioral viewing indexes, namely sum of fixation counts and fixation duration, showed corresponding variations. If such variations can be observed then these indexes can be used to map one's engagement within the VIGART, which may lead to new future intervention where objective metrics for engagement can be utilized within an intervention paradigm.

In the present study, for our analysis, we labeled trials as eliciting low engagement (LE) and high engagement (HE) states based on the lowest and highest ratings on engagement for each participant, as reported by the caregiver. Analysis of sum of fixation counts for the participants with respect to Face<sub>ROI</sub> and nonFace<sub>ROI</sub> (i.e., Object<sub>ROI</sub> and Others<sub>ROI</sub>) corresponding to trials reported as eliciting low engagement and high engagement states, reveals improved sum of fixation counts percentage towards Face<sub>ROI</sub> along with reduced fixation towards nonFace<sub>ROI</sub> with higher engagement for every participant (Table III). Group analysis shows an increase of 13.77% towards Face<sub>ROI</sub> and a decrease of 61.37% towards nonFace<sub>ROI</sub> with higher engagement (Table III). Not only did the participants visit the Face<sub>ROI</sub> more frequently, the average fixation duration of the participants indicates a trend towards improvement (i.e., increasing average fixation duration towards Face<sub>ROI</sub> in viewing patterns with increased engagement (Table III).

In short, our investigation reveals that the potential of VIGART to bring about an overall improvement in behavioral viewing patterns in terms of increased sum of fixation counts and average fixation duration towards the Face<sub>ROI</sub> accompanied with reduced distraction by NonFace<sub>ROI</sub>. Thus, both

TABLE IV  
AVERAGE FIXATION DURATION (FD<sub>AVG</sub>) AS MEASURE OF EMOTION RECOGNITION

	FD <sub>AVG</sub> (ms)		%Δ Neutral-To-Happy	FD <sub>AVG</sub> (ms)		%Δ Neutral-To-Angry
	(Neutral)	(Happy)		(Neutral)	(Angry)	
ASD1	315	342	8.62	315	339	7.84
ASD2	339	360	6.00	339	375	10.35
ASD3	298	329	10.39	298	356	19.66
ASD4	343	356	3.76	343	345	0.61
ASD5	272	267	-1.78	272	300	10.00
ASD6	295	307	4.09	295	309	4.63

the sum of fixation counts and the average fixation duration could be potential objective metrics for inferring engagement in VIGART-based intervention.

#### D. Objective Measures of Emotion Recognition Capability Through Behavioral Viewing Index Using VIGART

Further data analysis was carried out to examine the impact of the VIGART on the behavioral viewing indexes of our participants with ASD as function of their capability of recognizing emotions while they viewed their virtual classmates narrating personal stories and exhibiting context-relevant facial expressions. Note that our experiment was not designed to improve a participant's ability to correctly recognize an emotional expression, which will require a comprehensive investigation once the potential of VIGART is realized. However, since at the end of each trial, an emotion related question was asked to understand whether the participant was able to identify the facial expressions of his/her virtual classmates, we further performed an analysis to assess whether the emotion recognition capability and the gaze data were connected. In our usability study, we investigated fixation duration as a measure of emotion recognition [41] because of its greater importance when recognizing faces and emotional expressions as compared to the number of fixations as discovered in non-VR-based studies [21].

Our experimental findings revealed that all participants (except ASD5) fixated more on the face<sub>ROI</sub> of their virtual classmates displaying emotions, either positive (i.e., *Happy*) or negative (i.e., *Angry*), as compared to no emotion (i.e., *Neutral*). We found that there was an overall increase of 5.18% average fixation duration for *Neutral-to-Happy* expressions and an overall increase of 8.85% average fixation duration for *Neutral-to-Angry* expressions (Table IV). We examined the participants' responses to the emotion-identification questions and found that ASD5 could not identify the *Neutral* expression of his virtual classmate misidentifying it as *Happy* and we believe that this may account for his not looking more towards the Face<sub>ROI</sub>.

The above results indicate that VR-based social interaction can use average fixation duration as a quantitative measure to infer whether a participant with ASD has the ability to recognize emotion represented by facial expression, similar to the non-VR based tasks [21], [23]. In addition, the results show the potential of VIGART to map average fixation duration, measured in real-time to the participants' ability to recognize facial emotional expression of their virtual classmates. Thus, VIGART can be a step towards exploiting the behavioral viewing index in real-time as an objective measure which can serve as a tool while

TABLE V  
MEAN BLINK RATE  $BR_{MEAN}$  AS MEASURE OF EMOTION RECOGNITION

	$BR_{MEAN}$ (times/min)		% $\Delta$ <i>Neutral-To-Happy</i>	$BR_{MEAN}$ (times/min)		% $\Delta$ <i>Neutral-To-Angry</i>
	( <i>Neutral</i> )	( <i>Happy</i> )		( <i>Neutral</i> )	( <i>Angry</i> )	
ASD1	8.13	13.02	60.18	8.13	23.83	193.20
ASD2	5.42	8.13	50.00	5.42	8.63	59.23
ASD3	7.35	13.16	79.05	7.35	9.86	34.15
ASD4	12.39	14.42	16.38	12.39	15.20	22.68
ASD5	43.34	42.41	-2.16	43.34	71.50	64.96
ASD6	11.30	22.15	96.02	11.30	42.19	273.36

designing intervention strategies. However, a much larger study must be conducted before such findings can be generalized.

#### E. Eye Physiological Indexes (e.g., Blink Rate and Pupil Diameter) as Indicators of Emotion Recognition Capability

Children with ASD often experience states of emotional stress with autonomic nervous system activation without external expression [8] challenging their ability to communicate. However, their eye physiological indexes could be a valuable source to indicate their process of emotion recognition. In fact, literature indicates an important role of blink rate and pupil diameter in emotion recognition.

Past research shows startle potentiation for both positive and negative stimuli (static pictures) for children with ASD [42]. Additionally, pupil has been considered as an indicator of emotion recognition [43]. Studies indicate pupil to have sympathetic innervations with pupillary dilation to both pleasant and unpleasant auditory stimuli [29], being greater to unpleasant than pleasant visual stimuli [22].

Eye physiological indexes, namely blink rate and pupil diameter can be made continuously available within the VIGART using the feature extraction method (Section III-B2) when a participant socially interacts with the avatars. As a result, we analyzed to see whether the eye physiological indexes are also influenced by the interaction with VIGART.

*Analysis of Changes in Blink Rate:* In our present study, results as presented in Table V, reflect a higher change in mean blink rate for *Neutral-to-Angry* (an overall increase of 107.93%) than that for *Neutral-to-Happy* (an overall increase of 49.91%) for all participants, except ASD3. In addition, mean blink rate for all participants (except ASD5) while viewing the *Angry* and *Happy* facial expressions of their virtual classmates were greater than that while viewing *Neutral* expression. A detailed analysis revealed that ASD3 and ASD5 could not identify the avatar's *Angry* facial expression. ASD3 responded to the *Angry* face as *Not Sure* while ASD5 responded to it as *Happy*. Also, ASD5 could not identify the *Neutral* facial expression and misidentified this as *Happy* and he possessed a much higher mean blink rate in general, as compared to the other participants. Note that other researchers found similar trends in blink rate for non-VR based applications [42].

*Analysis of Changes in Pupil Diameter:* Our findings on pupil diameter as presented in Table VI, are also in line with some of the previous non-VR based findings ([22], [30]). We found that the mean pupil diameter of each participant was less while viewing the *Happy* faces than that for the *Angry* faces and both being greater than that for *Neutral* faces (except ASD5).

TABLE VI  
MEAN PUPIL DIAMETER  $PD_{MEAN}$  AS MEASURE OF EMOTION RECOGNITION

	$PD_{MEAN}$ (mm)		% $\Delta$ <i>Neutral-To-Happy</i>	$PD_{MEAN}$ (mm)		% $\Delta$ <i>Neutral-To-Angry</i>
	( <i>Neutral</i> )	( <i>Happy</i> )		( <i>Neutral</i> )	( <i>Angry</i> )	
ASD1	8.036	8.809	9.61	8.036	10.713	33.30
ASD2	6.213	6.272	0.96	6.213	6.394	2.92
ASD3	6.730	6.966	3.51	6.730	7.215	7.21
ASD4	6.741	6.817	1.12	6.741	7.262	7.73
ASD5	7.678	8.032	4.61	7.678	7.711	0.44
ASD6	7.209	7.264	0.77	7.209	7.311	1.42

Also, the change in mean pupil diameter for *Neutral-to-Angry* (overall increase of 8.84%) was found to be greater than that for *Neutral-to-Happy* (overall increase of 3.43%) for all participants (except ASD5) (Table VI). We examined the participants' responses to the emotion-identification questions and found that ASD5 could not identify the *Neutral* and *Angry* facial expressions, misidentifying them as *Happy*. We believe that this can be the reason for his percent increase in mean pupil diameter for *Neutral-to-Angry* to be lower than that for *Neutral-to-Happy*.

In the present usability study, with a limited sample size, we find that the blink rate of the participants is more sensitive to their ability to recognize different emotional expressions exhibited by their virtual classmates, than the pupil diameter. For pupil diameter, the percent change for *Neutral-to-Happy* and that for *Neutral-to-Angry* though quite small the overall trend is similar to that of other non-VR based studies. More importantly, the above results indicate the ability of VIGART to measure one's blink rate and pupil diameter in real-time and correlate these as function of his/her ability to recognize emotions while interacting socially with virtual classmates. The results suggest that a participant's eye physiological response in VR-based social interaction as presented in VIGART may indicate whether or not one is able to recognize emotion, which is similar to what has been observed in non-VR based tasks. Thus, it is reasonable to believe that VIGART could be used in intervention, perhaps as a supplementary tool, to allow a child with ASD to enhance his/her social skills.

## VI. CONCLUSION

Children with ASD are characterized by atypicalities in viewing patterns during social tasks. These deficits coupled with the wide spectrum of ASD necessitate the use of individualized intervention. Research indicates the potential of emerging technology to play a significant role in providing individualized intervention. With this motivation, in the current study, we design a novel VIGART, that seamlessly integrates the VR technology with dynamic eye-tracking to create a novel virtual gaze-sensitive social interaction platform and provides individualized feedback based on one's performance and real-time viewing.

In the present work, we designed a usability study with six adolescents with ASD, to test the acceptability of VIGART and also investigate how these adolescents respond in terms of their engagement level, behavioral viewing and eye physiological indexes. From the exit survey, we can infer that the VIGART has a potential to be accepted by the target population. The VIGART can measure behavioral viewing and eye physiological

indexes dynamically that have the potential to be used as objective measures of one's engagement as well as emotion recognition capability. Such capabilities, in our opinion, make VIGART unique in the sense that this intelligent system could be effectively used as a potential intervention tool where objective quantifiable measures could guide refinement of intervention strategies.

The preliminary results of the usability study are promising. Overall, our results indicate that the impact of VIGART is similar to other non-VR-based tasks. However, a much larger study must be conducted before such findings can be generalized. Also note that, at this current configuration, VIGART uses a wearable eye-tracker, which may not be suitable for children with low-functioning ASD. However, we plan to use a noncontact desktop-based eye-tracker to mitigate this problem. It is worth mentioning that VIGART is capable of providing extensive closed-loop interaction. For example, it can provide continuous information to the avatars about whether and for how long the participant is looking at them during interaction, and accordingly the avatars can change their mode of talking to allow a higher level of interactivity between the participant and VIGART. Such sophisticated interaction will be performed in the future. The present usability study shows, in principle, that VIGART has the potential to be used as a supplement to real-life social skills training tasks in an individualized manner. However, we acknowledge that current findings, particularly toward skill improvement, are preliminary. While demonstrating proof of concept of the technology and trends of "improved" behavioral viewing in a VR-based social task, questions about the practicality, efficacy, and ultimate benefit of the use of this technological tool for demonstrating clinically significant improvements in terms of ASD impairment remain, which will be addressed by empirical investigation in the future.

#### ACKNOWLEDGMENT

The authors would like to thank the participants and their families for making this study possible.

#### REFERENCES

- [1] "CDC, prevalence of autism spectrum disorders-ADDM network," *MMWR Surveill. Summ. U.S.*, 2006, vol. 58, pp. 1–20, 2009.
- [2] "APA, diagnostic and statistical manual of mental disorders: DSM-IV-TR," in *Am. Psychiatric Assoc.*, Washington, DC, 2000.
- [3] S. Baron-Cohen, *Mindblindness: An Essay on Autism and Theory of Mind*. Cambridge, MA: MIT Press, 1997.
- [4] M. D. Rutherford and M. T. Towns, "Scan path differences and similarities during emotion perception in those with and without autism spectrum disorders," *J. Aut. Dev. Disord.*, vol. 38, pp. 1371–1381, 2008.
- [5] C. Y. Trepagnier, M. M. Sebrechts, A. Finkelmeyer, W. Stewart, J. Woodford, and M. Coleman, "Simulating social interaction to address deficits of autistic spectrum disorder in children," *Cyberpsychol. Behav.*, vol. 9, no. 2, pp. 213–217, 2006.
- [6] K. A. Pelphrey, N. J. Sasson, J. S. Reznick, G. Paul, B. D. Goldman, and J. Piven, "Visual scanning of faces in autism," *J. Autism Develop. Disorders*, vol. 32, pp. 249–261, 2002.
- [7] NRC, *Educating Children With Autism* Washington, DC, Nat. Acad. Press, 2001.
- [8] R. W. Picard, "Future Affective Technology for Autism and Emotion Communication," *Philos. Trans. R. Soc. B*, vol. 364, no. 1535, pp. 3575–3584, 2009.
- [9] S. Parsons, P. Mitchell, and A. Leonard, "The use and understanding of virtual environments by adolescents with autistic spectrum disorders," *J. Autism Dev. Disord.*, vol. 34, no. 4, pp. 449–466, 2004.
- [10] A. Tartaro and J. Cassell, *Using Virtual Peer Technology as an Intervention for Children with Autism, Towards Universal Usability: Designing Computer Interfaces for Diverse User Populations*, J. Lazar, Ed. Chichester, U.K.: Wiley.
- [11] G. Pioggia, R. Iglizzi, M. Ferro, A. Ahluwalia, F. Muratori, and D. De Rossi, "An android for enhancing social skills and emotion recognition in people with autism," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 13, no. 4, pp. 507–515, Dec. 2005.
- [12] M. S. Goodwin, "Enhancing and accelerating the pace of autism research and treatment: The promise of developing innovative technology," *Focus Autism Other Dev. Dis.*, vol. 23, no. 2, pp. 125–128, 2008.
- [13] S. Parsons and P. Mitchell, "The potential of virtual reality in social skills training for people with autistic spectrum disorders," *J. Intell. Disabil. Res.*, vol. 46, pp. 430–443, 2002.
- [14] D. Strickland, "Virtual reality for the treatment of autism," in *Virtual Reality in Neuropsychophysiology*, G. Riva, Ed. Fairfax, VA: IOS Press, 1997, pp. 81–86.
- [15] M. Wilms, L. Schilbach, U. Pfeiffer, G. Bente, G. R. Fink, and K. Voegele, "It's in your eyes using gaze-contingent stimuli to create truly interactive paradigms for social cognitive and affective neuroscience," *Social Cognitive Affective Neurosci.*, vol. 5, no. 1, pp. 98–107, 2010.
- [16] S. Weider and S. Greenspan, "Can children with autism master the core deficits and become empathetic, creative, and reflective?," *J. Develop. Learn. Disorders*, vol. 9, 2005.
- [17] C. Y. Pan, "Age, social engagement, and physical activity in children with autism spectrum disorder," *Res. Autism Spectrum Disorders*, vol. 3, pp. 22–31, 2009.
- [18] D. B. Shalom, S. H. Mostofsky, R. L. Hazlett, M. C. Goldberg, R. J. Landa, Y. Faran, D. R. McLeod, and R. Hoehn-Saric, "Normal physiological emotions but differences in expression of conscious feelings in children with high-functioning autism," *J. Autism Develop. Disorders*, vol. 36, no. 3, pp. 395–400, 2006.
- [19] C. J. Anderson, J. Colombo, and D. J. Shaddy, "Visual scanning and pupillary responses in young children with autism spectrum disorder," *J. Clin. Exp. Neuropsych.*, vol. 28, pp. 1238–1256, 2006.
- [20] J. W. Denver, "The social engagement system: Functional differences in individuals with autism," Ph.D. dissertation, Dept. Human Develop., Univ. Maryland, College Park, 2004.
- [21] J. H. Hsiao and G. Cottrell, "Two fixations suffice in face recognition," *Psychol. Sci.*, vol. 19, no. 10, pp. 998–1006, 2008.
- [22] W. L. Libby, B. C. Lacey, and J. I. Lacey, "Pupillary and cardiac activity during visual attention," *Psychophysiology*, vol. 10, no. 3, pp. 270–294, 1973.
- [23] S. Ozdemir, G. Universitesi, G. E. Fakultesi, and O. E. Bolumu, "Using multimedia social stories to increase appropriate social engagement in young children with autism," *Turk. J. Edu. Tech*, vol. 7, no. 3, pp. 80–88, 2008.
- [24] L. Itti, N. Dhavale, and F. Pighin, "Realistic avatar eye and head animation using a neurobiological model of visual attention," in *Proc. Int. Symp. Opt. Sci. Technol.*, 2003, vol. 5200, pp. 64–78.
- [25] E. T. Hall, *The Hidden Dimension*. New York: Doubleday, 1966.
- [26] Dynamic indicators of basic early literacy skills 2007 [Online]. Available: [dibels.uoregon.edu/measures/](http://dibels.uoregon.edu/measures/)
- [27] Eye-Tracker goggles, Arrington Res. [Online]. Available: <http://www.arringtonresearch.com/>
- [28] L. A. Ruble and D. M. Robson, "Individual and environmental determinants of engagement in autism," *J. Autism Develop. Disorders*, vol. 37, no. 8, pp. 1457–1468, 2006.
- [29] T. Partala, M. Jokiniemi, and V. Surakka, "Pupillary responses to emotionally provocative stimuli," *Eye Track. Res. Appl.*, pp. 123–129, 2000.
- [30] C. J. Anderson and J. Colombo, "Larger tonic pupil size in young children with autism spectrum disorder," *Develop. Psychobiol.*, vol. 51, no. 2, pp. 207–211, 2009.
- [31] Data Collection. ViewPoint EyeTracker: PC-60 Software User Guide Arrington Res., 2002, p. 47.
- [32] R. J. K. Jacob, *Eye Tracking in Advanced Interface Design, in Advanced Interface Design and Virtual Environments*, W. Barfield and T. Furness, Eds. Oxford, U.K.: Oxford Univ., 1994.
- [33] H. Z. Li, "Culture and gaze direction in conversation," *RASK*, vol. 20, pp. 3–26, 2004.
- [34] M. Argyle and M. Cook, *Gaze and Mutual Gaze*. Cambridge, MA: Cambridge Univ. Press, 1976.
- [35] A. Colburn, S. Drucker, and M. Cohen, "The role of eye-gaze in avatar-mediated conversational interfaces," in *SIGGRAPH Sketches Appl.*, New Orleans, LA, 2000.

- [36] M. J. Wieser, P. Pauli, G. W. Alpers, and A. Mühlberger, "Is eye to eye contact really threatening and avoided in social anxiety?—An eye-tracking & psychophysio. Study," *J. Anxiety Disorders*, vol. 23, pp. 93–103, 2009.
- [37] L. M. Dunn and L. M. Dunn, *PPVT-III: Peabody Picture Vocabulary Test*, 3rd ed. Circle Pines, MN: Am. Guidance Ser., 1997.
- [38] J. N. Constantino, *The Social Responsiveness Scale*. Los Angeles, CA: Western Psych. Services, 2002.
- [39] M. Rutter, A. Bailey, S. Berument, C. Lord, and A. Pickles, *Social Communication Questionnaire*. Los Angeles, CA: Western Psych. Services, 2003.
- [40] C. Lord, S. Risi, L. Lambrecht, E. H. Cook, Jr, B. L. Leventhal, P. C. DiLavore, A. Pickles, and M. Rutter, "The autism diagnostic observation schedule-generic: A standard measure of social and communication deficits associated with the spectrum of autism," *J. Autism Develop. Disorders*, vol. 30, no. 3, pp. 205–23, 2000.
- [41] J. D. Buckner, J. K. Maner, and N. B. Schmidt, "Difficulty disengaging attention from social threat in social anxiety," *Cognitive Therapy Res.*, vol. 34, no. 1, pp. 99–105, 2010.
- [42] J. L. Wilbarger, D. N. McIntoshb, and P. Winkielmanc, "Startle modulation in autism: Positive affective stimuli enhance startle response," *Neuropsychologia*, vol. 47, pp. 1323–1331, 2009.
- [43] M. B. Bradley, L. Miccoli, M. A. Escrig, and P. J. Lang, "The pupil as a measure of emotional arousal and autonomic activation," *Psychophysiology*, vol. 45, no. 4, pp. 602–607, 2008.



**Uttama Lahiri** (S'11) received the Ph.D. degree from Vanderbilt University, Nashville, TN, in 2011. Currently she is working as a Postdoctoral fellow at Vanderbilt University.

Her research interests include adaptive response systems, signal processing, machine learning, human–computer interaction, and human–robot interaction.

Dr. Lahiri is a member of the Society of Women Engineers.



**Zachary Warren** received the Ph.D. degree from the University of Miami, Miami, FL, in 2005.

He completed his predoctoral internship at Children's Hospital Boston/Harvard Medical School in and his postdoctoral fellowship at the Medical University of South Carolina. Currently he is an Assistant Professor of Pediatrics and Psychiatry at Vanderbilt University. He is the Director of the Treatment and Research Institute for Autism Spectrum Disorders at the Vanderbilt Kennedy Center as well Director of Autism Clinical Services within the Division of Genetics and Developmental Pediatrics.



**Nilanjan Sarkar** (SM'10) received the Ph.D. degree from the University of Pennsylvania, Philadelphia, in 1993.

He joined Queen's University, Kingston, ON, Canada as a Postdoctoral fellow and then went to the University of Hawaii as an Assistant Professor before joining Vanderbilt University, Nashville, TN, in 2000. Currently he is a Professor of Mechanical Engineering, and Computer Engineering. His research focuses on human–robot interaction, affective computing, dynamics, and control.

Dr. Sarkar served as an Associate Editor for the IEEE TRANSACTIONS ON ROBOTICS.