



Design of an Intelligent and Immersive System to Facilitate the Social Interaction Between Caregivers and Young Children with Autism

Guangtao Nie¹(✉), Akshith Ullal¹, Amy R. Swanson²,
Amy S. Weitauf², Zachary E. Warren², and Nilanjan Sarkar^{1,3}

¹ EECS, Vanderbilt University, Nashville, TN 37212, USA
Guangtao.nie@vanderbilt.edu

² TRIAD, Vanderbilt University Medical Center, Nashville, TN 37212, USA

³ Mechanical Engineering, Vanderbilt University, Nashville, TN 37212, USA

Abstract. Children with autism spectrum disorder (ASD) have core deficits in social interaction skills. Intelligent technological systems have been developed to help children with ASD develop their social interaction skills, like response to name (RTN), response to joint attention (RJA), initiation of joint attention (IJA) and imitation skills. Most existing systems entail human-computer interaction (HCI) or human-robot interaction (HRI), in which participants interact with the systems to elicit certain social behaviors or practice certain social skills. However, because the robot/computer being the only therapeutic factor in HRI/HCI systems, this may result in the isolation effect. Therefore, in this work, an intelligent and immersive computer system is proposed for caregivers and their young children with ASD to interact with each other and help develop social skills (RTN and IJA). In this computer assisted HHI setting, caregivers deliver social cues to participants (young children with ASD) and give a decision-making signal to the system. The system also provides different non-social cues, to help caregivers to elicit and reinforce the social behaviors of participants. By including a caregiver in the loop, we hope to ameliorate the isolation effect by creating a more real-world HHI scenario. In this paper, we will show the feasibility of the proposed system and validate its potential effectiveness by both subjective measurements and objective measurements.

Keywords: Autism spectrum disorder (ASD) ·
Computer-assisted human-human interaction · Response to name (RTN) ·
Initiation of joint attention (IJA)

1 Background

1.1 Autism Spectrum Disorder

Autism Spectrum Disorders (ASD) are a group of developmental disabilities characterized by impairments in social interaction and communication [1]. According to estimates from CDC's Autism and Developmental Disabilities Monitoring (ADDMM) Network, 1 in 59 children is believed to be identified with ASD [2]. ASD occurs in all racial, ethnic and socioeconomic groups [2]. Intelligent technological systems have

been developed to help children with ASD develop their social interaction skills, like response to name (RTN) [3], response to joint attention(RJA) [4–6], initiation of joint attention(IJA) and imitation skills [7, 8]. Early intervention for young children with ASD may differ from treatment for older children due to the developmental differences in their social relationships, cognitive and communicative processes and learning characteristics [9]. The proposed system in this paper, which is designed for very young children with ASD, focuses primarily on setting up a natural learning environment for child initiative acts, development of nonverbal intentional communicative acts and reciprocal play with social partners [9]. The two tasks in this system, response to name (RTN) and initiation of joint attention (IJA), were designed based on the developmental considerations and utilize child gaze as the fundamental measurement to influence the process of interaction.

1.2 Computer-Assisted Human-Human Interaction

Over the past several decades, computer-assisted human-human interaction has been developed to facilitate cooperative work and job efficiency. The early and influential survey of computer-supported cooperative work was conducted by R. Johansen [10], who defined and illustrated multiple approaches and applications of computer-supported cooperative work, including software such as the online meeting, screen sharing, project management and calendar management software for groups. Here we specify the HCI scheme we adopted in the proposed system as computer-assisted human-human interaction, which theoretically belongs to an example of the groupware idea of Johansen in [10].

Most existing intelligent systems for children with ASD entail HCI or HRI, in which participants interact with the systems to elicit certain social behaviors or develop certain social skills. However, due to the robot/computer being the only therapeutic factor in HRI/HCI systems, the isolation effect has been reported, where after gaining some social skills within the HRI/HCI systems, one may not be able to transfer the skills back into real world HHI [11, 12]. Therefore, the computer-assisted HHI scheme was adopted here, which incorporated caregivers in the interaction loop to help ameliorate the isolation effect.

2 System Description

2.1 System Architecture and Environment

The proposed computer assisted HHI is depicted in Fig. 1, and the system environment is shown in Fig. 2, and the system architecture is shown in Fig. 3. This system was designed based upon our existing work in [3] which utilized a closed-loop interaction protocol between participants and the system. Introducing caregivers into the system has several advantages. Caregivers can provide real social cues, such as calling a child's name. In the previous work, the system provided the social cue by playing pre-recorded audio.

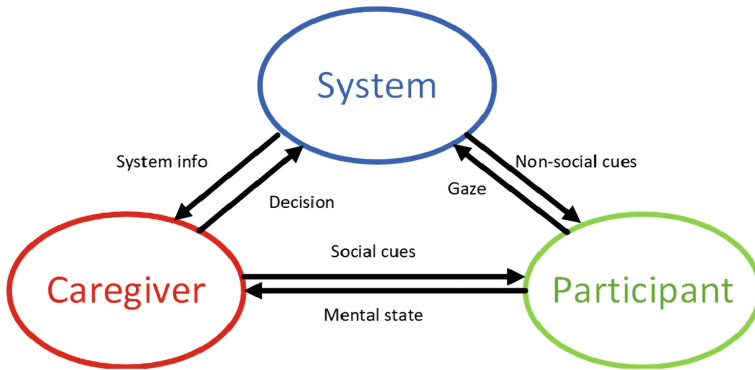


Fig. 1. Computer-assisted HHI

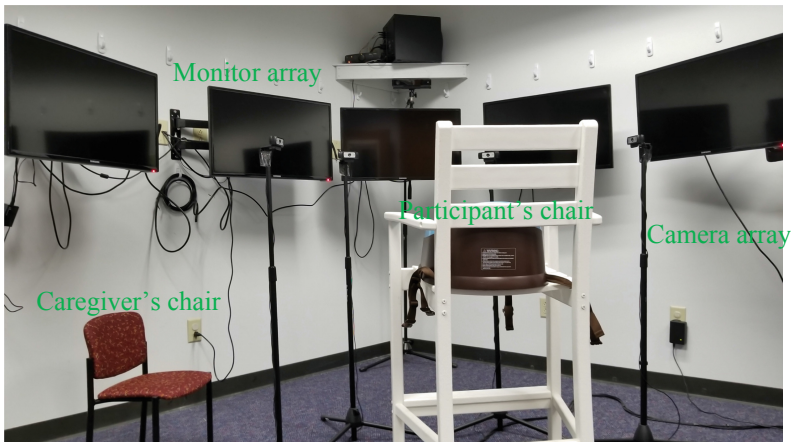


Fig. 2. System environment

We designed our system to reward children for responding to their actual caregivers, so that the elicited and reinforced social behaviors could be more easily transferred back to the real world. Our system also has caregivers use a tablet app to influence the system process based on real-time observation of participant behavior, which makes the system more adaptive and individualized. In our previous work, the system monitored nothing but the gaze of participant to do a closed-loop interaction in a fixed protocol. Also, participant mood and engagement (e.g., gestures, facial expressions) can now be considered based on caregiver input. All of these changes expand upon our previous work to create a more individualized, adaptable, and generalizable system for potential intervention.

From Fig. 2, one can see that the child sits in the center of a camera array in the shape of semi-circle with a radius of 90 cm. The caregiver sits in the small chair under the left most monitor.

The monitor array displays a video to attract and guide a participant's attention to the current target. It does this by displaying a red ball that bounces from where a participant is looking to where the current target is located to gradually transfer participant's attention and displaying reward video when a trial is completed. There is a speaker behind each monitor which forms a 5.1 surrounding sound effect.

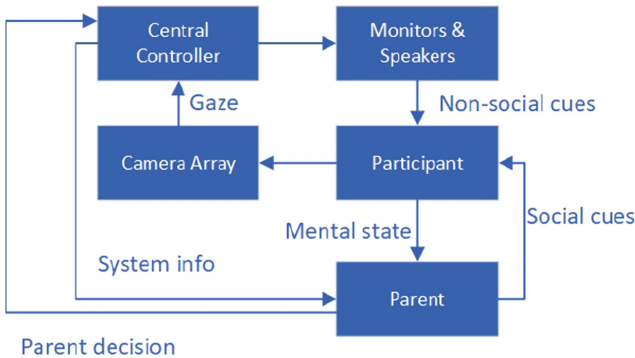


Fig. 3. System architecture

The camera array covers 180° in yaw in front of the participant to track the real-time head pose. This is used as input to the central controller to specify the starting position of the guiding ball or provide feedback to caregivers about participant engagement.

All the modules in Fig. 3 are connected through a Local Area Network with IPV4 Internet protocol.

2.2 Tablet App Design

The app for caregivers to provide input to the system process was designed and implemented using Unity with C#. As we wanted the caregiver to spend as much time as possible on observing participant's mental state and performance, the app was designed to be neutral and simple (e.g., showing only one large button to provide input; showing a question with no more than three response options available). The app had two primary functions:

1. Basic interaction process control function: Call name, Pause video, Reward. These buttons are available at different times, and only one at a time, based on the interaction stage. Details about their availabilities are described in Subject. 2.4.
2. System prompted question for decision making of non-social assistance (e.g., audio of the video and bouncing ball) and performance feedback (e.g., engagement level and task difficulty level).

2.3 Input and Output of System, Caregiver and Participant

Within our new system, the caregiver observes the participant's mental (such as emotional distress, engagement, and attention) and calls the child's name (the social

cue). The system monitors the real time gaze of participant and provides several non-social cues, including:

1. Pictures, audio and video
2. Moving objects (e.g., bouncing ball) across monitors (starting from where participant is looking at to the target monitor)

The information and options that the system provides for caregiver:

1. Task and trial information
2. Calling name/ Pausing video
3. Triggering moving objects

The system input from caregiver:

1. Feedback about participant’s mental state (e.g., engagement level, frustration level, etc.)
2. Decision to influence the system process

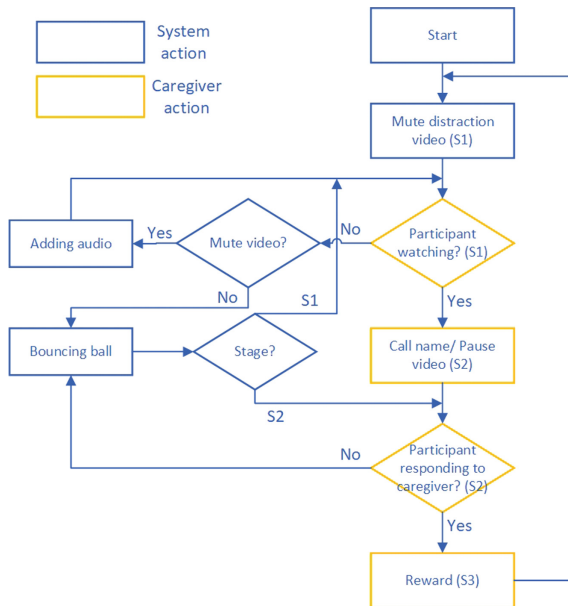


Fig. 4. Trial flowchart

2.4 Task Setup

The detailed information about the interaction scheme implemented in this system is plotted as a flowchart in Fig. 4. Two tasks are defined for this system: response to name (RTN) and initiation of joint attention (IJA). The general procedure of two tasks are described below:

RTN: 1. Play a video clip (e.g., distraction video) on a monitor to distract participant's attention away from caregiver to distraction video.

2. When caregivers press a button to confirm that participants are watching the distraction video, the app prompts the caregiver to call the participant's name.

3. If the child does look, caregivers press a button on the app to confirm that participants respond to them by looking at them. The system then plays a reward video on the monitor just over the caregiver's head and starts the next trial.

IJA: 1. Play a video clip (e.g., distraction video) on a monitor to distract participant's attention away from caregiver.

2. When caregivers press a button to confirm that participants are watching the distraction video, the system will hide the distraction video.

3. When participant responds to caregiver by looking at him/her, the system resumes playing the hidden video on the monitor over caregiver's head.

If a participant doesn't look at the distraction video or caregiver at a certain stage within 7 s, tablet app will prompt the caregiver to trigger non-social cues such as audio or bouncing ball.

3 Experiment

This study was approved by the Vanderbilt University Institutional Review Board (IRB). Caregivers (parents of participants) had all tasks explained verbally and then completed written consent documents. After formal consent, study personnel explained how the system and tablet app worked before the experiment began. This introduction script was neutral and comprehensive. After clarifying all questions from caregivers, the session started.

Each caregiver-participant pair experienced either 20 trials (10 RTN trials+ 10 IJA trials) or 20 min of interaction. Every 5 trials constituted a group. At the beginning of each group, an engaging, developmentally appropriate video clip played across the five target monitors to build participant awareness of potential targets in the system environment. After each group of 5 trials, the app prompted caregivers to give feedback about participant engagement level (not engaged, unclear or engaged) as well as perceived task difficulty level (too easy, OK or too hard) for their children.

At the very end of the experimental session, caregivers also completed a user experience survey about the system.

4 Data Analysis

Six participants (3 TD, 3 ASD) were recruited to validate the feasibility of the system. Their average age was 25.8 months (SD = 8.2). The ratio of male: female was 2:4. In this section, both subjective and objective measurements will be reported to validate the potential intervention effectiveness of the proposed system.

4.1 Objective Measurement

Response time was defined here as the time elapsed from when a caregiver called the child's name/paused the video to the time of caregiver's pressing the button and confirming that participants responded (turned and looked).

The averaged response time of RTN and IJA trials is shown in Figs. 5 and 6, respectively. Note that a shorter response time indicates a better performance.

Based on the Figs. 5 and 6, several preliminary findings are provided here:

1. For both RTN and IJA trials, TD generally performed better than ASD and the performance of both groups fluctuates with the session time going on.
2. From the pre-(Trial #1) and post-(Trial #10) comparison perspective: TD group's performance of RTN and IJA both decreased a bit. ASD group's performance of RTN decreased a lot while that of IJA increased a lot.

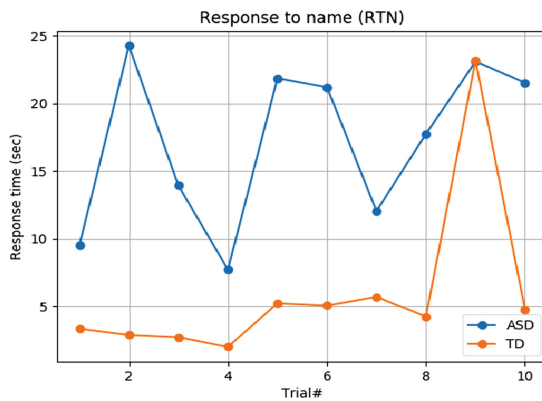


Fig. 5. Response time of RTN trials

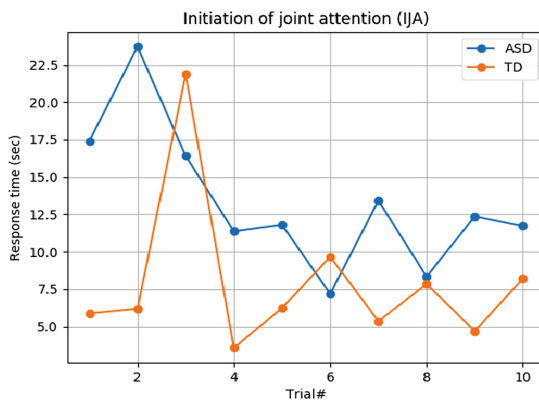


Fig. 6. Response time of IJA trials

4.2 Subjective Measurement

The caregivers' subjective feedback about the system design and user experience are shown in Figs. 7 and 8.

Here are the survey questions:

1. In general, what do you think of the design of the tablet app?
2. How much do you think the system could help you teach your child?
3. Do you like the RTN task design?
4. Do you like the IJA task design?
5. How much do you think your child liked the system?
6. In general, what do you think of the system?
7. Do you think your child's RTN skill improved through this session?
8. Do you think your child's IJA skill improved through this session?

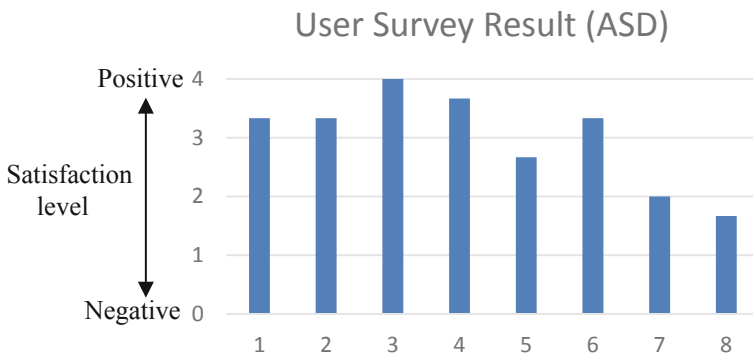


Fig. 7. User survey result (ASD)

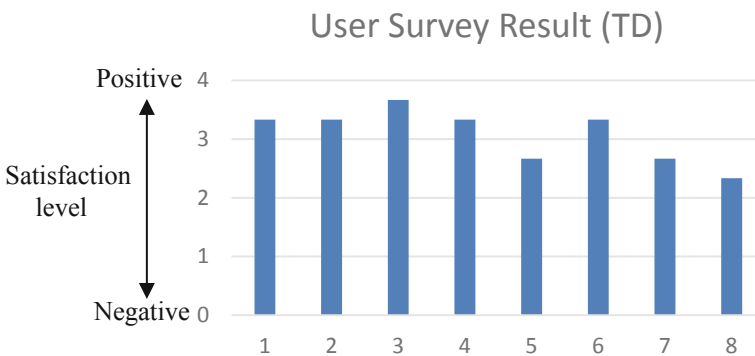


Fig. 8. User survey result (TD)

From the two user surveys above, we can see that caregivers were relatively satisfied with the tablet app design (3.33/4) and overall system design (3.33/4) and caregivers liked both task designs (RTN-3.83/4, IJA-3.5/4).

From the pre- and post- comparison perspective, a correlation test was conducted between the objective performance measurement of skill improvement and subjective performance measurement of skill improvement:

1. The correlation between RTN performance variation between Trial #1 and Trial #10 and user survey question 7: $r = -0.292$, $p = 0.574$
2. The correlation between IJA performance variation between Trial #1 and Trial #10 and user survey question 8: $r = 0.495$, $p = 0.318$.

These two correlation tests indicate that the caregiver's impression of RTN improvement is consistent with objective measurement while caregiver's impression of IJA improvement is inconsistent with objective measurement.

5 Conclusions

5.1 Achievements

We designed and implemented an intelligent and immersive system for caregiver-participant pairs to practice social interaction skills, specifically RTN and IJA. We incorporated caregivers into the system potentially to ameliorate the isolation effect which occurred in other HRI/HCI. We also provided uniform assistance for caregivers to trigger the social event and transfer participants' attention. Standardizing options for caregivers' behaviors within the system enabled us to compare different time and participant groups.

Subjective reports from caregivers showed positive results regarding system and task design. Children with TD performed better than children with ASD across both tasks, as predicted by the literature, which suggests that our system captured real differences in social responsiveness that distinguish diagnostic groups. Additionally, on IJA tasks, children with ASD showed increased performance by the end (Trial #10), compared with baseline (Trial #1), which is a promising result regarding the potential effectiveness of this system.

5.2 Limitations

The weaknesses of this paper are as follows:

First, more participants need to be recruited in the future to have a comprehensive test to validate the effectiveness of the proposed system.

Second, the intervention effect of RTN of ASD were not as promising as IJA of ASD so far, based on the objective analysis in Figs. 5 and 6. If with more caregiver-participant pairs recruited, we still cannot see a promising RTN skill improvement of ASD, we may need to modify our system or task design.

Third, the task designs lacked variation. The lack of variation could lead to participant losing social interest towards the system in a long-term interaction.

References

1. Lord, C., Cook, E.H., Leventhal, B.L., Amaral, D.G.: Autism spectrum disorders. *Autism Sci. Ment. Heal.* **28**(2), 217 (2013)
2. Baio, J., et al.: Prevalence of autism spectrum disorder among children aged 8 years—Autism and developmental disabilities monitoring network, 11 sites, United States, 2014. *MMWR Surveill. Summ.* **67**(6), 1 (2018)
3. Zheng, Z., et al.: Design of an autonomous social orienting training system (ASOTS) for young children with autism. *IEEE Trans. Neural Syst. Rehabil. Eng.* **25**(6), 668–678 (2017)
4. Zheng, Z., Zhao, H., Swanson, A.R., Weitlauf, A.S., Warren, Z.E., Sarkar, N.: Design, development, and evaluation of a noninvasive autonomous robot-mediated joint attention intervention system for young children with ASD. *IEEE Trans. Hum.-Mach. Syst.* **48**(2), 125–135 (2018)
5. Robins, B., Dickerson, P., Stribling, P., Dautenhahn, K.: Robot-mediated joint attention in children with autism: a case study in robot-human interaction. *Interact. Stud.* **5**(2), 161–198 (2004)
6. Bekele, E., Lahiri, U., Davidson, J., Warren, Z., Sarkar, N.: Development of a novel robot-mediated adaptive response system for joint attention task for children with autism. In: *ROMAN*, 2011, pp. 276–281. IEEE (2011)
7. Greczek, J., Kaszubski, E., Atrash, A., Matarić, M.: Graded cueing feedback in robot-mediated imitation practice for children with autism spectrum disorders. In: *The 23rd IEEE International Symposium on Robot and Human Interactive Communication*, 2014 *ROMAN*, pp. 561–566 (2014)
8. Zheng, Z., Das, S., Young, E.M., Swanson, A., Warren, Z., Sarkar, N.: Autonomous robot-mediated imitation learning for children with autism. In: *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2707–2712 (2014)
9. Zwaigenbaum, L., et al.: Clinical assessment and management of toddlers with suspected autism spectrum disorder: insights from studies of high-risk infants. *Pediatrics* **123**(5), 1383–1391 (2009)
10. Johansen, R.: *Groupware: Computer Support for Business Teams*. The Free Press (1988)
11. Feil-Seifer, D., Matarić, M.J.: Socially assistive robotics. *IEEE Robot. Autom. Mag.* **18**(1), 24–31 (2011)
12. Sharkey, N., Sharkey, A.: The crying shame of robot nannies: an ethical appraisal. *Interact. Stud.* **11**(2), 161–190 (2010)