Int J Soc Robot (2010) 2: 391–403
DOI 10.1007/s12369-010-0063-x

An Approach to the Design of Socially Acceptable Robots for Children with Autism Spectrum Disorders

Karla Conn Welch · Uttama Lahiri · Zachary Warren · Nilanjan Sarkar

Accepted: 21 June 2010 / Published online: 7 July 2010

Abstract Investigation into technology-assisted intervention for children with autism spectrum disorders (ASD) has gained momentum in recent years. Research suggests that robots could be a viable means to impart skills to this population since children with ASD tend to be fascinated by robots. However, if robots are to be used to impart social skills, a primary deficit for this population, considerable attention needs to be paid to aspects of social acceptability of such robots. Currently there are no design guidelines as to how to develop socially acceptable robots to be used for intervention for children with ASD. As a first step, this work investigates social design of virtual robots for children with ASD. In this paper we describe the design of a virtual environment system for social interaction (VESSI). The design is evaluated through an innovative experiment plan that combines subjective ratings from a clinical observer with physiological responses indicative of affective states from the participants, both collected when participants engage in social tasks with the social robots in a virtual reality environment. Two social parameters of importance for this population, namely eye gaze and social distance, are systematically varied to analyze the response of the participants. The results are presented to illustrate how experiments with virtual social robots can contribute towards the development of future social robots for children with ASD.

Keywords Virtual robots · Identification of emotional expressions · Social interaction · Autism intervention

K.C. Welch (✉)
Department of Electrical and Computer Engineering, University of Louisville, 448 Lutz Hall, Louisville, KY 40292, USA
e-mail: karla.welch@louisville.edu

U. Lahiri · N. Sarkar
Department of Mechanical Engineering, Vanderbilt University, Nashville, USA

U. Lahiri
e-mail: uttama.lahiri@vanderbilt.edu

Z. Warren
Department of Pediatrics, Vanderbilt Kennedy Center, Treatment and Research Institute for Autism Spectrum Disorders (TRIAD), Nashville, USA
e-mail: zachary.warren@vanderbilt.edu

N. Sarkar
Department of Computer Engineering, Vanderbilt University, Nashville, USA
e-mail: nilanjan.sarkar@vanderbilt.edu

1 Introduction

Autism encompasses a wide variety of symptoms but generally is characterized by impairments in social interaction, social communication, and imagination, along with repetitive behavior patterns [1]. Emerging research suggests that prevalence rates as high as approximately 1 in 110 for the broad autism spectrum [8]. While there is at present no single accepted intervention, treatment, or known cure for autism spectrum disorders (ASD), there is growing consensus that intensive behavioral and educational intervention programs can significantly improve long term outcomes for individuals with ASD and their families [10, 38, 44].

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 [39, 45]. 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 [5, 6, 53], robotic systems [15, 31, 37], and virtual reality environments [40, 52, 54]. There is increasing consensus in the autism community that development of assistive tools that exploit advanced technology will make application of intensive intervention for children with ASD more efficacious.

The design of social robots is a developing field in which standards are still being established. In general, a social robot is an autonomous agent that can act in a socially appropriate manner based on its role in an interaction [16, 21], such as interacting with a child with ASD. Initial results indicate that robots may hold promise for rehabilitation of children with ASD. Dautenhahn and Werry [15] have explored how a robot can become a playmate that might serve a therapeutic role for children with autism. Robots can allow simplified but embodied social interaction for typical children [55] and may offer interaction that is less intimidating or confusing than human-to-human interaction specifically for children with ASD. Investigation of the impact of robot design on interactions with children found a need to emphasize systems that are versatile enough to adapt to the varying needs of different children [37]. Pioggia et al. [41] developed an interactive life-like facial display system for enhancing emotion recognition in individuals with ASD. Robots have also 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 [15, 31, 43]. Robotic technology poses 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/or diagnosis [47]. Our earlier work [34, 35] suggests that endowing a robot with an adaptive ability to recognize and respond to the affective states of a child with ASD based on physiological information could be a viable means for autism intervention. The results demonstrated for the first time that affect-sensitive adaptation improved the performance as well as enhanced how much the children with ASD liked interacting with the robot.

While the above research indicates the usefulness of robot-assisted intervention for children with ASD, there are no design guidelines as to how to develop socially acceptable robots to be used for social skill intervention for children with ASD. In particular, it is important to know how these robots should display intentions (e.g., through the use of facial expressions, gestures, verbal communication, etc.) and how they should interact (e.g., amount of eye contact, proximity to the child, etc.) to ascertain the intended social skill teaching to this population. Additionally, it is also important to develop an evaluation method that is not dependent on self-report because of the known difficulty of self expression exhibited by children with ASD [27].

It is thus imperative to systematically develop social robots and study social interaction with children with ASD in a step-by-step manner. Similar to how sophisticated machines are designed in virtual environments prior to fabrication, we design social robots in virtual environments that would enable a systematic evaluation and manipulation of different components of social interaction through virtual social robots.

Virtual reality (VR) represents a medium well-suited for creating interactive intervention paradigms for skill training in the core areas of impairment for children with ASD. VR-based therapeutic tools can partially automate the time-consuming, routine behavioral therapy sessions and may allow intensive intervention to be conducted at home [52]. Furthermore, VR has also shown the capacity to ease the burden, both time and effort, of trained therapists 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 [50].

We describe the design and development of VESSI, a virtual environment system for social interaction that is capable of systematic manipulation of various design parameters that are important for the development of social robots. VESSI 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 social robots for children with ASD.

This paper describes an investigation into socially-driven virtual reality interactions to guide future intervention of children with ASD. We describe the socialization and expressivity of the VR characters. The VR environment monitors affective changes during social contexts for children with ASD. In particular, we study how the affective state of anxiety; 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 on VESSI. Finally, we discuss how our findings, regarding these virtual social robots and their interactions, can be useful in the development of future social robots for this target population.

2 Task Design

For ASD intervention, VR is often effectively experienced on a desktop system using standard computer input devices [39]. 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 (e.g., nausea, headaches, or dizziness) potentially associated with head-mounted devices [9]. Therefore, users view VESSI on a computer monitor from the first-person perspective.

We created realistic VR scenarios for interaction with virtual social robots (i.e., expressive humanoid avatars). Vizard (worldviz.com), a commercially available VR design package, was employed to develop the environments. Within the controllable VR environment, components of the interaction were systematically manipulated to allow users to explore different social compositions. The virtual robots made direct and averted eye contact. They conversed by matching their mouth movements to recorded sound files. The user responded to the virtual robots via pop-up text boxes.

2.1 Social Parameters

Eye gaze and social distance, the social parameters of interest, are organized in a 4×2 experimental design, allowing investigation of eight distinct situations. Although several parameters make up a social situation; such as group size, vocal tone, facial expressions, gestures/body movement, surrounding environment, etc.; eye gaze and social distance are chosen because they play significant roles in social communication and interaction [1–4]. Future studies will explore additional social parameters on a step-by-step basis; however, these parameters are chosen for initial study because they represent key factors of possible stressors and areas of deficiency for children with ASD [15, 43]. 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 [30, 51]. Each trial of an experiment session includes one virtual robot 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 robot 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 [2, 48, 49]. Social parameters such as facial expression, vocal tone, and environment 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 a virtual robot 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 virtual robot 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 [3] and algorithms adopted by the artificial intelligence community to create realistic virtual characters [11, 23], 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 robot in VESSI is speaking, we use the "normal" gaze definitions for a person speaking, which is approximately 30% straight gaze and 70% averted gaze [3]. Research represents averted gaze as looking more than 10° away from center in evenly-distributed, randomly-selected directions [23, 28]. 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 virtual robot and the user. Two types of social distance, termed "invasive" and "decorum," are examined. In VESSI, distance is simulated but can be appropriately represented to the view of the participant [32]. For invasive distance, the virtual robot 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 [25]. A distance of 1.5 ft. apart has been investigated by several research groups in social interaction experiments with similar experimental setups to ours in which two people are specifically positioned while one participant listens as the other introduces himself/herself and discusses a personal topic for approximately 2 min [2, 48, 49], 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 [2]. Decorum distance means the virtual robot 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 [26], and research indicates this distance results in a more comfortable conversation experience than the invasive distance [2]. Using Vizard software we project virtual social robots who display different eye gaze patterns at different distances; two examples are shown in Fig. 1.

In this work, our aim is to identify and evaluate the strength of systematic manipulation of social parameters (i.e., eye gaze and social distance) through VR-based social tasks in eliciting variations in an affective state (i.e., anxiety) and map such variations with changes in physiological

Fig. 1 At *top* a virtual robot displays straight gaze at an invasive distance, while on *bottom* a virtual robot stands at a decorum distance and looks to her left in an averted gaze

signals of the participants. This work is the first step towards building a more back-and-forth interaction. If there is effect on physiological response even in this task, then we can assign the effect on the social parameter manipulation alone.

2.2 Humanoid Avatars

The virtual social robots 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, 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.

The stories the virtual robots share are adapted from DIBELS (Dynamic Indicators of Basic Early Literacy Skills; dibels.uoregon.edu/measures/) reading assessments. 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 our previous research in the design of an easy-to-medium level of the Anagrams game used with a similar recruitment pool [35]. In each trial of the experiment, a virtual robot 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.

2.3 Social Interaction

The interaction involves a virtual robot telling a story while a participant listens. At the end of the story, the virtual robot 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. The correct choice is spoken at least five times during the story, which is sufficient for the information to be relayed [29], and the incorrect choices are never spoken in the story. For example, one story is about a bus breaking down on the way to a school picnic. At the end of the story the virtual robot 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.

3 System Refinement

Efforts were made to minimize reactions due solely to viewing a virtual robot by choosing the 10 most-neutral avatar heads based on a survey of 20 participants. 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 virtual robot alone.

Participants for the avatar head survey were recruited from undergraduate engineering and psychology courses at a private Southeastern 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 avatar heads (see Table 1). Participants were asked to rate each avatar head on four questions. Two questions were designed to measure elicited reactions from viewing the avatar heads (Q1 and Q2), and two were designed to determine participants' perceptions of each avatar head's display of emotion (Q3 and Q4). Following descriptive terms from [33], participants were asked to rate how they felt when looking at each avatar head on a 5-point scale of valence and arousal. To gauge affective reactions to images, valence and arousal are important measures to consider [7]. Participants were also asked two questions about what emotion the avatar head was conveying [18, 22].

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

Unhappy, Annoyed, Despaired		Neutral		Happy, Pleased, Hopeful
○	○	○	○	○

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

Calm, Sleepy, Unaroused		Neutral		Excited, Wide-awake, Aroused
○	○	○	○	○

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 Table 1) 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.

3.1 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.

$$E_{norm,i} = \frac{\sqrt{(V_{mean,i} - 0)^2 + (A_{mean,i} - 0)^2}}{2\sqrt{2}} \tag{1}$$

Equation (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 arousal 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 (1) the shortest distance away from 0,0 in the valence-arousal space, (2) the least portion of respondents answering "a specific emotion" to Q3, and (3) 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 been shown to be a reliable assessment of affect and proven to sufficiently cover the affective space [7]. In this survey we also took into account, but to a lesser extent, whether the avatar heads were perceived

Table 1 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. Overall 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 (2))

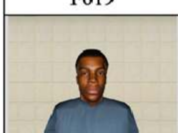
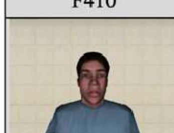
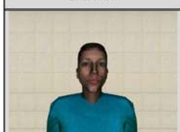
Screenshot	Mean Survey Measurement across Respondents (N=20)					
	Q1[a]	Q2[b]	Eqn. 1[c]	Q3[d]	Q4[e]	Eqn. 2[f]
 F074	-0.80	-0.65	0.36	0.45	0.34	0.37
 M304	-0.25	-0.05	0.09	0.20	0.10	0.10
 M097	-0.50	-0.15	0.18	0.60	0.28	0.24
 F232	-0.40	-0.60	0.25	0.25	0.10	0.24
 F619	-0.75	0.20	0.27	0.60	0.25	0.30
 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
 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
 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

Table 1 (*Continued*)

Screenshot	Mean Survey Measurement across Respondents (N=20)					
	Q1[a]	Q2[b]	Eqn. 1[c]	Q3[d]	Q4[e]	Eqn. 2[f]
 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
 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]Equation (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]Equation (2): Weighted sum of (1), Q_3, and Q_4

as expressing little to no emotion/neutral emotion. The combined equation used to rate the avatar heads is as follows,

$$R_i = 0.8(E_{norm,i}) + 0.1(Q_{3,i}) + 0.1(Q_{4,i}) \quad (2)$$

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 1).

3.2 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 robots in the experiment. 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.) [19, 22], Vizard has not tested user perception of the morphs to establish if the morphs convey the designed emotion to the viewer.

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 2, are sorted in ascending order when scanned from left to right, top to bottom. The ten avatar heads with the lowest scores were used for the virtual social robots in the VR experiments. The four lowest female avatar heads and four lowest male avatar heads were used in the eight experiment conditions. The female and male avatar heads with the highest rating of the bottom 10 were used for the demonstration of the VR interaction during session one of the experiment. Each of the eight experiment conditions were shown three times, creating 24 trials in the experiment, which were divided over two sessions on two different days for each participant.

4 Experiment Protocol

4.1 Participant Recruitment

Participants were 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.

Table 2 Shown are the 10 avatar heads with the most-neutral ratings from the survey. These avatar heads were used for the virtual social robots 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)						EC	Screenshot	Mean Survey Measurement across Respondents (N=20)					
		Q_1[a]	Q_2[b]	Eqn. 1[c]	Q_3[d]	Q_4[e]	Eqn. 2[f]			Q_1[a]	Q_2[b]	Eqn. 1[c]	Q_3[d]	Q_4[e]	Eqn. 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

For all participants, the Social Responsiveness Scale (SRS) [12] profile sheet and Social Communication Questionnaire (SCQ) [46] were completed by a participant's caregiver before the first session to provide an index of current functioning and ASD symptom profiles. The SRS is a 65-item questionnaire designed to quantitatively measure the severity of autism-related symptoms. It generates a total 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 SCQ is a brief instrument for the valid screening or verification of ASD symptoms that has been developed from three critical autism diagnostic domains of qualitative impairments in reciprocal social interaction, communication, and repetitive and stereotyped patterns of behavior. Selection was also based on a receptive vocabulary standard score of 80 or above on the PPVT-III (Peabody Picture Vocabulary Test-3rd Edition). The PPVT-III [17] was used as an inclusion criterion in our previous research [34, 35]. 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 were able to perform the interaction tasks involved in this study. All written components of the current design were accompanied by audio readings, thus alleviating some of the language requirements and could open the prospect of including younger participants or those with less language and/or reading skills in future studies.

4.2 Procedure

The commitment required of participants was a total of two sessions (i.e., approximately 2.5 hrs). The first session ran approximately 1.5 hrs, due to gathering consent and assent, administering the PPVT-III, and running demonstrations of the social task. The second session lasted about 1 hr. For each completed session, a participant received compensation in the form of gift cards. A parent of each participant observed their child's experiment sessions and provided feedback in a brief post-interview.

Our hypothesis was that manipulation of the social parameters may elicit variations in affective reactions [3] and physiological responses [20, 24]. 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 virtual social robots. 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 (i.e., low anxiety or high anxiety) from a clinical observer were recorded while the participant interacted with VESSI. Given the fact that clinical observers' judgment based on their expertise is the state-of-the-art in most autism intervention approaches and the results about the reliability of the subjective reports in our previous studies [35], the reports from the clinical observer were used as the reference points linking the objective physiological data to the participant's affective state. The physiological signals were processed to 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 with the clinical observer's report to relate the participant's affective reactions and physiological responses with respect to the various social stimuli.

The physiological signals recorded in this work included electrocardiogram (ECG), impedance cardiogram (ICG), galvanic skin response (GSR), and photoplethysmogram (PPG). These signals were collected using a Biopac MP150 system (biopac.com) and small, wearable sensors placed on a participant's chest (ECG), neck and torso (ICG), and first through third fingers on the participant's left hand (GSR and PPG). 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 and children with ASD [35, 42]. Results from variation of features extracted from these signals are shown in the next section. The features of interest included GSR phasic response rate (GSR_{prr}), PPG peak maximum (PPG_{pmax}), and time of pre-ejection period (PEP). GSR_{prr} is measured in responses per minute (rpm) and represents a rapid increase in skin conductance similar to a peak. PPG_{pmax} 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. A detailed description of the sensor placement, signal processing, feature extraction routines, and significance of these physiological features for indicating an affective response can be found in our previous work [35].

The equipment setup includes a computer dedicated to the social interaction tasks where the participants interacted with VESSI, Biopac biological feedback equipment that collected the physiological signals of the participant, and another PC dedicated to acquiring signals from the Biopac system. The Vizard Virtual Reality Toolkit ran on a computer 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. We also video recorded the sessions to cross-reference observations made during the experiment. The clinical observer and a participant's parent watched the participant from the view of the video camera, whose signal was routed to a television hidden from the view of the participant. The signal from the participant's computer screen where the task was presented was routed to a separate computer monitor so that the clinical observer and parent could view how the task progressed.

Each participant engaged in two VR-based social interaction sessions on two different days. During the first session, the participants were told about the experiment purpose, the sensors, and the VESSI tasks. After the physiological sensors were placed, the participants were asked to relax quietly for three minutes while a baseline recording of physiological signals was taken that was used to offset day-variability. The first session included two demonstrations of the VR task, the baseline physiological measurement, and a set of eight 2-min trials with different virtual social robots. The second session consisted of the baseline physiological measurement and the remaining 16 trials of social interaction tasks. After each trial, the participant answered the story question and the clinical observer rated what she thought the level (i.e., low or high) of the affective state of anxiety was for the participant during the finished trial.

5 Results

A group of 13 (10 male) children with ASD and a matched group of TD children completed initial testing of the system. Their characteristics are shown in Table 3. The children with ASD had a confirmed diagnosis using DSM-IV criteria as well as scores from SRS (cutoff = 60) and SCQ (cutoff = 12) assessments [1, 12, 46]. The TD children did not meet cutoffs for ASD on either the SRS or SCQ. Each child in the two groups completed two sessions with the virtual social robots. 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 ($p > 0.05$, $p = 0.2601$).

Evidence of overt behaviors as well as more subtle reactions to the different experiment conditions was demonstrated. Several ASD children showed considerable reactions to the virtual social robots standing at the invasive dis-

Table 3 The participant group characteristics are listed. The participants were matched by gender, age, and PPVT standard score

Group	Age (yrs)	PPVT[g]	SRS[h]	SCQ[i]
ASD ($N = 13$)				
M (SD)	16.0 (1.7)	105.9 (14.0)	79.5 (9.9)	21.9 (6.3)
TD ($N = 13$)				
M (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

[g]Peabody Picture Vocabulary Test-3rd edition Standard score [16]
[h]Social Responsiveness Scale Total T-score [12, 13]
[i]Social Communication Questionnaire Total score [14, 46]
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)

tance 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, the parents of these children were surprised to observe such a stark reaction to the change in stimuli. Although accustomed to withdrawing behavior in complex, overwhelming social situations, the parents agreed that the story content and virtual robots' facial expressions were neutral and were therefore perplexed to see such a reaction from their children to the change in distance or eye gaze alone. These reactions and reflections highlight an advantage that systems like VESSI 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; VESSI can systematically manipulate each element of an interaction and observe the effect. Therefore, VESSI can go beyond identifying a broad scope of situations that are anxiety-inducing. VESSI 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 overt reactions reflected ratings on affective states from the clinical observer and the subtle variations in physiological signals during the experiment trials. The ASD group's physiological signals showed significant changes to trials rated as eliciting "low anxiety" (LA) versus "high anxiety" (HA) according to the clinical observer label (COL). 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. As shown in Table 4, both the ASD and TD group had a significant increase in GSR_{prr}, 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, GSR_{prr} significantly increased. 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_{pmax}, but the ASD group did not (see Table 5). The experiment condition of Averted eye gaze with distance varying elicited a significant increase in PEP for the ASD group but not the TD group (see Table 6).

Table 4 Listed are results of GSR_{prr} 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

COL	Percent (%) of trials rated as LA or HA by Group	ASD group GSR_{prr} (rpm) M (SD)	TD group GSR_{prr} (rpm) M (SD)
LA	ASD 24, TD 69	4.43 (2.75)	3.23 (2.78)
HA	ASD 76, TD 31	5.80 (3.55)	4.46 (3.57)
	t-Value	−2.18	−2.33
	p-Value	<0.05*	<0.05*
	Exact p-value	0.0311	0.0211

Significant differences, $^{*}p < 0.05$

Table 5 Listed are results of PPG_{pmax} 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

COL	Percent (%) of trials rated as LA or HA by Group	ASD group PPG_{pmax} (μV) M (SD)	TD group PPG_{pmax} (μV) M (SD)
LA	ASD 50, TD 77	2.93 (3.12)	1.21 (1.18)
HA	ASD 50, TD 23	4.24 (4.39)	3.42 (4.67)
	t-Value	−1.52	−3.37
	p-Value	ns	<0.05*
	Exact p-value	0.1329	0.0012

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 in terms of their physiological reaction (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 for the Invasive and Decorum distance, 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.

Table 6 Listed are results of PEP 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

COL	Percent (%) of trials rated as LA or HA by Group	ASD group PEP (ms) *M* (SD)	TD group PEP (ms) *M* (SD)
LA	ASD 69, TD 69	156.61 (23.49)	145.54 (20.91)
HA	ASD 31, TD 31	143.85 (26.57)	146.79 (19.63)
	t-Value	2.13	−0.25
	p-Value	<0.05*	ns
	Exact p-value	0.0367	0.8044

Significant difference, $^{*}p < 0.05$

Therefore, the VR system shows it can elicit variations in both affective ratings and physiological signals to changes in social experimental stimuli. The presented physiological features were chosen to showcase that significant reactions were observed. The results give an introduction on how the participants reacted to VESSI; the full psychophysiological impact will be told with further analysis. No one physiological feature is necessarily more revealing than another. However, some physiological features are more commonly studied than others, which present richer opportunities for future comparisons between the current experiment and previous studies. The current findings are similar to observations in social anxiety research of typical adults in real-world settings [2, 48, 49] but have now been examined with observations and physiological signals for ASD and TD children in a virtual interaction. This research is the first step towards examining how children react to and accept the virtual social robots as realistic to real-world settings. Establishing realistic interactions builds a basis for creating more complex settings for intervention and will guide design of real-world social robots for embodied social communication intervention.

6 Conclusions and Future Work

Social communication and social information processing are thought to represent core domains of impairment in children with ASD. The results show the VR system elicits variations in both affective ratings and physiological signals to changes in experimental social stimuli for children with ASD and TD children. Further analysis of the variation of additional physiological features compared to the manipulation of all conditions of the social parameters is warranted and forthcoming. These current results establish a statistically significant basis for continuing the research. This research may enhance our ability to understand the specific vulnerabilities in social communication of children with ASD. This system can determine patterns of physiological response across participants and may identify specific social communication deficits for individual children.

This work used virtual social robots to systematically manipulate specific aspects of social communication and provides a vital step towards development of future social robots for this target population. Systematic manipulation of facial expressions, eye gaze, social distance, vocal tone, and gestures need to be studied with virtual social robots where such manipulation is easy to perform, repeatable, and highly controllable. Studies like the one presented here will provide insight on how such social robots display intentions, how they should interact, and how their interactions with children with ASD should be regulated. Investigation using virtual social robots is not only cost and time efficient, it is also necessary to understand the complexity of social tasks. Such studies will answer important questions about the design requirements and control functionalities of real-world social robots. For example, questions like what amount of eye gaze modulation abilities needs to be designed, how much vocal tone modulation is required, what facial expressions are necessary, and similar questions on social communication can be exhaustively explored using controlled studies in virtual environments with virtual social robots interacting with the target population. In that sense, this work is one of the first that presents a design platform for social robots for specific applications that is analogous to well-adopted practices in the manufacturing industry where computer-aided design inevitably precedes any manufacturing.

The implications of this work show that if a social robot is to be used to interact with a child with ASD, its amount of eye contact and distance must be adjustable. The results show that TD and ASD children significantly react to variations in eye gaze and social distance, as evidenced by affective ratings and physiological responses. Therefore, the design of social robots must allow for manipulation of these parameters. Furthermore, it would be desirable for a social robot to possess the ability to detect affective changes and respond in an affect-sensitive manner to a child to maintain an optimally positive social interaction. Establishing a supportive social interaction can lead to increased learning opportunities [4], which is an advantageous provision of effective intervention for children with ASD [38].

A limitation of this work includes the level of interaction currently possible between the user and virtual robot. The current work utilized transparent text-based menus superimposed in the corner of the VR scene to provide structured responses from which a participant could choose using a keypad. This type of interaction does not reflect natural communication, but employing VR in assistive settings relies on structured menus for communication because that is the level of communication where the technology currently stands [40, 50, 52]. In the future, an extensive database of appropriate and diversionary statements could allow for shifts in the interaction within the same or different settings and social parameter configurations. This flexibility could be essential for generalization of skills learned within an ASD intervention, because children with ASD have shown rigidity for associating learned behaviors to specific settings [52]. Also, speech recognizers along with a database of available dialogue could help move the interaction towards a natural and free-flowing exchange between user and robot. An increase in the flexibility of the interaction would prevent the autonomous functionality of the social robots from being stifled while moving the research towards its ultimate goal of developing an autonomous robot that can interact with a user for social skills improvement.

In the current research, we measured the influence the virtual robot had on the user, but the user did not influence the virtual robot. The future direction of this research into a closed-loop system would include the user having an influence on the robot's actions and then measuring how those changes in-turn influence the user. Having the robot (virtual or embodied) react to the user would require the robot to have some objective, either random response or the affect-sensitive goal of minimizing the anxiety level of the user. Based on our previous work [34], we would venture that a closed-loop interaction with that goal would be able to decrease the user's anxiety level, but this type of back-and-forth influence between robot and user has yet to be fully studied in a social setting. Creating a closed-loop system that could mitigate the effects of anxiety in the user (i.e., child with ASD) and would possess the ability to adapt to the user's anxiety level in an individualized manner will be our next level of research.

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. Future work will involve developing an expanded set of VR social interaction scenarios for exploration of aspects of social communication that may elicit affective responses. Comparing the current findings to similar experiments with an embodied robot or human is also of interest for continued analysis. For example, the social situations 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. Matarić et al. [36] contends that the role of physical embodiment of socially assistive robots remains an important yet open question in need of further study which compares embodied robots to 3D simulations and other representations of assistive technologies. We also plan to investigate the application of fast and robust learning mechanisms to permit a social robot's adaptive response within complex social interaction tasks. In the future, an autism intervention paradigm could use the system for adaptively responding to the effects of elements of social interaction that lead to struggles in social communication for children with ASD.

Acknowledgements The authors gratefully acknowledge the Autism Speaks grant for financial support of this work and Vanderbilt TRIAD for guidance during the development of the experiment protocol. The authors also thank Ann Conn and Kathy Welch for proofing this manuscript, Danny May for assistance in the recruitment of individuals for the voices of the virtual social robots, and the parents and participants who took part in this work.

References

1. American Psychiatric Association (2000) Diagnostic and statistical manual of mental disorders: DSM-IV-TR, 4th edn. American Psychiatric Association, Washington
2. Argyle M, Dean J (1965) Eye-contact, distance and affiliation. Sociometry 28(3):289–304
3. Argyle M, Cook M (1976) Gaze and mutual gaze. Cambridge University Press, Cambridge
4. Bancroft WJ (1995) Research in nonverbal communication and its relationship to pedagogy and suggestopedia. ERIC
5. Bernard-Opitz V, Sriram N, Nakhoda-Sapuan S (2001) Enhancing social problem solving in children with autism and normal children through computer-assisted instruction. J Autism Dev Disord 31(4):377–384
6. Blocher K, Picard RW (2002) Affective social quest: emotion recognition therapy for autistic children. In: Dautenhahn K, Bond AH, Canamero L, Edmonds B (eds) Socially intelligent agents: creating relationships with computers and robots. Kluwer Academic, Dordrecht
7. Bradley MM (1994) Emotional memory: a dimensional analysis. In: Van Goozen S, Van de Poll NE, Sergeant JA (eds) Emotions: essays on emotion theory. Erlbaum, Hillsdale, pp 97–134
8. CDC (2009) Prevalence of autism spectrum disorders-ADDM network, United States, 2006. MMWR Surveill Summ 58:1–20
9. Cobb SVG, Nichols S, Wilson JR, Ramsey A (1999) Virtual reality-induced symptoms and effects. Presence 8(2):169–186
10. Cohen H, Amerine-Dickens M, Smith T (2006) Early intensive behavioral treatment: replication of the UCLA model in a community setting. J Dev Behav Pediatr 27(2):145–155
11. Colburn A, Drucker S, Cohen M (2000) The role of eye-gaze in avatar-mediated conversational interfaces. In: SIGGRAPH sketches and applications
12. Constantino JN (2002) The social responsiveness scale. Western Psychological Services, Los Angeles
13. Constantino JN, Davis SA, Todd RD, Schindler MK, Gross MM, Brophy SL, Metzger LM, Shoushtari CS, Splinter R, Reich W (2003) Validation of a brief quantitative measure of autistic traits: comparison of the social responsiveness scale with the autism diagnostic interview-revised. J Autism Dev Disord 33(4):427–433
14. Corsello C, Hus V, Pickles A, Risi S, Cook EH, Leventhal BL, Lord C (2007) Between a ROC and a hard place: decision making and making decisions about using the SCQ. J Child Psychol Psychiatry 48(9):932–940
15. Dautenhahn K, Werry I (2004) Towards interactive robots in autism therapy: background motivation and challenges. Pragmat Cogn 12(1):1–35
16. Duffy BR (2003) Anthropomorphism and the social robot. Robot Auton Syst 42(3–4):170–190
17. Dunn LM, Dunn LM (1997) Peabody picture vocabulary test, 3rd edn. American Guidance Service, Circle Pines
18. Ekman P (1992) An argument for basic emotions. Cogn Emot 6:169–200
19. Ekman P (1993) Facial expression and emotion. Am Psychol 48:384–392
20. Farroni T, Csibra G, Simion F, Johnson MH (2002) Eye contact detection in humans from birth. Proc Natl Acad Sci USA 99(14)

21. Fong T, Nourbakhsh I, Dautenhahn K (2003) A survey of socially interactive robots. Robot Auton Syst 42(3–4):143–166
22. Frijda NH (1986) The emotions. Cambridge University, New York
23. Garau M, Slater M, Bee S, Sasse MA (2001) The impact of eye gaze on communication using humanoid avatars. In: SIGCHI human factors in computing systems, pp 309–316
24. Groden J, Goodwin MS, Baron MG, Groden G, Velicer WF, Lipsitt LP, Hofmann SG, Plummer B (2005) Assessing cardiovascular responses to stressors in individuals with autism spectrum disorders. Focus Autism Dev Disabil 20:244–252
25. Hall ET (1955) The anthropology of manners. Sci Am 192:84–90
26. Hall ET (1966) The hidden dimension. Doubleday & Company, New York
27. Hill E, Berthoz S, Frith U (2004) Brief report: cognitive processing of own emotions in individuals with autistic spectrum disorder and in their relatives. J Autism Dev Disord 34(2)
28. Jenkins R, Beaver JD, Calder AJ (2006) I thought you were looking at me: direction-specific aftereffects in gaze perception. Psychol Sci 17:506–513
29. Jonides J, Lewis RL, Nee DE, Lustig CA, Berman MG, Moore KS (2008) The mind and brain of short-term memory. Annu Rev Psychol 59:193–224
30. Keppel G (1991) Design and analysis: a researcher's handbook, 3rd edn. Prentice Hall, New York
31. Kozima H, Michalowski MP, Nakagawa C (2009) Keepon: a playful robot for research, therapy, and entertainment. Int J Soc Robot 1(1):3–18
32. Krikorian DH, Lee JS, Chock TM, Harms C (2000) Isn't that spatial? Distance and communication in a 2-D virtual environment. J Comput-Mediat Comm 5(4)
33. Lang PJ, Bradley MM, Cuthbert BN (1999) International affective pictures system (IAPS): technical manual and affective ratings. University of Florida, Center for Research in Psychophysiology, Gainesville
34. Liu C, Conn K, Sarkar N, Stone W (2008) Online affect detection and robot behavior adaptation for intervention of children with autism. IEEE Trans Robot 24(4):883–896
35. Liu C, Conn K, Sarkar N, Stone W (2008) Physiology-based affect recognition for computer-assisted intervention of children with autism spectrum disorder. Int J Hum-Comput St 66:662–677
36. Matarić MJ, Eriksson J, Feil-Seifer D, Winstein C (2007) Socially assistive robotics for post-stroke rehabilitation. Int J NeuroEng Rehabil 4(5)
37. Michaud F, Theberge-Turmel C (2002) Mobile robotic toys and autism. In: Dautenhahn K, Bond AH, Canamero L, Edmonds B (eds) Socially intelligent agents: creating relationships with computers and robots. Kluwer Academic, Dordrecht
38. NRC (2001) Educating children with autism. National Academy Press, Washington
39. Parsons S, Mitchell P (2002) The potential of virtual reality in social skills training for people with autistic spectrum disorders. J Intellect Disabil Res 46(Pt 5):430–443
40. Parsons S, Mitchell P, Leonard A (2004) The use and understanding of virtual environments by adolescents with autistic spectrum disorders. J Autism Dev Disord 34(4):449–466
41. Pioggia G, Igliozzi R, Ferro M, Ahluwalia A, Muratori F, Rossi DD (2005) An android for enhancing social skills and emotion recognition in people with autism. IEEE Trans Neural Syst Rehabil 13(4):507–515
42. Rani P, Liu C, Sarkar N, Vanman E (2006) An empirical study of machine learning techniques for affect recognition in human-robot interaction. Pattern Anal Appl 9:58–69
43. Robins B, Dickerson P, Stribling P, Dautenhahn K (2004) Robot mediated joint attention in children with autism: a case study in robot–human interaction. Interact Stud 5(2):161–198
44. Rogers SJ (1998) Empirically supported comprehensive treatments for young children with autism. J Clin Child Psychol 27:168–79
45. Rogers SJ (2000) Interventions that facilitate socialization in children with autism. J Autism Dev Disord 30(5):399–409
46. Rutter M, Bailey A, Berument S, Lord C, Pickles A (2003) Social communication questionnaire. Western Psychological Services, Los Angeles
47. Scassellati B (2005) Quantitative metrics of social response for autism diagnosis. In: RO-MAN, pp 585–590
48. Schneiderman MH, Ewens WL (1971) The cognitive effects of spatial invasion. Pac Sociol Rev 14(4):469–486
49. Sommer R (1962) The distance for comfortable conversations: a further study. Sociometry 25(1):111–116
50. Standen PJ, Brown DJ (2005) Virtual reality in the rehabilitation of people with intellectual disabilities: review. Cyberpsychol Behav 8(3):272–282; discussion pp 283–278
51. Stormark KM (2004) Skin conductance and heart-rate responses as indices of covert face recognition in preschool children. Infant Child Dev 13:423–433
52. Strickland D, Marcus LM, Mesibov GB, Hogan K (1996) Brief report: two case studies using virtual reality as a learning tool for autistic children. J Autism Dev Disord 26(6):651–659
53. Swettenham J (1996) Can children with autism be taught to understand false belief using computers? J Child Psychol Psychiatry Allied Discipl 37(2):157–165
54. Tartaro A, Cassell J (2007) Using virtual peer technology as an intervention for children with autism. In: Lazar J (ed) Towards universal usability: designing computer interfaces for diverse user populations. Wiley, Chichester
55. Weiss A, Wurhofer D, Tscheligi M (2009) "I love this dog"—children's emotional attachment to the robotic dog AIBO. Int J Soc Robot 1(4):243–248

Karla Conn Welch received her Ph.D. in 2009 from Vanderbilt University, Nashville, Tennessee. In 2005 she was awarded a National Science Foundation Graduate Research Fellowship. In 2010 she joined the University of Louisville as an Assistant Professor of electrical and computer engineering. Her current research interests include machine learning, adaptive response systems, and robotics.

Uttama Lahiri is a research assistant in the Robotics and Autonomous Systems Laboratory (RASL), Vanderbilt University. She received the M.Tech. degree in electrical engineering in 2004 from Indian Institute of Technology, Kharagpur, India. She is currently working toward the Ph.D. degree in mechanical engineering at Vanderbilt. From 2004 to 2007 she worked for Tata Steel, Jamshedpur in plant automation. Her current research interests include machine learning, adaptive response systems, signal processing, human-computer interaction, and robotics.

Zachary Warren received his Ph.D. in clinical psychology from the University of Miami, Miami, Florida in 2005. He joined the Division of Genetics and Developmental Pediatrics at the Medical University of South Carolina as a Postdoctoral fellow. In 2006 he joined Vanderbilt University as Director of the Parent Support and Education Program for TRIAD. Currently he is an Assistant Professor of clinical psychiatry and clinical pediatrics. His research focuses on early childhood development and intervention for children with autism spectrum disorders.

Nilanjan Sarkar received his Ph.D. from the University of Pennsylvania, Philadelphia, Pennsylvania in 1993. He joined Queen's University, Canada as a Postdoctoral fellow and then went to the University of Hawaii as an Assistant Professor. In 2000 he joined Vanderbilt University as Director of the RASL. Currently he is an Associate Professor of mechanical engineering and computer engineering. His research focuses on human–robot interaction, affective computing, dynamics, and control.