Virtual Conversation Partner for Adults with Autism

Cheryl Y. Trepagnier, Ph.D., 1 Dale E. Olsen, Ph.D., 2 Laura Boteler, B.S., 2 Corinne A. Bell, M.A. 1
1 The Catholic University of America, Washington, DC
2 SIMmersion LLC, Columbia, MD

Summary

Abstract

Autistic Spectrum Disorder (ASD) is notable for severely impaired reciprocal social interaction skills relative to language and intellectual abilities, presenting a major barrier to social integration and vocational success. Evidence-based interventions to address these needs are lacking. We report on the development of a small, prototype conversation simulation to teach conversational skills to adolescents and adults with ASD and average to superior intellectual abilities, and on a test of the feasibility and acceptability of the simulation approach with a sample of the target population. The simulation engages the user in a virtual conversation with an on-screen partner, whose reactions provide naturalistic feedback geared to the appropriateness of the learner’s response choices. The prototype simulation, which provides for up to twelve potentially unique multi-turn conversations, was used over a period of two weeks by 16 adolescents and adults, who then rated statements about the system on a linear scale of 1 (disagreement) to 5 (high agreement).

The participants highly endorsed the majority of positive statements about the quality and credibility of the interaction and the virtual conversation partner. In contrast, agreement with positive statements about instructional features external to the conversation was moderate. Unexpectedly, most participants strongly agreed that using the simulation had been helpful to them. Further development and testing in the context of a controlled study with randomized assignment to control and experimental groups are needed to determine whether this approach is effective in improving real-world pragmatic language behavior of high-functioning adults with ASD.

Introduction

Autism Spectrum Disorder (ASD), a syndrome of impairments in reciprocal social interaction and language, accompanied by narrow, obsessive interests and repetitive behavior, is diagnosed in as many as one in 110 children in the United States. 1 Over 60% of this population may not be otherwise intellectually disabled. 2 ASD is a lifelong diagnosis 3,4 with marked impairment in pragmatic use of language to interact with others, even in the presence of relatively spared intellectual and language abilities. 5,6 While interactions with familiar adults may show improvement over time, peer interactions continue to be severely limited in adulthood. 7,8,9,10
ASD has long been recognized as highly heritable.\textsuperscript{11} Recent studies have implicated multiple, and different, genes, having in common that they damage neurodevelopment.\textsuperscript{12} Processing of nonverbal communication is impaired and participation in naturally-occurring social learning opportunities is severely restricted, beginning in early childhood.\textsuperscript{13,14}

Several theories of specific cognitive deficits underlying ASD social interaction impairment have stimulated research activity, including impaired theory of mind, which posits a key role for delay in the development of the ability to attribute beliefs different from one's own to another person;\textsuperscript{15,16} impairment in processing complex, non-visuo-spatial information;\textsuperscript{17} and weak central coherence, manifested as over-attention to detail, to the detriment of appreciation of relations among components.\textsuperscript{18} Neuroimaging has strongly implicated deficiencies in long-range connections within cortical regions.\textsuperscript{19} While findings from genetic, cognitive, and anatomical and functional neurological research perspectives remain to be integrated, they are largely consistent with the view of ASD that includes early deficiency in nonverbal communication leading in turn to an atypical developmental course that is particularly impoverished in social interaction.

Intervention research in ASD has focused primarily on pre-school and school-aged children. While numerous small-scale studies have been successful in teaching social skills to children with ASD, there are few large-scale studies, and limited bases for comparing among different comprehensive models of intervention.\textsuperscript{20} In practice most intervention is offered through the schools, using both evidence-based and non-evidence-based approaches informed by behavioral, developmental and/or eclectic principles.\textsuperscript{21} Studies of the effect of childhood intervention on eventual adult functioning are lacking.\textsuperscript{22}

Enhanced computer displays and virtual environments (VEs), which can offer learning opportunities without exposure to unpleasant real-world consequences of errors, are particularly attractive platforms for social skills training of adults with ASD.\textsuperscript{23,24} Recent intervention studies with the adult ASD population, using enhanced computer platforms or in vivo group intervention, are few in number and have been directed primarily at social cognitive skills hypothesized to underlie impaired social interaction. Gains in social cognition have been achieved, including higher scores on tests of theory-of-mind skills, more appropriate judgments about pragmatically appropriate behavior, and improved recognition of facial expressions of emotion, with limited generalization to untrained stimuli.\textsuperscript{25,26,27,28} Direct measurement has not, however, detected any effect of improved social cognition on subjects' real-world social interaction.\textsuperscript{28}

In marked contrast to the non-autistic population, many adults with ASD may have had little or no experience of sustained conversational interaction, whether in spontaneous settings or even in therapeutic contexts. The approach described here differs from the interventions mentioned in the previous paragraph in that it focuses primarily on participation in an approximation of the target behavior. The simulation is designed to engage learners in structured conversational interaction, with multiple opportunities to practice monitoring the conversation partner's verbal and nonverbal behavior and responding in ways that support the conversation and build rapport over multiple turns, with the goal of teaching by experience the cooperative nature of social conversational interaction.
Materials and Methods

The simulated conversation was produced using algorithms originally developed by SIMmersion LLC for teaching occupationally important social communication skills, such as culturally appropriate interviewing techniques, to government and industry personnel. In the prototype simulation developed for this study, the conversation partner, 'Sam Martin' (Figure 1), a young man in his early twenties portrayed by a professional actor, is encountered at a party given by a mutual acquaintance. The learner's task is to meet Sam, chat with him as long as a pleasant conversation can be sustained, and then end the interaction on good terms. At each turn, after the character speaks, his image remains visible until the learner speaks a response, which is input by means of a speech recognition system. The dialog options for both sides of the conversation, and directions for the actor, are written by professional script-writers who are trained in the use of the simulation algorithms. When the simulation is in operation, the system takes account of the conversation's history at each conversational turn to select a video-clip of the actor, and a set of options for the learner, to display on-screen. Figure 1 provides an example of response options available to the learner at a particular point in the conversation. The learner can opt instead to select a different conversation topic, such as school, work or ending the conversation, in which case a new set of response options is displayed. The simulated character expresses positive and negative emotions when appropriate, including mild irritation and surprise, and he gets up and leaves at the end of conversations that have gone poorly. He does not, however, engage in overtly rude behavior, to maintain the simulation as a safe environment and avoid evoking the anxiety that frequently accompanies ASD.29 The small, prototype simulation constructed for the purposes of the feasibility study includes 300 character video-clips and 125 learner choices, and is adequate to support only 10 to 12 novel conversations of about ten minutes each.

While the primary feedback to simulation users' performance is the character's spoken and non-verbal behavior, some auxiliary features are included in the VE to support motivation and provide optional explanations. There is a score which increments by one point with each conversational exchange, to reward conversation length. Response quality is separately scored: a response that is appropriate to the conversation scores one point, and a response that is not only appropriate, but also includes a relevant question or remark that helps to sustain the conversation, earns an additional point. An inappropriate response loses one point; while an egregiously poor choice (such as a critical personal comment) reduces the quality score by two points. A female coach displayed in the lower left corner of the screen (Figure 1) applauds good response choices and shakes her head sadly at responses that are especially off the mark. Help buttons can be accessed to view an explanation for the character's behavior or feedback on the appropriateness of the response chosen by the learner. Conversations can be reviewed in part or in their entirety, and help functions can also be accessed from the review screen. Instructional screens provide conversation guidelines. All explanations and instructional guidance are couched in straightforward language. As an alternative to speech-recognition, responses can be entered by mouse-click. The system then plays pre-recorded audio, so that both sides of the conversation are heard.

Training Objectives

A global set of objectives (Table 1) was developed based on recognized issues in the language pragmatics of this population30,31,6 and descriptive accounts of English conversational
The objectives were then reviewed and approved by a panel of three consultants, an autism scientist-clinician, an autism diagnostic assessment specialist, and the parent of an affected adolescent. Conversations were videotaped between 5 individuals with ASD (age range 19 to 26; 2 female), and SIMmersion staff member peers. The videotaped conversations were viewed by both conversation partners and an investigator (CT) to identify and resolve miscommunications. Pragmatics errors by one or more participants were observed corresponding to Objectives 2, 3, 4 and 5 in **Table 1**, for example, brief responses that left their conversation partner to do the work of sustaining the conversation, abrupt changes of topic, and references to information their partner did not share. In view of limits to the size of the prototype simulation, it was decided to construct the simulation to exemplify and provide experience in social conversation (Objective 2), including attending to the partner's state of mind in choosing a response (Objective 3), and cooperating to sustain a conversation (Objective 4). The simulation also addresses appropriate beginning and ending of conversations (Objective 1).

The five participants also tried out and offered comments on a demonstration simulation of a brief interaction with a middle-school-aged child, that had been constructed to illustrate how conversation simulation might be used with the autistic population. All participants found the simulation easy and engaging to use, and differed in their opinions of the help and scoring features. No barriers were revealed to proceeding with development of a prototype simulation for adults.

**Participant Recruitment and Informed Consent**

A call for participants was posted with internet groups of individuals and parents of individuals with ASD in Washington, DC and neighboring suburban Maryland and Virginia. When a parent responded, telephone screening was conducted first with the parent and then with the prospective participant, to determine transportation availability, adequacy of expressive and receptive language and access to a home computer. Appointments were booked with 20 individuals, of whom three (2 female) did not appear or reschedule.

The Protocol, Informed Consent, Assent and Call for Participants were approved by the Committee for the Protection of Human Subjects of The Catholic University of America. Informed Consent was carried out with parents of minors, and minors participated using either the Consent or the Assent form, depending on individual preference. A separate Consent requested permission to use recordings made during the study in scientific presentations. Participants had the option to permit or disallow the use of recordings of themselves for this purpose with no bearing on their participation in the study. Those who were undecided were asked to decline, and were free to give permission later if they so chose. Participants received a gift certificate for $25.00 for each of the two sessions.

**Procedure**

Sixteen adolescents and adults (1 female) completed the study. The gender distribution is consistent with the elevated proportion of males relative to females in the high-functioning ASD population. A seventeenth participant (male) attended an initial session but declined to make the return visit because of transportation difficulties. All had an ASD diagnosis by a psychiatrist.
or licensed clinical psychologist, cognitive ability within the normal or superior range (77 to 139; mean 109.4, SD 17.7) according to the two-subtest version of the Wechsler Abbreviated Scale of Intelligence (WASI), unimpaired or corrected to normal vision and unimpaired hearing. Ages ranged from 16 to 30 (mean 19.77, SD 3.54). The Autism Diagnostic Interview – Revised was completed with parents of 12 participants. Eleven of these qualified for a diagnosis of Autistic Disorder and one for Asperger Disorder.

Participants attended two sessions, two weeks apart (one participant returned in one week, because of his schedule constraints). At the initial session, a mean of 24.4 minutes (SD 5.4) was spent on introducing the simulation, using mouse click input, and participants were then asked to play two games independently. The instructional features external to the conversation were then demonstrated, training of the speech recognition system was carried out, and participants used the simulation by means of speech. Each participant was then given a copy of the program on a DVD to take home, along with a headset microphone and instructions for home installation, and was asked to play two games each week before returning for the second, final session. The program was designed for use over the internet, and data representing usage was automatically acquired under a unique code assigned to each participant. Participants for whom no at-home conversations had been logged by the end of week one received a reminder telephone call. At the second session, participants played two more games, and they were then asked to provide an independent rating of their agreement with each of seventeen printed statements about the system (Table 2) by putting a mark on a linear scale from 1 (complete disagreement) to 5 (complete agreement). Participants then met briefly with an investigator to explain their ratings, and their comments were invited.

**Hypotheses**

It was hypothesized that participants would find the simulation experience to be realistic and adequately engaging, that they would indicate interest in a fully-developed simulation able to sustain a greater number of diverse conversations, and that they would have a positive expectation of benefit from using a large simulation, as represented by agreement with statements 1-5, 7 and 16-17 (Table 2); playing more than the four requested at-home games, and using personal pronouns and verbs imputing personhood to the simulated character. In contrast, it was expected that the limited size of the prototype and the consequent restrictions on the range of response options, and number of different simulated conversations possible, would be reflected by low ratings of agreement with positive statements about the number of response options and the benefit of the simulation experience (statements 6 and 15 in Table 2).

No predictions were associated with the remaining statements, which were included for comparison with the statements of interest, as well as to obtain participants' opinions of instructional and technical features for possible design changes in the future.

**Results**

**Ratings**
Participants endorsed all 10 positive statements about the conversational experience and the character's realism (statements 1-7, and 15-17 in Table 2), and all but one of the 7 statements about auxiliary instructional features and technical features.

Voluntary Use

Usage data collected over the internet from 10 individuals showed that one played fewer than the four games requested, 3 played the requested number; and 6 played extra games. Technical errors prevented collection of data from 3 additional participants who reported playing more than four games. Three others were unable to load the program on their home computer, and one did not attempt any use of the program at home.

Performance

The higher response quality score from the two games played at the first session was compared to that of the two games at the last session. The mean was 6.56 (SD 1.63) at Session 1, and 7.56 (SD 1.79) at Session 2.

Personifying Language

Post-rating discussions between each participant and an investigator (CT) were audio recorded and transcribed. The number of participants who referred to the simulated character using personal pronouns and verbs of human action, thought or feeling, and the number of occurrences of each type of language, are shown in Table 3, along with counts omitting the discussion of statements 4 and 5, which strongly imply personhood.

Discussion

Since all of the statements to be rated were phrased positively (Table 2), a positive response bias cannot be ruled out. If only high agreement (>4) is counted, positive statements about the conversation and the simulated character are seen to be endorsed more often than statements about technical and instructional features. Seven of 10 assertions about the interaction and partner were strongly endorsed, with 3 abstentions from rating; in contrast, one of 7 assertions about extrinsic features was strongly endorsed, and one elicited disagreement, with 19 abstentions. These differences suggest that participants were expressing an authentic, relatively positive response to the conversation experience, at least as opposed to the extrinsic, instructional features.

It had been hypothesized that participants would find the number of dialog choices offered to them limiting, but the choices themselves realistic (statements 6 and 7 in Table 2). Participants did not express strong agreement with either statement, suggesting that they perceived limitations in both the number and variety of responses available to them in the small prototype system. As expected, participants strongly agreed that a larger simulation would be helpful to them and they would likely to use it (statements 16 and 17 in Table 2). Contrary to expectation, they also strongly agreed that their experience with the prototype simulation had been beneficial,
and their comments were consistent with their high agreement, e.g., "It got me very interested in wanting to keep the conversation going."

Auxiliary features were not highly endorsed, with the exception of the introductory instructional screens. Some participants reported liking the help agent (see Figure 1), while others found her distracting or superfluous and felt that they got the information they needed from the simulated partner's response. Engagement in the conversational interaction itself may have militated against paying attention to the help buttons and the coach's actions. It is also possible that the pragmatics challenges offered in the simulated conversation were not difficult enough to require recourse to explanations.

The language that most participants used in referring to the character and the interaction (Table 3) is consistent with their highly positive ratings of the credibility of the experience (Table 2), e.g., "I liked that he likes working at a pizza restaurant and he likes The Simpsons," and "...I felt like I was talking to a real person rather than some scripted thing."

Technical problems with loading the program on some home computers are reflected in participants' disagreement with statement 14, regarding ease of system installation at home. Lack of high agreement with statement 3, about enjoyment of talking to the character using speech recognition, may in part reflect the difficulties experienced by some participants in using the recognition system. Two participants abstained from rating the statement because they had not utilized speech input at home. In future work it will be important to assure adequate training of the speech recognition system to be used, and to track use of speech input and determine its role, if any, in effectiveness.

Participants' scores for quality of response choice trended to improvement from Session 1 to Session 2 but the improvement was not significant. The quality score is affected by conversation length, number of topics discussed, and dialog choices offered and taken scores, and these can vary widely. Accordingly it is difficult to interpret a two-point comparison. In a large-scale study performance will need to be measured at multiple points in order to assess progress in the training environment.

None of the participants had difficulty with the conceptual aspects of the simulation. While no direct comparison has been made with non-ASD peers, participants' willingness to engage with the system was no less than that of industry and government workers who have used training simulations to acquire occupational interactive skills. It is of particular interest that, in the brief interviews conducted at the end of the study, several participants readily imputed feelings and thoughts to the virtual character; and referred to their use of the system as conversation. Further, it was surprising that participants perceived their experience as useful and reported that it had increased their interest in talking with people. Their simulated conversational interaction may have amounted to more multi-turn sustained conversation than they had previously experienced. Some participants mentioned that the simulated conversation was less stressful than interaction in the real world, e.g., "I try to steer away from social situations but I actually felt comfortable talking to him." Participants' largely positive response to the simulation experience supports the concept of using virtual reality to provide training in social conversational skills, and suggests that the simulated conversation technique feasibility-tested in this study is an acceptably credible
and engaging method of providing this training. A randomized, controlled trial is needed, using a multi-character simulation large enough to support several hours of conversation, to determine whether this technique has efficacy for participants' real-world interaction. If simulated conversational experience can be shown to bring about improvement in interactive language pragmatics, it will be possible to provide an inexpensive, accessible and, if desired, private addition to the social skills supports greatly needed by many adults with autism spectrum disorders.

Acknowledgments

This work was partially supported by Grant No. IR43MH076307-10A2 from the National Institutes of Health of the USA, through the Small Business Innovation Research Program of the National Institute of Mental Health.

We wish to acknowledge the contributions of Dr. Deborah Fein, Ms. Susan Bacalman and Ms. Monical Adler Werner, and that of the study participants and their families.

Author Disclosure Statement

The second author is the principal and the third author is an employee of SIMMersion, LLC, the developer of the system. SIMMersion, LLC provided facilities and services for the conduct of the study.

SIMmersion, LLC provided $5,932.00 to The Catholic University of America, Washington, DC, between September 2005 and March, 2007, in partial support of the first author's work on preparation of the funding proposal.

References


5. Martin I, McDonald S. An exploration of causes of non-literal language problems in individuals with Asperger Syndrome. Journal of Autism and Other Developmental Disorders


Table 1. Conversational Learning Objectives

1. Appropriate Greeting and Parting Routines
1.1 Initiate greeting or respond to partner's greeting
1.2 Introduce self and ask partner's name or respond to question about name
1.3 Inquire about other and respond to inquiries about self
1.4 Initiate good-bye or respond to other's good-bye

2. Nature and Purpose of Social Conversation
2.1 Try to get to know, and establish rapport with, partner; expect that partner is interested in getting to know you and establishing rapport with you
2.2 Expect partner to introduce some topics; think about these topics and build on them even if they are not of great interest to you

3. Interpretation of Partner's State of Mind
3.1 Recognize from verbal and nonverbal signals (facial expression, direction of gaze &/or intonation) when partner is feeling a little negative, or very negative
3.2 Identify possible cause within the interaction and repair the interaction
3.3 Recognize when partner is merely polite, neutral/pleased or very enthusiastic
3.4 If partner is enthusiastic, validate his/her enthusiasm
3.5 If partner is at least neutral, you may continue the topic

4. Cooperative Nature of Conversation
4.1 Refrain from changing topic if it appears that partner has more to say
4.2 If introducing new topic, bridge the change
4.3 Accept partner's change in topic
4.4 Avoid repetition and avoid telling partner more than s/he needs to know
4.5 Avoid tangential and personal remarks
4.6 Use colloquial & common words rather than technical or learned words

5. Inference and Implication
5.1 Be aware of what partner probably knows, avoid allusions to things or events about which s/he doesn't know, and don't retell what s/he already knows

5.2 Keep track of what has already happened in the interaction and avoid topics in which you know partner is not interested

5.3 Expect that partner will not verbally insult you, and avoid contradicting her/him in ways that can insult

5.4 Expect that words near beginning of partner's sentence represent what s/he perceives the conversation to be about, and words at end of partner's sentence represent comment on that topic; and use word order that way

5.5 Recognize when partner is using emphasis, and understand that s/he intends to draw your attention to what s/he is emphasizing; use emphasis that way

5.6 Interpret and use pronouns to refer to things most recently mentioned or indicated, and use pronouns that way

5.7 If partner doesn't specify a topic, interpret what s/he says as related to the most recent topic mentioned or to the most evident circumstances; specify your topic unless you are commenting on the most recent topic mentioned or most evident circumstances

\(^1\) Simulation includes examples pertaining to this objective