

Virtual social environment for preschoolers with autism – preliminary data

C Y Trepagnier, M M Sebrechts, A Finkelmeyer, J Woodford and W Stewart Jr

Department of Psychology, The Catholic University of America,
4001 Harewood Road NE, Washington, D.C. USA

trepagnier@cua.edu

psychology.cua.edu/autism

ABSTRACT

Preliminary results are presented of a feasibility study, still in progress, of a virtual social environment designed to stimulate the social attention of pre-school-aged children with Autism Spectrum Disorder (ASD). The system uses eye-tracking and provides gaze-contingent rewards of clips from preferred videos. Of six children reported on here, most find the experimental setting appealing, and the rewards compelling; they voluntarily engage with it across numerous sessions, and demonstrate learning, with large inter-individual differences in rate of progress. Implications are discussed for the pilot study to follow.

1. INTRODUCTION

Autistic Disorder is characterized by three features, each of which can be present in various degrees: impairments in reciprocal social behavior, impairments in communication, and restricted and/or repetitive patterns of behaviors and interests (American Psychiatric Association, 2000). Autism Spectrum Disorder (ASD) includes Autistic Disorder and related disorders, all of which have in common significant impairments of reciprocal social behavior. A recent study by the Centers for Disease Control and Prevention of the United States has found that 5.5 to 5.7 per thousand children aged 4 to 17 have received a diagnosis of an Autism Spectrum Disorder (ASD) (CDC, 2006).

The value of autism-specific early educational intervention has long been considered incontrovertible (e.g., Harris & Handleman, 2000). However, despite considerable improvements from treatment for some children, there no evidence that any treatment offers a ‘cure’ for this disorder.

ASD is notable for the relative lack of engagement of the affected individual in interactions with other people (e.g., Osterling & Dawson, 1994, Leekam & Ramsden, 2006). Nonverbal communication is severely limited. Young children with ASD may not attend to others’ facial expressions or follow others’ direction of gaze, nor share enjoyment or enthusiasm with others. They attempt to direct others’ attention solely to achieve a goal, e.g., pulling mother’s hand to the doorknob to open the door, and, for some children, pointing to something they want and cannot reach by themselves.

It has been hypothesized that interruption or avoidance of social interaction during critical, experience-expectant early months of life may play a role undermining the child’s further cognitive, social, and communicative development (Trepagnier, 1996). Very young children with ASD, then, may not yet have acquired such basic social skills as following joint attention. The earlier they acquire these skills, the greater their potential to participate in nonverbal interactions of everyday life. Tuning in to nonverbal communication will in turn enable gains in language, social and cognitive development. This hypothesis has motivated the development of a virtual social environment, called Virtual Buddy, to entice very young children with ASD (ages 24 to 60 months) to participate willingly in a training program that draws their attention to the face, and provides multiple experiences of being rewarded for extracting and acting upon information transmitted nonverbally.

2. STUDY OBJECTIVES

A key feature of the Virtual Buddy system is that it be under the child's control, in order to boost motivation and thereby facilitate learning. Accordingly, training proceeds only when the child chooses to enter the training environment. The primary objective of this feasibility study is to determine the system's acceptability and appeal, over multiple sessions, to a small sample of the varied and challenging population of very young children with ASD, as well as to refine training criteria and other independent parameters. An additional objective is to identify, if possible, individual characteristics associated with progress.

The feasibility study will be completed once data from 12 children participating in up to 20 sessions each have been acquired and analyzed. Results, and on-going observation, are being used to identify needed changes. At the time of writing, 10 children have been enrolled, three of whom are currently in progress.

3. PARTICIPANTS

Results of feasibility trials are reported for six children (one female) ranging in age from 24 to 58 months at the beginning of their participation in the study. These children represented Asian (one), African-American (one), Caucasian (two) and Other (2) ethnicities, reflecting the diversity of the Washington DC area population. Preference was given to children with an older sibling. Four of the six had an older sister, one was an only child, and one had a dizygotic twin sister.

4. PROCEDURE

4.1 Instrumentation

A children's arcade helicopter was modified for the purposes of this project. The interior was gutted and equipped with a tilt-adjustable car safety seat, facing a flat screen monitor and an eye-tracking camera connected to an ISCAN RK-726I system (software version 3.58) in the adjoining room, which also houses the computers and investigator displays. The two rooms are separated by a one-way mirror. The camera and monitor in the helicopter are protected by a Plexiglass shield, and the eye-camera is concealed by darkness. Barriers prevent the child from accessing the side of the helicopter from which cables emerge. The interior of the helicopter is as free as possible of moveable objects that would lend themselves to repetitive play and distract the child from looking at the monitor. The monitor displays bids for attention by a Virtual Buddy, and gaze-contingent video rewards. The video rewards, chosen according to parent report, are individualized to each child.

4.2 Objectives

The objectives of training are to draw the child's attention to the face, and to reward the child not only for looking at the face but also for extracting information from it and other nonverbal behavior, and acting on this information. The specific behaviors addressed by the training that has been implemented are attention to the face and the eye area; and joint-attention-following. Joint-attention-following is operationalized as the child's looking in the direction indicated by the on-screen Virtual Buddy.

4.3 Calibration

The ISCAN eye-tracking system requires calibration, usually accomplished by instructing the participant to look at a particular target and then pressing a key to capture the position of the pupil and the corneal reflection (first Purkinje image) while the participant is looking at that target. The same procedure is carried out for each of 5, or each of 9 targets. Only then can the system acquire eye-tracking data.

To achieve calibration with small children whose language status and cooperativeness may be in question, a semi-automatic calibration procedure is used. Once the child has viewed a couple of minutes of the preferred video, the investigator initiates the capture of one calibration point: the screen shrinks down to a small, 6 by 4 degree area centered on the calibration point, and after 1.5 seconds the eye-tracking program captures the relative position of pupil and corneal reflection for that calibration point. The video then returns to being full-screen. This is repeated, with intervening video watching, until all five calibration points have been captured. Calibration is tested by briefly displaying a bouncing ball with a sound-effect (boing). If the point of regard (POR) veers off from the ball, one or more calibration points can be re-done, while the values for points that are good can be retained.

4.4 Training

4.4.1 Dyadic Attention. For Dyadic Attention (DA) training, a Virtual Buddy, represented by video clips of members of the research team, appears on the monitor, greets the child and offers social praise and comments. Each Buddy appearance constitutes an opportunity to produce the target behavior, gaze at face and eyes. The child's gaze behavior in response to the Buddy receives a composite score reflecting the distribution of gaze among the regions of interest (ROIs): eye area, central face (CF), Buddy and Background. The latency of gaze at the most central region looked at also figures in the score. The components are weighted according to their relative importance, and the score in response to each prompt is compared with a running average of the four previous scores. An improved score earns a reward sequence: the Buddy may say "Wow, you're terrific! Let's watch a video!", and a video ensues. Decreases in score trigger additional cues. If the child is not looking at the face at all, a face-cue is displayed: the screen is masked except for the central face area of the Buddy. If the child has received maximum score for gaze at the CF but not at the Eyes, the screen is masked except for the eye area.

Dyadic Attention training continues only until the child has demonstrated gazes of .5 seconds at the eye area.

4.4.2 Joint Attention Training. Levels in Joint Attention (JA) training are defined according to the number and salience of the directional cues the Buddy provides. The most salient cueing is provided by manual pointing accompanied by head turn, gaze at the target, and verbalization (e.g., "What's that?"). Initially (Level 1) the response is rewarded as long as the child's gaze moves toward the correct side of the screen (left or right). At Level 2, the child's gaze needs to be detected in the correct ROI. Levels 3 and 4 remove the manual pointing; and Levels 5 and 6 remove the head turn as well, so that only the gaze direction cue is displayed. Figure 1, below, illustrates the three different types of prompt. The first level of each requires only lateral accuracy of gaze.



Figure 1. Three types of joint attention prompts.

The reward for looking in the indicated direction is the appearance of a small clip of the child's preferred video, which then expands to fill the whole screen. Four target locations are distinguished, and shown as pastel-colored squares on the screen. Each visible target is located within an invisible ROI, which extends beyond it towards the central window in which the Buddy appears. Detection of the child's POR within the ROI corresponding to the correct target triggers display of the video. If the child's gaze is detected in an incorrect ROI, no video is provided. In case tracking is lost for half or more of the duration of the trial, so that the system cannot detect whether the child looked in the correct direction, a default video is provided. Clips of moving, noise-making toys serve as default videos. Figure 2 depicts three frames captured from the record of a successful response to a manual pointing prompt. In the bottom row of photographs, the change in the child's direction of gaze can be seen by the changing relationship between the two reflections in the child's eye. In the upper row of photographs the cursor representing the child's POR is shown. In the first image the cursor is on the Buddy. In the second it has moved in accordance with the direction indicated by the Buddy. In the third it has arrived at the target. The signal to play the video is sent when gaze is detected in the ROI containing the correct target.



Figure 2. Following a manual pointing cue to obtain a video reward.

5. DATA COLLECTION

5.1 Standardized Testing

Clinical diagnosis of an ASD was confirmed by means of the Autism Diagnostic Interview (ADI-R) (Lord et al., 1994), administered by the first author.

Two play-based assessments, the Communication and Symbolic Behavior Scales (CSBS), (Wetherby & Prizant, 1993), and the Mullen Scales of Early Learning (MSEL), (Mullen, 1995), were administered when possible to gain a rounded picture of participating children's developmental and communicative characteristics

5.2 Computerized Acquisition of Gaze Data

Once calibrated, the child's gaze is sampled at 60 Hz. Entry into and exit from the regions of interest defined in DA and JA training are recorded and used in scoring the child's responses, to determine whether a reward is offered and whether the next level of training should begin.

5.3 Probes for Generalization

A series of structured 'real-world' probes was developed, to be offered in the playroom during the 'break'. Probes were designed to be carried out near the beginning of training, mid-way through training, and at the final sessions. The majority of probes were delivered by one of the investigators. These included

5.3.1 Name. The investigator calls the child's name, on entering the room, when the child is facing in the opposite direction, to see whether the child will turn and look;

5.3.2 Request. The investigator tries to elicit requesting behavior by operating some toy which the child enjoys and which requires adult skills, and then waiting to see whether and how the child requests a repetition;

5.3.3 Questioning Intent. The investigator's hand is placed over the child's, while the child is manipulating a toy, blocking the child from continuing to play with it for a few seconds. The objective is to see whether the child looks at the investigator;

5.3.4 Joint Attention. The investigator attracts the child's attention or waits for the child to be looking and then offers one of the bids for joint attention described in Figure 1, to assess joint-attention-following.

5.3.5 Social Praise. The investigator praises the child (e.g., "What a handsome boy you are!") to see whether the child will look and smile.

5.3.6 *Sibling Probes*. Siblings who were able to participate were coached to carry out two simple probes. The first was to make a teddy bear or other doll carry out a repeated action, like dancing. An identical toy was given to the child with ASD, in order to assess imitation. The second probe was for the sibling to have her doll or teddy bear interact with her sibling's identical toy by hitting or kissing it. Here the objective was to see whether the child with ASD would engage in the game. Some siblings had been videotaped providing the manual pointing prompts, to serve as Buddies.

6. RESULTS

6.1 *Acceptability of Training Environment and Schedule*

6.1.1 *Children's Acceptance of the Training Environment*. Most children got into the helicopter and began training at their first session. All but one continued to do so throughout.

6.1.2 *Session Length*. Sessions lasted approximately one hour. The session always started with the helicopter. After a play break in another room, children usually returned to the helicopter for a shorter, second period. Periods in the helicopter lasted from 4.2 to 33.7 minutes ($M = 16.5$, $SD = 6.5$). The helicopter period was ended either by the investigator, or by the child. The child was then escorted to the playroom after the first period in the helicopter, to play with an investigator and participate in probes, or for administration of the play-based assessments. If the child was 'having a bad day' a second period in the helicopter was not offered. If the child insisted, however, a brief second period in the helicopter was offered.

6.1.3 *Schedule*. Families were invited to come to the lab at as intensive a schedule as they wished (up to 5 times/week). No family chose a 5 days /week schedule. The number of sessions per week ranged from one to four. The majority of families attended two sessions per week.

6.2 *Calibration*

Calibration was accomplished without any need for children to follow directions, and with minor interruptions to the video. No child showed distress at these interruptions. Because of the stability of children's position, calibration carried out prior to a play break was usable when the child returned to the helicopter in the same session.

6.3 *Training Achieved*

Participants varied in their progress through the training, ranging from no demonstration of learning to completion of all six levels.

6.4 *Probes*

A longer list of probes was originally planned. The large number of probes resulted in failure to acquire an adequate baseline for all of them because of time limitations. There was no evidence from the probes carried out of carry-over to interaction with investigators or siblings.

6.4 *Pre-training Tests*

The ADI-R was carried out for 5 of the six children. All met social and communicative criteria for Autistic Disorder, and all had age of onset before three years. None met the repetitive behavior/narrow intense interest criterion. Narrow intense interests are less often seen in pre-school-aged children.

Play-based assessments (the Mullen and the CSBS) were successfully carried out with four of the six children. These data have yet to be analyzed.

7. DISCUSSION

7.1 *Child Appeal*

Concern that some children might not choose to engage with the training environment was quickly allayed. With the sounds of a preferred, familiar video coming from the helicopter, children were usually eager to get in. An exception was one child who is not among the six reported on here. At the first session his family urged him to get into the helicopter and tried to place him in it against his will. Despite his interest in the video, he had not entered the helicopter by the end of the second session.

Part of the challenge of creating a training environment under the child's control is to make it appealing enough so that the child wants to participate, while at the same time providing rewards contingent on gaze behavior. The continued enthusiasm of five of the six children, the time they spent in the helicopter and their willingness to return to it after the play break demonstrate that the environment is appealing. Data to measure interest in the training, in terms of how often children looked at the screen when the training was going on, have not been extracted. Tracking failure because of looking away from the screen is not distinguished from tracking failure for other reasons, such as a limb in front of the camera or head leaning to one side. However investigators' impression is that after numerous sessions, some children look less often at the training video clips. Because of this observation, ways of reducing the number of sessions and numbers of trials are being considered for the Pilot Study.

Video has proven to be a compelling reward for most children. Often siblings request a turn in the helicopter, and they are given the opportunity while their ASD sibling is in the playroom.

One child used the helicopter for several sessions and then in the midst of the next session became unhappy, and would not reenter the helicopter despite two more visits. Parents are interviewed prior to beginning the study to identify anything the child dislikes or fears (e.g., for one child, barking dogs), so that video clips can be chosen which avoid arousing anxiety. Sometimes, however, the reason for onset of anxiety is not obvious to the observer.

7.2 Session Length and Time in Helicopter

Most sessions begin with the child's entry into the helicopter. As long as the child appears comfortable and engaged, he is allowed to remain in the helicopter at the investigator's discretion. In order to maintain interest and to avoid having young children sitting still for long periods without moving around, periods in the helicopter were usually limited by the investigator, especially with the younger children. To help the child accept leaving the helicopter, video curtains close, the sound is shut off, and the child is invited to come and play with toys in the playroom. Playroom sessions are usually terminated by asking whether the child would like to go back to the helicopter. Sometimes playroom sessions are cut short because the child requests return to the helicopter.

7.3 Schedule

All of the families of the children reported on here, and all but one of the families enrolled in the study so far, live outside of the District of Columbia itself, so that they are driving into the city from counties in Maryland and Virginia, in a traffic environment known for its congestion. Constraints due to distance, traffic, and the child's, sibling's and parent's schedule resulted in the need to hold regular Saturday sessions for two families. On only one occasion did a family fail to show up for a scheduled session without prior email or telephone notification.

7.4 Calibration

The use of a car seat was helpful in that it was familiar, and children were comfortable with having the safety harness attached, or attaching it themselves. The car seat was also slightly tilted back. Both these features helped to maintain the child in a stable position favorable to eye-tracking, and meant that the child returned to the same position, so that calibration did not need to be repeated. There were nevertheless some difficulties with eye-tracking. Data was lost if the child's limb was blocking the camera, whether because legs were extended to the Plexiglas covering the monitor, or because the child's arm was moved in front of the camera. Tracking was lost when the child was leaning to one side while watching the screen. While a wearable tracker might have reduced these losses, mounting of equipment on the child was avoided so that the child would feel free to come and go. In addition, it is likely that pre-school-aged children with ASD would not tolerate a head-mounted display or eye-camera, especially for multiple sessions, and even more likely that they would use it for repetitive play.

7.5 Training

The child who took a dislike to the helicopter after a few sessions was the youngest child in the study (24 months of age at his first session), and did not, in the few sessions in which he took part, demonstrate learning. Two children completed all 6 levels of JA training with few or no errors in the early levels. Their performance suggests that these skills were already established or emerging. These two children were the most advanced in language, with functional, spontaneous phrase speech. Two children demonstrated learning and progressed to Level 6. These two children used speech functionally and spontaneously, but as single-word utterances rather than multi-word phrases. The two children who made little or no progress did not use

speech spontaneously. Data from the play assessments will be examined once all twelve children have completed the study, in order to determine whether to use language levels for inclusion and exclusion criteria for participation in the Pilot Study.

7.6 Generalization

Probes used were reduced in number, over the course of these children's participation. Some of the probes retained were not carried out early and often enough to be useful in detecting change.

Parents whose children are in the study have the opportunity to monitor their child's performance and see where he is looking in response to the joint attention prompts. The investigator's display includes the child's face and expression as well as the location of the POR superimposed on the video. Some of the parents were interested in watching, while others used the opportunity to devote some one-to-one time to the child's sibling. Although parents were not instructed to transfer the skills being taught to the home, some did so. If this technique were in clinical use, it would be important to include a parent component.

8. CONCLUSIONS

The use of an inviting setting and clips from preferred video has been successful in enticing children to participate in this training program. For the most part children perceive it as a treat rather than an imposition.

Further, there is evidence, for two of the six children, that they are acquiring joint-attention-following skills, which they did not already possess. Training needs to be revised, however, to minimize the numbers of trials, so that children's interest in the training, as opposed to the video, will still be high when the more subtle indicators of direction of attention are being presented. The observations and data acquired from this feasibility study will inform changes to be made prior to carrying out a Pilot Study of efficacy.

Virtual Buddy represents a first attempt to address social skills in this population by means of a potentially stand-alone virtual environment, and the first attempt to provide an intervention that would be freely chosen by children rather than being imposed upon them. Results to date support the feasibility of this approach.

Acknowledgements: We wish to acknowledge the support of the National Institute of Mental Health (MH064924-01A2), and the volunteer families and individuals with ASD without whose participation none of this work would be possible.

9. REFERENCES

- S L Harris and J S Handleman (2000), Age and IQ at intake as predictors of placement for young children with autism: a four- to six-year follow-up. *J Autism Dev Disord*, **30**, 2, pp. 137-142.
- S R Leekam and C A Ramsden (2006), Dyadic orienting and joint attention in preschool children with autism. *J Autism Dev Disord*, **36**, 2, pp. 185-97.
- C Lord, M Rutter and A LeCouteur (1994), Autism Diagnostic Interview – Revised: A revised version of a diagnostic interview for caregivers of individuals with possible pervasive developmental disorders. *J Autism Dev Disord*, **24**, pp. 659-85.
- J Osterling, and G Dawson (1994), Early recognition of children with autism: a study of first birthday home videotapes. *J Autism Dev Disord*, **24**, 3, pp. 247-58.
- C Y Trepagnier (1996), A possible origin for the social and communicative deficits of autism. *Focus on Autism and Other Developmental Disabilities*, **11**, 3, pp. 170-182.
- A M Wetherby and B M Prizant (1993), *Communication and Symbolic Behavior Scales – Manual*, Riverside, Chicago.