Accessing Peer Social Interaction: Using Authorable Virtual Peer Technology as a Component of a Group Social Skills Intervention Program

ANDREA TARTARO, Furman University
JUSTINE CASSELL, Carnegie Mellon University
CORINA RATZ, The Chicago School of Professional Psychology
JENNIFER LIRA, Francis W. Parker School
VALERIA NANCLARES-NOGUÉS, Advocate Illinois Masonic Medical Center

Autism spectrum and related communication and social disorders can severely affect some children's ability to engage in peer social interaction. In this article, we describe and evaluate an Authorable Virtual Peer (AVP), technology designed to help children access peer interactions by supporting them in developing critical social skills. Children interact with the AVP in three ways: (1) engaging in face-to-face interaction with a life-sized, computer-animated child; (2) creating new social behaviors for the AVP; and (3) controlling the AVP using a graphical user interface to select appropriate responses while the AVP interacts with another person. Our evaluation suggests that when an AVP is used as an activity during a social group intervention, a common intervention approach used with children with social and communication difficulties, that children's use of specific social behaviors critical to successful social interaction increases during role-play of common social situations with another child.

Categories and Subject Descriptors: H.5.1 [**Multimedia Information Systems**]: Artificial, Augmented and Virtual Realities; K.4.2 [**Social Issues**]: Assistive Technologies for Persons with Disabilities

General Terms: Design, Human Factors

Additional Key Words and Phrases: Autism, Constructionism, Virtual Peers

ACM Reference Format:
Andrea Tartaro, Justine Cassell, Corina Ratz, Jennifer Lira, and Valeria Nanclares-Nogués. 2014. Accessing peer social interaction: Using authorable virtual peer technology as a component of a group social skills intervention program. ACM Trans. Access. Comput. 6, 1, Article 2 (December 2014), 29 pages.
DOI: http://dx.doi.org/10.1145/2700434

1. INTRODUCTION

Just as many of us take our mobility or sight for granted, children's ability to play and interact with their peers is typically taken for granted. Children readily collaborate in the classroom, engage in structured clubs for sports, art or music, and interact

This work was supported by Autism Speaks.
Authors' addresses: A. Tartaro, Furman University Computer Science Department, 3300 Poinsett Highway, Greenville, SC 29609; email: andrea.tartaro@furman.edu; J. Cassell, Carnegie Mellon University Human Computer Interaction Institute, 5000 Forbes Avenue, Pittsburgh, PA 15213; email: justine@cs.cmu.edu; C. Ratz, The Chicago School of Professional Psychology, 325 N. Wells Street, Chicago, IL 60654; email: corinaratz@thechicagoschool.edu; J. Lira, Francis W. Parker School, 330 West Webster Avenue, Chicago, IL 60614; email: jlira@fwparker.org; V. Nanclares-Nogués, Pediatric Developmental Center, Advocate Illinois Masonic Medical Center, 836 W. Wellington Avenue, Chicago, IL 60657; email: valeria.nanclares@advocatehealth.com.

DOI: http://dx.doi.org/10.1145/2700434

informally on the playground. These interactions with peers are critical to learning, developing friendships, and eventually finding a job.

However, for some children, autism spectrum and related communication and social disorders severely affect the ability to engage in peer social interaction. Given the importance of social interaction to education, relationships, and future employment, supporting successful social behaviors is valuable. In fact, mothers of children with autism spectrum disorders (ASDs) list social skills development as a priority, but report insufficient availability of opportunities to develop these skills [Little 2003].

In this research, we describe and evaluate a technology designed to help children access social interaction with peers by supporting them in developing critical social skills: an Authorable Virtual Peer (AVP). Children interact with the AVP in three ways: (1) children engage in face-to-face interaction with a virtual peer, a life-sized, computer-animated child that uses verbal and nonverbal social behaviors; (2) children create new social behaviors for the virtual peer; and (3) children control the virtual peer using a graphical user interface and select appropriate social behaviors while the virtual peer interacts with another person. Our evaluation suggests that when an AVP is used as an activity during a social group intervention, a common intervention approach used with children with social and communication difficulties, that children's use of specific social behaviors critical to successful social interaction may increase during role-play of common social situations with another child.

In what follows, we define autism spectrum disorders (ASDs) and describe how they affect social behaviors critical to peer social interaction. We then use previous research on both technological and nontechnological interventions for ASD to motivate the need for the AVP system as well as its design and use in a social skills group context. We describe the system implementation and detail the methods and analysis used in our evaluation. We provide evidence that suggests AVP interactions prior to interactions with peers may increase children's use of social behaviors aimed at engaging a conversation partner, and discuss the contributions of this study toward establishing a new intervention.

2. AUTISM SPECTRUM DISORDER (ASD)

Autism and other related disorders such as Social Communication Disorder are developmental disorders, which means symptoms begin in early childhood and affect a child's development. Autism is characterized by two main features: (1) impaired social communication (e.g., awkward around others, avoid social interactions, respond inappropriately during a conversation, or misunderstand gestures) and (2) patterns of restrictive and repetitive behaviors (e.g., focus on a particular topic or activity) [American Psychiatric Association 2013]. The severity of ASD varies; some children are nonverbal and cannot take care of their own basic needs, whereas, at the other end of the spectrum, some children are highly verbal and may have savant abilities. Since our focus is peer social interaction, our research involves children who are considered high-functioning.

Children with high-functioning ASD use language, but their pragmatics, meaning their ability to use language appropriately in a given context or with a given audience, is severely affected [Baltaxe 1977]. In other words, although there are not any specific difficulties with grammar or vocabulary characteristic of ASD, individuals with ASD have difficulty using language meaningfully [Tager-Flusberg 1994]. Wilkinson [1998] describes this characteristic of ASD language as "a dissociation of form (language structure) and function (language use)... the forms by which language is expressed remained relatively unimpaired, but the functions of language show significant disturbances" [Wilkinson 1998, p. 77]. For example, as a listener, they may not provide the kind of feedback that demonstrates they are following a conversation [Loveland et al.

1988]. As a speaker, they may not provide necessary new information others need to understand them [Dennis et al. 2001]. Thus, conversations with individuals with ASD are often less cohesive [Fine et al. 1994].

High-functioning ASD can manifest in a number of specific behaviors that contribute to awkward interactions with peers: saying odd things, avoiding eye contact, insisting on a particular way to do things, focusing exclusively on a particular topic, and using toys in repetitive, odd or unimaginative ways (e.g., spinning the wheels on a toy bus). Research confirms peer interaction is difficult for children with ASD [Barry et al. 2003; Travis and Sigman 1998]. Individuals with ASD report fewer friendships, and their ability to make friends is, as expected, negatively affected by the severity of their social skills impairment [Orsmond et al. 2004].

2.1. Reciprocity

The pragmatic difficulties experienced by children with ASD result in a number of differences in conversational social behaviors compared to other children. Children with ASD respond less often [Capps et al. 1998; Loveland et al. 1988], offer less feedback that they are following the conversation [Loveland et al. 1988], initiate topics less often [Loveland et al. 1988], and make fewer new contributions [Capps et al. 1998]. In contrast, successful interactions and conversations require contingency, making contributions that relate to and expand on the previous topic, and more specifically reciprocity—forms of contingency aimed at engaging with another person. Contingent and reciprocal contributions are essential to establishing, developing, and changing a topic of conversation, and can take a number of different forms, such as following a basic social routine (e.g., "Hi, how are you?" "I'm fine, how are you?") or rephrasing and expanding on previous conversational content. Research suggests children with ASD are less contingent than other children and use contingency differently [Tager-Flusberg and Anderson 1991]. Their responses are often off topic or otherwise noncontingent [Hale and Tager-Flusberg 2005]. Improving children's contingency could prove valuable to improving peer social interaction because successful interactions rely so heavily on maintaining reciprocal conversation. In this study, we target reciprocity, the specific forms of contingency aimed at engaging another person.

2.2. Theoretical Explanations

Researchers have developed a number of theoretical explanations to explain the underlying pattern of symptoms of ASD. These theories provide some guidance for how to develop interventions for ASD. One theory proposes that the pattern of behaviors of ASD can be explained by an underlying deficit in Theory of Mind (ToM), the ability to understand that others have thoughts and feelings (mental states) that are different from one's own thoughts [Baron-Cohen 1995; Frith et al. 1994]. Thus, some of the social interaction difficulties come from not being able to take the perspective of the other person in the conversation. Research suggests that contingency is correlated with performance on ToM tasks [Happe 1993; Tager-Flusberg 2000; Tartaro and Cassell 2008].

Another theory explains the deficits of ASD as Weak Central Coherence (WCC) [Frith 2003], which is a tendency to focus on parts over the whole. Frith [2003] uses the metaphor of a puzzle to explain WCC. While neurotypical individuals unite the pieces of a puzzle into one picture, individuals with ASD tend to focus on the individual pieces. WCC is reflected in social interactions by difficulties with understanding sentences [Frith 2003] and drawing inferences [Norbury and Bishop 2002].

A third theory describes ASD as a preference for systemizing (i.e., systematically applying rules to a logical system) over empathizing [Baron-Cohen 2002]. Baron-Cohen and colleagues have extended the ToM theory just described by incorporating ToM as one component of empathy [Sucksmith et al. 2013]. The ability to recognize that others

have their own mental states is the cognitive component of empathy [Sucksmith et al. 2013]. Empathy also includes an affective component that involves appropriate recognition and response to emotions [Sucksmith et al. 2013]. Recent work examining this affective component suggests individuals with autism have difficulties with empathy and emotion recognition [Sucksmith et al. 2013].

Theoretically derived intervention designs prescribe approaches that help children with ASD take the perspective of the other and break complex social skills into components or rules that can be systematically applied. The two aims of the current research are (1) to help users employ ToM to reflect on their partner's perspective in the conversation and (2) to utilize strengths from WCC and systemizing theory to create a systematic approach to learning the components of social interactions and how to put them together.

2.3. Social Interventions

Reichow and Volkmar [2010] conducted a meta-analysis of social skills interventions and describe eight main categories of approaches to helping children with ASD develop social skills that demonstrate some success. However, they identify only one "established evidence-based practice" (EBP) that is "shown to be effective across multiple methodologically sound studies conducted by at least two independent research groups" [Reichow et al. 2008]: social skills groups, treatment that occurs in a group context where children can learn and practice social skills with peers. In addition to empirically supported effectiveness, group settings are a popular approach to social skills instruction because they are less expensive than many other treatments because multiple children are treated at the same time. In a group setting, children have plenty of opportunities to practice social skills with the other group participants. While the format of social groups can take a number of different forms, many involve formal social skills training [Jackson et al. 1991; Krasny et al. 2003; McGinnis and Goldstein 1984, 1990; Webb et al. 2004]. Baker [2003] uses a format called Structured Learning, which balances instructional content with more natural activities that encourage the use of the skills. Structured Learning consists of four components: (1) didactic instruction on skills that breaks each skill into a series of steps, (2) modeling of the skills by therapists, (3) role-play scenarios the children perform to practice the skills, and (4) opportunities to practice skills during group activities and outside the group [Baker 2003]. These components are common in other formal social skills programs.

2.4. The Role of Interactive Technology in Social Intervention

Recently, researchers have begun to focus on the potential of interactive technology as a component of social interventions. To date, projects have employed a variety of technologies including:

—Virtual worlds and agents [Bosseler and Massaro 2003; Cole et al. 2003; DeAngelis 2009; Jarrett 2009; Kerr et al. 2002; Parsons et al. 2004; Porayska-Pomsta et al. 2012; Trepagnier et al. 2010; Wise et al. 2007]
—Robots [Dautenhahn and Weery 2004; Dautenhahn et al. 2002; Feil-Seifer et al. 2009; Feil-Seifer and Mataric 2010; Michaud and Theberge-Turmel 2002; Robins et al. 2004; Scassellati 2005a, 2005b]
—Pervasive, sensor, and mobile technologies [Cramer et al. 2011; Escobedo et al. 2012; Hayes et al. 2004; Hirano et al. 2010; Kientz et al. 2006; Kientz et al. 2007; Marcu et al. 2009; Picard 2009; Tentori and Hayes, 2010]
—Multitouch and tabletop displays [Gal et al. 2009; Hourcade et al. 2012; Piper et al. 2006]
—Eye tracking [Ramloll et al. 2004; Shic et al. 2006, 2007, 2008]

—Multimodal interaction and motion-based interaction [Bartoli et al. 2014; Hailpern et al. 2009]

Although these projects are in the early stages of design, many have demonstrated some initial success and point to a number of benefits of incorporating interactive technologies into interventions. Kientz et al. [2014] provide a comprehensive review and framework for categorizing the current state-of-the art in interactive technologies for autism. The review discusses challenges and future research for the different approaches, such as those just listed.

3. AUTHORING A VIRTUAL PEER

Although current social interventions, particularly social groups, have yielded some promising results, they have limitations. Especially relevant to this research, Barry [2003] found that while play skills improved from participation in a social group, conversation skills did not have clear improvements. Conversation skills, including initiating topics, asking questions, and responding, are vital to social interaction. In social groups, children have peers with whom they can practice these skills, yet outcomes are not ideal. Thus, additional methods are needed to target these skills.

In a review of social interventions, McConnell [2002] recommends extending treatment to other activities to address some of the limitations of current instructional approaches. AVPs apply a new activity, instructional strategy, and theoretical approach. Although approaches like structured learning [Baker 2003] offer didactic instruction, modeling of skills, and opportunities to practice skills, AVPs aim to give children the ability to construct their own understanding of skills by building those skills into a virtual peer. Children create interactions by specifying the speech and gesture behaviors for the virtual peer, organizing these behaviors into a graphical user interface, and choosing the behaviors while the virtual peer interacts with another person. These activities are based on the constructionist theory in education of learning by creating artifacts [Harel and Papert 1991].

Projects based on constructionist theory examine how building artifacts such as computer games [Kafai 1995; Steiner et al. 2006], robots [Resnick et al. 1996], or storytelling characters [Bers and Cassell 1998] help children learn about topics such as math, science, and writing. The goal is to give children a creative task where they can incorporate content that is personally meaningful while making connections to educational content [Resnick et al. 1996]. For example, Resnick et al. [1996] describe a series of "computational construction kits": systems that enable children to work with educational content such as writing and computer programming, while drawing inspiration from popular culture and personal interests such as Star Trek or learning to drive. Previous studies explore the mechanisms of learning from constructionist systems, including (1) metacognitive skills associated with a creative task [Robertson and Nicholson 2007], (2) the role of considering the audience for the artifact [Steiner el al. 2006], and (3) "decentering"—taking the perspective of others [Bers and Cassell 1998].

Systems based on constructionist theory often (though not exclusively) focus on math and science content. In contrast, our work focuses on learning language and social skills. With virtual peer authoring, our aim is to engage metalinguistic skills: metacognitive activities related to the use of language [Gombert 1992]. Planning and reflecting on the interactions of the virtual peer may promote language learning and developing social communication skills.

Interactive technologies for autism have incorporated authoring components, in particular, Bourjawah et al. [2012] developed authoring tools for parents and educators to use with their Re-flex system. Re-flex is a social problem-solving tool; children complete

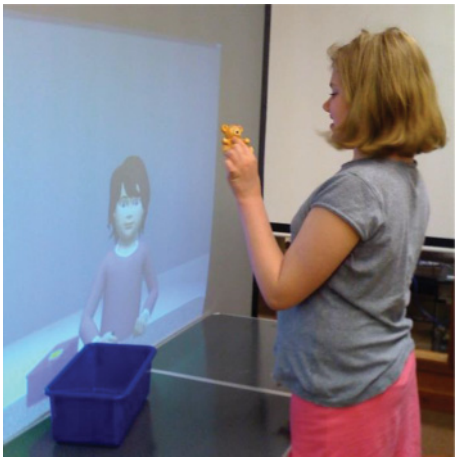

Fig. 1. Face-to-face interaction.

Fig. 2. Virtual peer.

modules that walk them through a social situation in which an obstacle arises and help them resolve the situation. The authoring tools enable parents and educators to create new modules. Although other stakeholders (e.g., parents, clinicians) can use the AVP tools, the main contribution of AVPs is that the authoring tools are designed as an educational tool for the children to use. The constructionist task of authoring the virtual peer is a key component of the intervention. Other technologies for children with autism develop collaboration skills by having children build stories [Gal et al. 2009] or music [Hourcade et al. 2012]. These examples use construction tasks that result in story and music artifacts as a context for practicing turn-taking and other cooperative skills. In our system, users are building social interaction as the artifact.

3.1. AVP Functionality

Figures 1 through 5 illustrate the functionality of the AVP system. Users engage in three types of interactions with the system. First, a user can *engage in face-to-face interaction* and tell stories with the virtual peer (Figures 1 and 2). The virtual peer is projected life-sized and interacts using speech and gesture. It is set up behind a table, and the table appears to extend in to the virtual peer's world. The child has toys or other artifacts he or she can play with while interacting with the virtual peer, and the virtual peer has her[1] own set of toys. This arrangement is designed to give a sense of "shared reality," [Cassell et al. 2000] meaning rather than the child interacting with the virtual peer in a virtual reality or virtual world as an avatar, the virtual peer becomes part of the child's world. While the user is interacting with the virtual peer, another person is observing the interaction, typically from another room via one-way glass or a video camera, and selecting from prerecorded behaviors to make the virtual peer respond. This face-to-face interaction was modeled after how children tell stories together [Cassell et al. 2000, 2007; Wang and Cassell 2003], and was evaluated with children with autism with promising results for intervention [Tartaro and Cassell 2008].

Second, a user can *operate* the virtual peer while another person interacts with her. Here, the user is observing the virtual peer's face-to-face interaction with another person and selecting the behaviors to make the virtual peer respond. The tools to

[1]The virtual peer is designed to be gender ambiguous. For simplicity, we will refer to it as a girl.

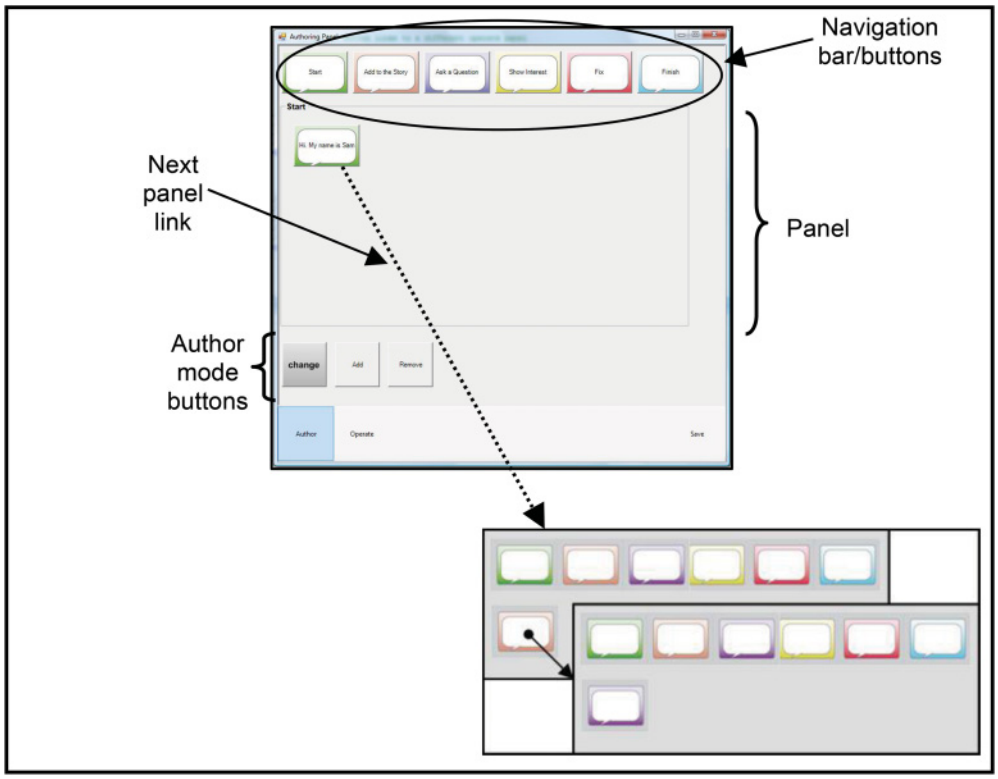

Fig. 3. Interface design: The "Operate" and "Author" interfaces are the same, except that the "Author mode buttons" only appear when the system is in "Author" mode.

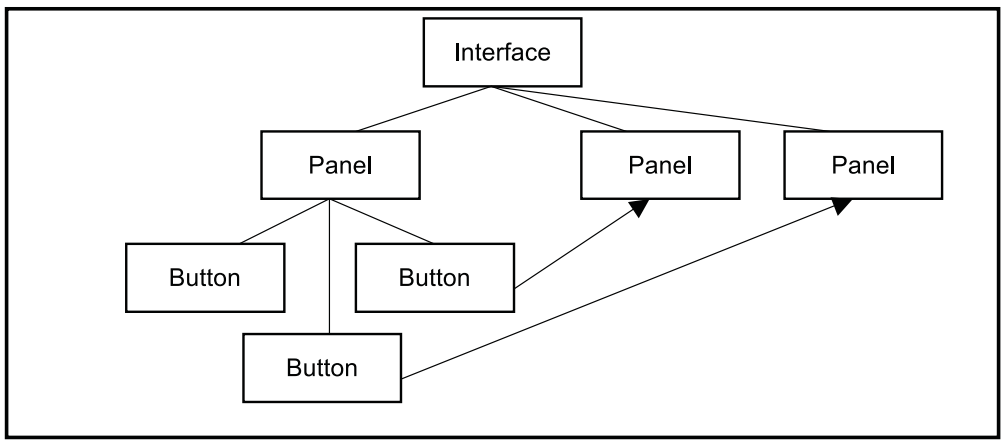

Fig. 4. Operate functionality: The interface is comprised of panels. Each panel is comprised of buttons. Each button links to another panel.

operate the virtual peer (Figures 3 and 4) are designed as a series of linked panels. At the top of the interface is a navigation bar of all the panels in the interface. Below is the collection of buttons for the selected panel. Each button, when pressed, makes the virtual peer perform a specified social behavior; for example, the virtual peer speaks

Fig. 5. Authoring.

"Hi, I'm Sam" and waves. Then, another panel of buttons appears, as specified by the selected button. In other words, each button is associated with a social behavior and acts as a link to the panel that comes next. By default, the system is comprised of six panels: (1) a *Start* panel for social behaviors to start an interaction (e.g., greetings), (2) an *Add* panel that contains social behaviors to add information to the current discourse, (3) an *Ask* panel for questions, (4) a *Show Interest* panel for social behaviors that demonstrate the virtual peer is listening or to provide positive or negative feedback, (5) a *Fix* panel with social behaviors for cases where the interaction goes awry (e.g., apologies), and (6) a *Finish* screen with social behaviors to end the interaction.

Third, a user can *author* the virtual peer by editing or adding social behaviors for the virtual peer and organizing buttons into the various panels (Figure 5). When the system is in author mode, the buttons still launch virtual peer behaviors, but instead of linking to another panel of buttons, the button is selected and can be modified or removed. If the user wants to change a button, he or she can modify any content of the button, including the audio recording (i.e., speech, by making a new recording), label text, and next linked panel. The interface provides buttons for the user to record speech through the computer's default microphone (we used the built-in microphone) and play back the recording. The previous recording for that button is replaced with the new recording. Users can also add a new button either by selecting and editing an existing button from a large library of buttons that are not currently used in any of the panels, or creating a new button from scratch. When they create a new button, they specify the label for the button, record the audio (using the same controls described above for editing button content), choose from some basic gestures like wave or point, and select the next linked panel. To select an appropriate next panel, children must think about the structure of a contingent and reciprocal conversation. The operate and author functionality were designed in previous iterative design studies described elsewhere [Tartaro 2011].

The AVP system is typically used with one child authoring, operating, or engaging in face-to-face interaction at a time. In this study, children were paired up such that they

authored and operated stories for their partner. Each child independently authored a story for the system, and the two children then took turns operating the AVP while the other child engaged in face-to-face interaction.

3.2. System Implementation

The AVP system has two main components: the interface for operating and authoring the virtual peer, implemented in C#, and the graphics animation, composed of an animation controller and renderer. While the animation system is used by several virtual peer research projects (e.g., Cassell et al. [2009]), the interface for operating and authoring the virtual peer are unique to the current research. There are three main screens to the interface: the linked panels, a screen that displays a library of possible buttons that can be added to panels, and the screen for creating and editing button functionality. Panels are described in two XML files: one is a list of the script identification numbers for all buttons that can be added to panels (these are the buttons that appear in the library); the second file describes the layout and button functionality of the various panels. The buttons launch scripts from a database that uses the Body Markup Language (BML) to describe coordinated animation and speech behaviors. BML is a standard representation language, using XML, designed to be an interface between behavior planning and realization [Kopp et al. 2006].

The graphics animation component uses SmartBody [Thiebaux et al. 2008], a framework for animating agents, and PandaBMLR (BML Realizer), the rendering engine and wrapper for SmartBody. PandaBMLR is a module for Panda3D, a 3D rendering and game development engine from Carnegie Mellon University and Disney that is free and open software [Carnegie Mellon University 2010]. The PandaBMLR component contains a database of animations for the virtual character model (created by an animator in Maya3D) and a database of audio recordings. The audio database contains not only the .wav files for the speech but also a description of the lip shapes for animating the given speech. These are generated by a utility that takes the audio file and a text file transcript of the audio and generates lip shapes and timings. When BML information is sent to PandaBMLR from the interface, body movements and audio are carefully coordinated and synchronized by SmartBody.

When a user creates a new button, the audio file is saved as a .wav file in the audio database. Users do not need to provide text specifying the speech exactly; the only text supplied is a button label. However, as described earlier, a text transcript is needed so that lip shapes can be estimated. We used speech recognition to estimate the content of the audio recording. Trying to use speech recognition on children's speech is difficult, and thus the results are imprecise. However, because we are not trying to extract meaning from the audio files, imprecise recognition is acceptable and estimates the sound well enough to generate lip shapes. The text file and lip shape descriptions are also stored in the audio database. A basic BML script is generated that launches speech and any specified gestures. This is added in the BML script database. Finally, the XML file specifying the button script identification numbers is updated. When a button is added to a panel, a temporary XML file for the interface is saved. When the "Save" button is selected, these temporary files are removed and the panel descriptions are saved as permanent files in the panel database.

4. AVP SOCIAL GROUP CURRICULUM

In this article, the AVP is evaluated as a component of a social group program. A social group is an ideal environment for evaluating the AVP because of the opportunity for frequent peer interactions and the similar curricular goals. In addition, the AVP may benefit social group programs by addressing some of their limitations. For example, the AVP targets conversational skills that Barry et al. [2003] found difficult to address.

With an AVP, children are programming a virtual character to have a conversation prior to engaging in that conversation and then selecting the responses for the virtual character during the conversation. This adds a new approach to traditional methods of teaching conversation skills (i.e., didactic instruction, modeling skills, and role-play) by giving a different perspective on the skill. It breaks the conversation down into units that the child can reflect on to build their own understanding of what is happening in the conversation in a constructivist manner. Children can use the AVP to experiment with conversation in a way that leverages their ability to systemize [Baron-Cohen 2002], encourages them to take the perspective of others, and allows them to reflect on and revise conversations.

Thus, the motivation for integrating an AVP into a social skills group is strong. While a number of technologies have been used or evaluated in a group context (see Tartaro and Ratz [2013] for a review and discussion), the technology is often a supplement to the program. In contrast, we developed an integrated social skills curriculum incorporating the AVP as a key component of the program. The curriculum development team consisted of two therapists at a local therapeutic center for children with developmental disabilities (the third and fourth authors), who then led the group program, and a researcher (the first author, at the time a Computer Science and Communication Studies Ph.D. candidate with an M.A. Instructional Technology). We used the Structured Learning program described by Baker [2003] in his book, *Social Skills Training for Children and Adolescents with Asperger Syndrome and Social-Communication Problems*, to design the group. As described earlier, Baker's program uses both didactic lessons and modeling to tell and show participants about a social skill, followed by role-play as well as other activities to encourage children to use the skills.

Baker's book outlines 70 specific social skills lessons. The lessons include a list of explicit steps for each social skill, suggested activities and possible role-play scenarios. The "Taking Turns (the Two-question rule)" lesson is one example:

1. When others greet you, greet them back
   a. If they say "Hello," then say "Hello" back to them"
2. Two-question rule: When others ask you a question and you answer it, you can ask a similar question right back. Example: "How was your weekend?" "Good. I went to the movies. How was your weekend?" [Baker 2003]

One of the suggested role-play scenarios for this lesson is, "Pick topics about which people have different preferences (movies, TV show, food, school subjects, etc.). Prompt each student to take turns sharing their preferences about each topic" [Baker 2003].

Based on questionnaires on children's needs administered to parents of our participants, we picked nine lessons from Baker's curriculum:

1. Getting to know someone new (Lesson 15)
2. Taking turns talking, the two-question rule (Lesson 8)
3. How and when to interrupt (Lesson 5)
4. Having a respectful attitude (Lesson 70)
5. Compromising (Lesson 26)
6. Dealing with making a mistake (Lesson 55)
7. Dealing with losing (Lesson 30)
8. Dealing with teasing (Lesson 64)
9. Dealing with peer pressure (Lesson 44)

We developed a 2-hour weekly program for a social group that included the AVP as an additional activity within the Structured Learning program. Each week focused on one of the nine lessons. The general structure of the group included: 20 minutes for

introduction and skill instruction; 90 minutes for group activity, AVP, and role-play rotations; and 10 minutes to wrap up. The social skills group met for 11 weeks. The first session did not include a skill lesson or the AVP and was used to get participants acquainted with each other, the staff (the first three authors and an undergraduate research assistant), and the rules and structure of the group. Similarly, the last session was used to wrap up the group.

The clinical therapists led the general instruction of the lessons. Based on Baker's prescription for teaching social skills, the instruction time included didactic instruction on specific skill steps and modeling of the skill. The therapists also discussed practicing the skill in and outside of the group by encouraging children to use the skill during the group activities for the rest of the session and giving children handouts about the skills to take home and practice.

During the rotations, all children participated in the group activity, led by one of the therapists. An undergraduate research assistant helped the therapist during the group activity. All children also left the group in dyads for 10 minutes to participate in a role-play activity, which was facilitated by the other therapist. In addition, some children left the group individually and in dyads to work with the AVP prior to role-play, which was facilitated by the experimenter.

Successfully using the social skills described by the nine lessons listed earlier requires a number of aspects of reciprocity that can be targeted by the AVP intervention. Table I summarizes how we decomposed each lesson into key components. First, we list the specific components outlined by the didactic lesson for that skill (*lesson components*). Then, we break down those lesson components into their related *reciprocity components*.

5. AVP EVALUATION

The AVP adds a new approach to the Structured Learning program. After receiving instructions on a social skill and viewing a model of the social skill, children "construct" that skill by authoring the social behaviors of a virtual human. However, we do not know if the AVP interactions will affect behaviors during face-to-face interactions with peers. Thus, this study asks:

Do AVP interactions help children with ASD employ reciprocal social interaction skills during interactions with their peers?

5.1. Hypotheses

We thought this constructionist task, when completed prior to the role-play task and interspersed with unstructured group activities, would increase children's use of reciprocity related to the skills. Therefore, we hypothesized:

H1. After performing the virtual peer task, children with ASD will use more appropriate reciprocity skills (the lesson components and reciprocity components in Table I) in the dyad role-play task than without prior AVP interactions.

Many of the lessons use similar reciprocity components, and the AVP activities are designed to carry some skills over from previous weeks. In addition, constructionist learning theory suggests the construction task should help children develop their own understanding of a skill in a way that may improve their learning of the skill over time. Thus, we hypothesized:

H2. Over multiple sessions in the group, appropriate use of reciprocity skills (the lesson components and reciprocity components in Table I) will increase in the dyad role-play task.

Table I. Decomposition of Lessons into Reciprocity Components

Lesson	Lesson Components	Reciprocity Components
Getting to know someone new	• Ask a question about something in the present moment or that you know you share an interest in • Introduce yourself • Ask questions about the other person	Ask Respond Share information Introduce oneself Greet
Taking turns talking (2 question rule)	• Asking a similar question back after being asked a question	Ask Respond Share information
How and when to interrupt	• Wait for a pause in a conversation others are having to interrupt/join • Say excuse me or sorry to interrupt • Ask a question about something relevant in others' conversation	Interrupt Ask
Having a respectful attitude	• Use request words and tone rather than demand words • Compromise or accommodate others • Ask permission to touch something not yours • Talk about feelings when upset	Request Compromise[2]
Compromising	• Ask what another person wants • Tell what you want • Offer to do some of what you and another person want	Ask Share information Compromise
Dealing with making a mistake	• Ask for help • Apologize • Try again	Ask Apologize
Dealing with losing	• Give positive feedback (e.g., Congratulations) • Acknowledge losing is ok	Give feedback
Dealing with teasing	• Tell someone to stop • Ignore teasing • Talk back to teasing • Compliment someone teasing • Act as if teasing is a joke • Tell an adult • Don't care	Respond
Dealing with peer pressure	• Saying yes and no appropriately • Explaining why • Identifying good and bad peer pressure	Respond Share

5.2. Study Design

We used a within-subjects, counter-balanced design to compare the effects of using the AVP versus not using the AVP on appropriate use of reciprocity skills during role play scenarios, and employed repeated measures to evaluate appropriate use of reciprocity skills over time. Specifically, all children used the AVP prior to participating in role-play scenarios for three to five consecutive sessions. So children could use the AVP for consecutive sessions, half the children used the AVP during lessons 1 through 5 of the interventions, and the other half of the children used the AVP during lessons 6 through 9 of the intervention. As described earlier, each week children participated in a dyad role-play task. Each child was matched with a partner whom they primarily worked

[2]By compromising, children are being reciprocal by combining some of what they want with some of what their partner wants.

Table II. Participant Characteristics

Participant	Age	Gender	TONI quotient	PPVT standardized	PPVT age equivalent	SRST
1	10:0	M	125	95	9:6	67
2	9:3	M	121	127	13:7	85
3	12:0	F	91	71	7:6	90
4	11:5	M	85	91	9:11	68
5	8:11	F	84	87	7:8	75
6	8:11	M	85	79	6:8	83
7	10:2	M	87	72	6:7	70

with for both the role-play task and the AVP task. Occasional substitutions were made because of absences. We analyzed these interactions for appropriate use of reciprocity skills as the dependent measure.

5.3. Participants

We recruited eight children from among clients of the developmental disorders clinic; seven of the children completed the intervention and were included in the analysis. We administered the Social Responsiveness Scale (SRS) to parents both pre- and post-intervention. The SRS is particularly well suited for this research because of its focus on reciprocity skills. It is a measure of the severity of autistic social impairment that yields an overall standardized (T) score as well as scores on five treatment subscales: social awareness, social cognition, social communication, social motivation, and autistic mannerism. All pre-test T scores were clinically significant: in the moderate range (60–75) for 2 children and severe range (above 75) for 5 children. Scores in the moderate range "indicate deficiencies in reciprocal behavior that are clinically significant and are resulting in mild to moderate interference in everyday social interactions" [Constantino 2005]. Scores in the severe range are "strongly associated with a clinical diagnosis of" ASDs and "suggest a severe interference in everyday social interactions" [Constantino 2005].

In addition to the SRS, we administered the following standardized measures to the participants:

—Test of Nonverbal Intelligence-3 (TONI-3). The TONI is a brief screening of nonverbal intelligence that is administered without the use of any language. We used the TONI to screen participants for a nonverbal intelligence quotient (IQ) of 75 or above.[3]
—Peabody Picture Vocabulary Test-IV (PPVT-IV). The PPVT is a language ability scale that measures children's receptive language. Children's standardized scores ranged from 71 to 127,[4] with age equivalents ranging from 6 years, 7 months to 13 years, 7 months.

Table II summarizes the characteristics of the children included in the analysis.

5.4. Treatment Task: AVP Interactions

We designed the children's first session with the AVP to primarily introduce them to the various functionality of the system. During the first session:

1. Children engaged in a short face-to-face interaction with the virtual peer in which the virtual peer asked about an interest. An experimenter selected prerecorded

[3]TONI mean standardized IQ = 100; standard deviation = 15; higher scores reflect higher nonverbal intelligence.
[4]PPVT-IV mean standardized score = 100; standard deviation = 15; higher scores reflect higher receptive language ability.

responses for the virtual peer. These responses consisted of basic greetings, questions, and information to share such as, "Hi, my name is Sam," "What video games do you like to play?", and "I like to play on my Wii."
2. The experimenter introduced children to operating the AVP. After children familiarized themselves with the interface, they selected responses for the virtual peer while the experimenter interacted with the virtual peer.
3. Children learned the concept of authoring by re-recording an existing utterance so it was in their own voice.
4. Children learned to create new buttons for the interface. They made one to three new buttons (ask, add, and/or feedback) related to the lesson of the day.
5. Children used their interface to operate the AVP for their peer partner. If a child was the second person in her dyad to work with the AVP, then her partner interacted face-to-face with the virtual peer immediately following authoring. If a child was the first person in her dyad to work with the AVP, she returned to the group activities while her partner was introduced to how the VP worked, and then operated the virtual peer after her partner completed all the AVP tasks.

Each subsequent week, children built on their interface from the previous week. They could change or remove anything that was already on the interface. Then they added a new button to the interface, and thus something new for the virtual peer to say, that acted as the prompt for the week. This button had a specific script for the child to follow. We designed each week's button to work with the topics from the previous week, mirror the topics the children would cover during the role-play task, and enable the virtual peer to direct the conversation to the topic of the week. For example, the prompt for week 2 (Taking turns talking) was: "Let's say we are meeting for the first time and we start talking about our favorite video games."

After they added this button, children added additional "ask" and "add" buttons to continue the interaction prompted by the new button. We encouraged children to employ the skills they had just learned about. For example, week 3 (How and when to interrupt), the experimenter told them, "Sam can say this to your partner, and then you can make a story for Sam to describe how to interrupt and join a conversation, just like you learned from [the therapist]."

5.5. Dependent Measures Task: Dyad Role-Play

Each session, all children worked with their partner for 10 minutes on a series of role-play scenarios related to the topic of the week. One of the therapists facilitated the role-plays. This task served two purposes. First, it fulfills step 3 of the structured learning technique (role-play). Second, it acts as a measure of transfer: Can interactions with the AVP increase reciprocal behaviors of children with ASD when they interact with a peer? Several studies of the effects of social skills groups use role-play or peer interaction tasks to evaluate behavioral outcomes. For example, Ruble [2008] used a pre-/post- test design to measure conversational outcomes. In an evaluation by Barry et al. [2003], children participated in two 5-minute play sessions with typically developing peers at the end of each treatment session. Kroeger [2007] used a pre-/post- test design to evaluate initiation, response, and interaction behaviors in a 30-minute unstructured play session.

We selected role-play scenarios from those recommended by Baker [2003] for each social skill lesson. All participants completed the first prompt, which is most closely related to the AVP activities. Then they completed as many other prompts as they could in their 10-minute time slot. We designed the prompts to use both the current lesson as well as skills learned in the previous weeks as much as possible. For example, the prompts for week 1 (getting to know someone new) included: "Introduce yourself

to your partner and get to know them by talking about:" (1) "your interests;" (2) "your school;" (3) "your family;" and (4) "your neighborhood" (each category was given as a separate prompt with the same introduction).

5.6. Additional Dependent Measures

In their evaluation of a group social skills intervention program for children with ASD, Ozonoff and Miller [1995] emphasize using two levels of evaluative measures. First, "specifying measures" determine whether a specific target behavior has changed. In this study, the target behaviors are the reciprocity components and lesson components outlined in Table I. "Impact measures" determine if the intervention influences everyday life; for example, using more reciprocity when talking to friends and family. Typically, research uses parent and teacher reports that are either standardized measures or specially designed for the study. However, these measures are rarely sensitive to the changes that occur within a short-term intervention. In several studies, including the study conducted by Ozonoff and Miller [1995], changes on the scales were not significant [Barry et al. 2003; Howlin and Yates 1999; Ozonoff and Miller 1995; Webb et al. 2004].

Despite the insensitivity of impact measures, we included two pre-/post- rating scales as dependent measures. The pre-/post- measures evaluate the social group as a whole, not the AVP on its own. As mentioned earlier, we administered the SRS to parents both pre and post treatment. Based on evaluative methods recommended by Baker [2003], parents also rated the nine target skills, combined with seven other randomly selected skills, on a 5-point scale (1 = almost never uses the skill; 2 = seldom uses the skill; 3 = sometimes uses the skill; 4 = often uses the skill; 5 = almost always uses the skill) at the first session and the last session.

6. ANALYSIS

6.1. Role-Play Task

The analysis of the role-play task focused on reciprocal social interaction skills children used with their peers and what factors affect appropriate reciprocity. While many evaluations of social interventions use some form of peer interaction as part of their dependent measures, such as role-play [Barry et al. 2003; Ruble et al. 2008; Webb et al. 2004] or free-play sessions [Barry et al. 2003; Kamps et al. 1992; Kroeger et al. 2007], researchers use a wide range of methods to analyze these interactions. Each study uses a coding scheme specific to the skills targeted in their program and the teaching strategies used. Ruble [2008] evaluated the presence or absence of four specific skills in the conversation (getting a person's attention, asking a question or making a statement, listening to a response, and ending the conversation). Barry et al. [2003] assessed conversations for presence or absence (or not applicable) of 43 skills specific to their study, including greeting, conversation, and play skills. Kamps et al. [1992] coded initiations, responses, duration of social interactions and the frequency of 21 specific social behaviors. Other studies rate the quality of skills. Kroeger et al. [2007] identified initiations, responses, and interaction events in free-play sessions and coded each as positive or negative. Webb et al. [2004] rated the facial expression, sound quality, and eye contact used by participants while performing the skills targeted by their intervention on a 3-point scale. Clearly there is a wide range of assessments used, all tailored to the program's specific instructional strategies, targeted skills, and participant abilities.

The coding scheme for the current study draws from these previous evaluations, studies of conversational skills, and the decomposition of specific lessons targeted by

Table III. Criteria for Identifying Appropriate and Inappropriate Utterances

Appropriate	Inappropriate
• Comprehensible • Socially appropriate • Contains shared referents (evident from prior discourse or introduced within the utterance)	• Unclear referents • Uninterpretable • Confirmable confabulations • Introduces a new topic immediately following a topic introduction • Nonreciprocal (not directed at the other child) • Taboo/impolite topics • Socially inappropriate or awkwardly worded • Indiosyncratic • Repetitive • Noncontingent • Incorrect based on lesson components: ○ Interrupting at an inappropriate time ○ Interrupting with an inappropriate contribution ○ Evaluating good or bad peer pressure incorrectly ○ Saying yes or no inappropriately to peer pressure

the current intervention (Table I). In addition, because human behavior is typically too complex to predict all behaviors that will occur during a specific task, we further refined the annotation scheme using a more grounded approach derived from observation of task performance [Weingart 1997].

We first transcribed videotapes of the role-play sessions and segmented the data into utterances, where an utterance is an independent clause and its modifiers. We identified specific instances (one or more utterances by one participant) of the lesson components (second column in Table I). For example, for week 5's lesson (compromising), we identified instances where a child offered to do some of what he wants and some of what the other wants. We also coded each utterance for reciprocity components (column 3 in Table I): "greet," "ask" (a question or for help), "respond," "share," "introduce," "request," "compromise," "apologize," and "give feedback." We then labeled utterances as appropriate or inappropriate. Following Briton et al. [1997], appropriate utterances are comprehensible, are socially appropriate, and contain referents that are shared (evident from the prior discourse or introduced within the utterance). Inappropriate utterances, also described by Briton et al. [1997], have unclear referents, are uninterpretable, are confirmable confabulations, or introduce a new topic immediately following a topic introduction by the other interlocutor with no acknowledgement of the other topic. Based on a grounded approach [Weingart 1997], to this list we added utterances that are nonreciprocal (not directed towards the other child), taboo/impolite based on the topic, socially inappropriate, awkwardly worded, idiosyncratic, repetitive, or noncontingent. Finally, we designated some utterances inappropriate based on specific components of the lessons: interrupting at an inappropriate time or with an inappropriate contribution (how and when to interrupt), incorrectly evaluating if peer pressure is good or bad (dealing with peer pressure), and responding inappropriately to peer pressure (dealing with peer pressure). These criteria are summarized in Table III. We excluded utterances directed at the therapist/role-play facilitator unless they were part of the role-play task. Two annotators coded 20% of the data and we calculated inter-rater reliability using Cohen's Kappa: for lesson components $K = 0.71$, for reciprocity components $K = 0.79$, and for appropriate versus inappropriate $K = 0.77$.

We analyzed reciprocity components and lesson components using least-squares regression models to see if use of the AVP, the number of weeks in the program, or language ability were statistically significant factors for appropriate use.

Table IV. Participant's Use of AVP by Lesson

	Participant						
Lessons	1	2	3	4	5	6	7
1. Getting to know someone new		AVP					AVP
2. Taking turns talking	AVP	AVP				AVP	AVP
3. How and when to interrupt	AVP	AVP				AVP	
4. Having a respectful attitude	AVP	AVP				AVP	AVP
5. Compromising	AVP	AVP				AVP	AVP
6. Dealing with making mistakes			AVP	AVP	AVP		
7. Dealing with losing			AVP	AVP	AVP		
8. Dealing with teasing			AVP	AVP	AVP		
9. Dealing with peer pressure				AVP	AVP		
Total AVP Interactions	4	5	3	4	4	4	4

6.2. Pre-/Posttest Measures

We used matched pairs Wilcoxon signed rank test with continuity correction to analyze pre-/post- data from the SRS. Since different subscales have differing relevance to the actual skills targeted by the intervention, we made comparisons for the total score as well as each of the five subtest scores. For example, the intervention focuses heavily on social communication but does not directly target autistic mannerisms.

We used standard least squares regressions to test for significant pre-/post- test differences on the parent skills ratings for those skills not covered by the intervention (or for which the child was absent), those skills covered by the interaction but where the child did not interact with the AVP, and those skills for which the child did interact with the AVP. Participant was included as a random effect.

7. RESULTS

7.1. Using the AVP

Most children used the AVP system four times. Due to absences, one child used the system three times, while another child used the system five times (to substitute for an absent child). Table IV shows the lessons for which each child used the AVP system.

Anecdotal comments from the children and their parents throughout the course of the program suggest the AVP was an engaging and relevant component of the program. Children frequently requested their turn "on the computer" and mentioned "Sam," the virtual peer, on occasion during their role-play interactions. For example, one child said, "Did you say to her a [favorite] TV show?" Since children only used the system 3–5 times, engagement did not appear to diminish over the course of the program.

Table V illustrates the kinds of things children authored for the AVP. The table includes examples from each lesson and different types of panels (i.e., Add, Ask, Show Interest). We chose these examples to illustrate some of the ways children applied the lessons they were learning to their AVP authoring.

7.2. Reciprocity Components

To analyze significant factors in the appropriate use of reciprocity components, we fit a standard least squares regression model for the rate of appropriate reciprocity components each week by participant (56 observations). The observations in this model include all seven participants—each observation is the rate of appropriate use of reciprocity components for one participant during one role-play session. We included the following factors in the model ($R^2 = 0.622$):

—Week (i.e., time—how far along in the program)
—Whether the child used the AVP that week

Table V. Examples of Authored Turns for the AVP

Lesson	Panel	Participant	Example (recorded speech)
Getting to know someone new	Ask	7	What's your name?
	Add	2	I have an Xbox 360
Taking turns talking (Two-question rule)	Ask	2	What is your favorite TV show?
	Add	7	My favorite TV show is Frasier
	Show Interest	1	That's cool
How and when to interrupt	Ask	2	What sports do you like?
	Add	6	She started to interrupt
Having a respectful attitude	Ask	7	That game sounds fun, do you want to play again?
	Add	2	That was sort of rude, I'm sorry. Can we play *Sorry*?
	Show Interest	1	That's awesome.
Compromising	Ask	1	Do you really want to play this game? Or do you want to change your mind and play another game?
	Add	7	I want to play a different game. How about *Guess Where*? It's a good game.
	Show Interest	2	By the name it might tell me that it will be fun.
Dealing with making a mistake	Ask	5	Guess my favorite game.
	Add	3	Your favorite game is Barbie.
Dealing with losing	Ask	5	When did you lose?
	Add	4	Try to win next time.
Dealing with teasing	Ask	5	When were you teased?
	Add	3	I should tell them to stop.
Dealing with peer pressure	Add	4	Let's do our homework and play a game.

—Language ability measured as the child's age equivalent on the PPVT
—Week and AVP interaction
—Week and language interaction
—AVP and language interaction
—Week and language and AVP interaction
—Speaker as a random effect to control for multiple points per participant

We found use of the AVP ($p < 0.003$) significantly predicted appropriate use of reciprocity components, with a higher rate of appropriate responses predicted if the child interacted with the AVP prior to engaging in the role play. Week ($p < 0.002$) also significantly predicted appropriate use of reciprocity components, with a higher rate of appropriate responses predicted during later weeks of the intervention program. In other words, over the course of the intervention, children's use of appropriate responses increased. We also found an interaction between language and week ($p < 0.01$), such that language ability negatively impacted the increase in appropriate responses over the course of the intervention.

Figure 6 illustrates the group's mean appropriate use of reciprocity components during role play when they first interacted with the AVP versus did not interact with the AVP over the weeks of the intervention. The vertical dotted line marks week 6, when

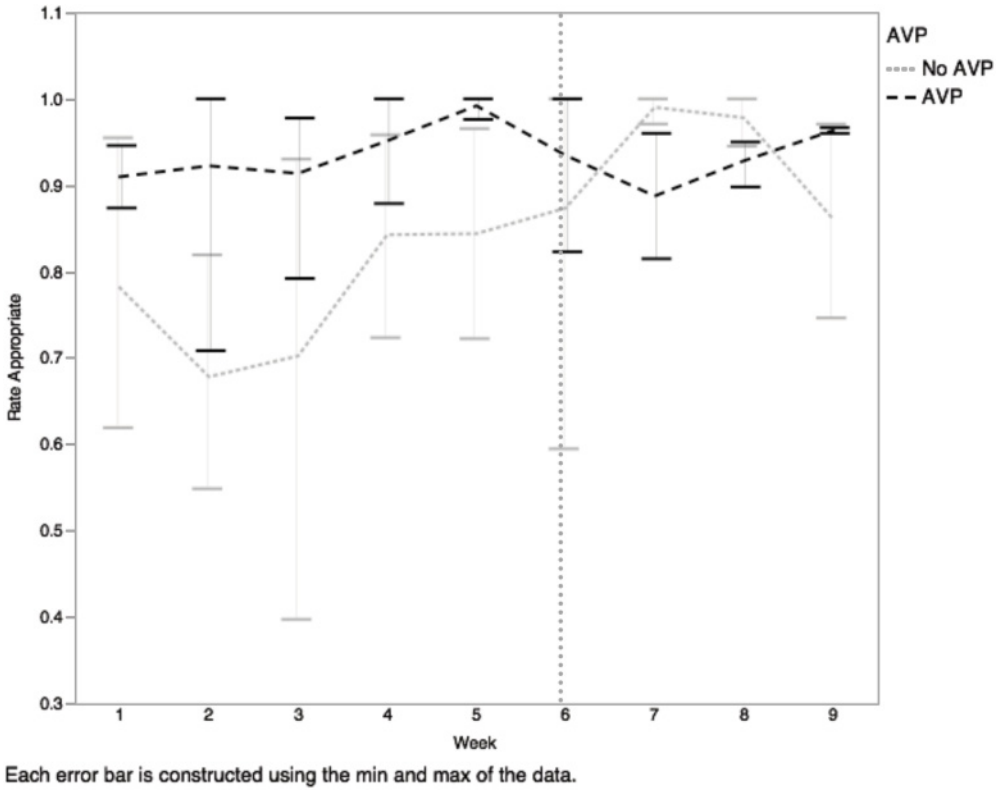

Fig. 6. Mean rate of appropriate use of reciprocity components over weeks when children first interacted with the AVP and did not interact with the AVP. The vertical dotted line at week 6 is when the groups using the AVP changed.

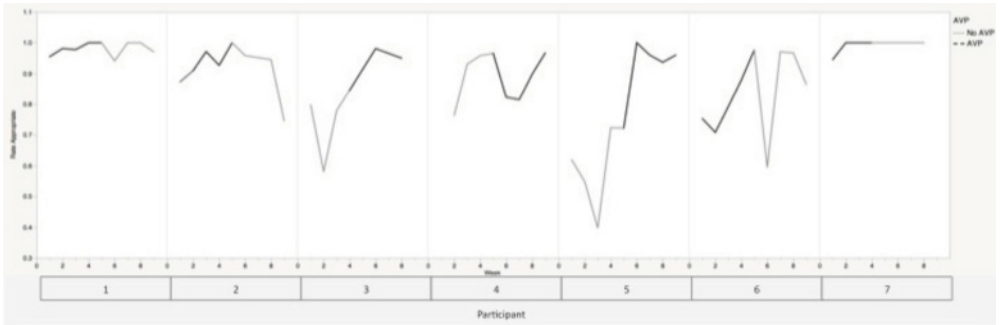

Fig. 7. Rate of appropriate use of reciprocity components over weeks for each participant when children first interacted with the AVP and did not interact with the AVP.

the groups using the AVP changed. There is some drop in the mean rate of appropriate response in week 6, but overall this figure illustrates the overall increased rate of appropriate response with the AVP, as well as over the course of the program, suggested by the regression analysis. Figure 7 illustrates the results for each individual participant, and reflects the results as well. In particular, participants 1, 2, 3, and 5 illustrate

clear increased rate of appropriate responses with the AVP, and participants 3 and 5 illustrate clear increased rate of appropriate responses over the course of the program.

To test if specific reciprocity components were particularly affected by the intervention, we ran another least squares regression by component, with factors of week and AVP to yield nine models (though not all models included enough observations to build the model). Three reciprocity components yielded significant factors for AVP use, week, or both. For the feedback component (51 observations; Whole Model Test: $R^2 = 0.42$; $p < 0.0001$), AVP use ($p < 0.03$), and week ($p < 0.03$) both significantly predicted appropriate responses, with a higher rate of appropriate responses when children first interacted with the AVP and over the course of the intervention. For the respond component (59 observations; Whole Model Test: $R^2 = 0.31$; $p < 0.0001$), AVP use ($p < 0.04$) significantly predicted appropriate responses, with a higher rate of appropriate responses when children first interacted with the AVP. Finally, for the share information component (55 observations; Whole Model Test: $R^2 = 0.51$; $p < 0.0001$), week ($p < 0.01$) significantly predicted appropriate responses, with a higher rate of appropriate responses over the course of the intervention.

7.2.1. Example. Role play scenarios between participant 3 (P3) and participant 5 (P5) provide an example for how appropriate responses during role play may have been affected by the AVP and over the course of the intervention. We chose this example because it is a clear illustration of differences. In week 1, the rates of appropriate responses for the participants were 0.797 for P3 and 0.691 for P5. During the role play, they were prompted to introduce themselves to each other by talking about their interests. P5 began the conversation with nonreciprocal utterances unrelated to getting to know each other:

> P5: I just went to school. It was hot today. They turned off the lights and [unintelligible]. Someone got in trouble. I'm so mad.

When prompted by the therapist to talk about their interests, P5 awkwardly asked:

P5: Uh, what's it about you anyways

In week 6, P3 and P5 began using the AVP. The rate of appropriate responses during role-play were 0.981 for P3 and 1 for P5. During this role-play scenario, they were first prompted to have a conversation that the therapist could them quiz them on. The therapist would then ask them questions, intentionally asking some they would get wrong so they could practice that week's lesson, dealing with mistakes. During one of the quizzes, P5 correctly guesses P3 favorite animal. Then P3 begins a conversation about P5's favorite animal:

P3: What's your favorite animal?
P5: Lizards.
P3: Lizards? Do you like cats?
P5: No.
P3: I think they're cute.
P5: I don't like cats... I like lizards.

This on-topic, reciprocal interaction contrasts starkly with P5's awkward, off-topic and nonreciprocal contributions during week 1.

7.3. Lesson Components

To analyze predictors of appropriate use of specific lesson components, we ran a least squares regression on the rate of appropriate use of lesson components each week for each participant (52 observations). The observations in this model include all seven

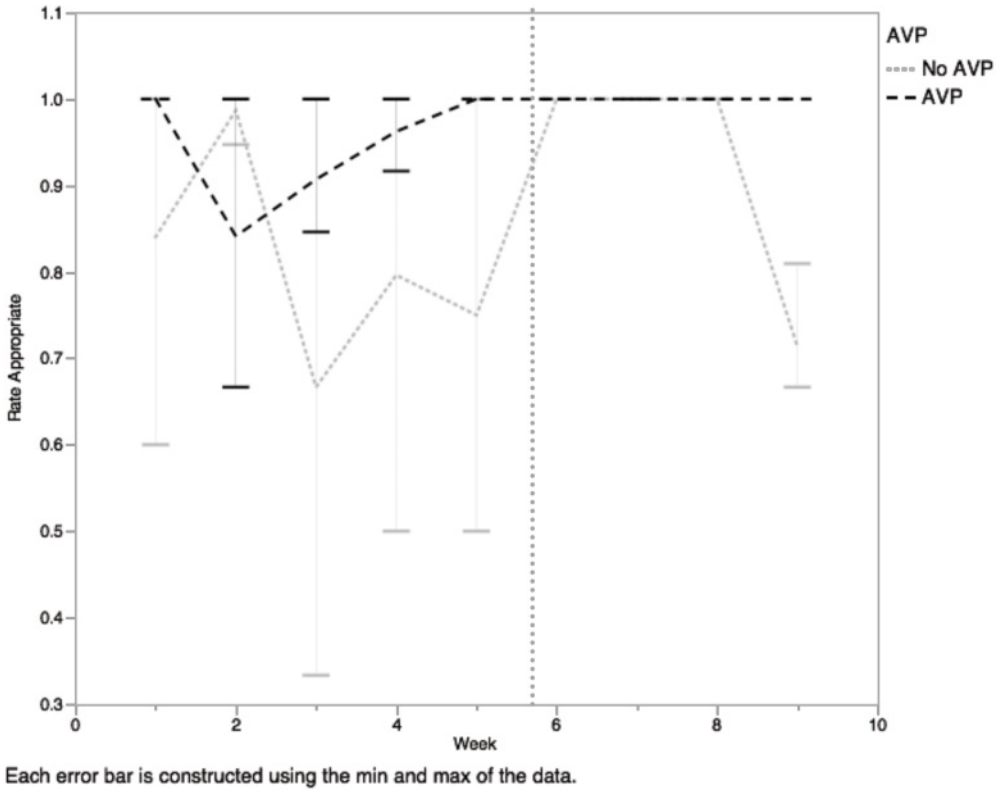

Fig. 8. Mean rate of appropriate use of lesson components over weeks when children first interacted with the AVP and did not interact with the AVP. The vertical dotted line at week 6 is when the groups using the AVP changed.

participants—each observation is the rate of appropriate use of lesson components for one participant during one role-play session. We included the same factors as in the model of reciprocity components (AVP, week, language, as well as interactions, and speaker as a random effect; $R^2 = 0.359$). Similar to the results for reciprocity components, we found a trend due to the use of the AVP, with a higher rate of appropriate responses predicted if the child first interacted with the AVP ($p < 0.059$). This suggests that if children first interacted with the AVP, their use of specific lesson components was more appropriate. We also found a trend for the interaction of language and week ($p < 0.056$), suggesting the use of appropriate responses over the course of the intervention was negatively impacted by language ability. Week was not a significant factor ($p = 0.306$). However, this is not unexpected: only some of the lesson components carried over week to week. Furthermore, lessons became more difficult over the course of the intervention.

Figure 8 illustrates the group's mean rate of appropriate use of lesson components during role play when they first interacted with the AVP versus did not interact with the AVP over the weeks of the intervention. The vertical dotted line marks week 6, when the groups using the AVP changed. There is some drop in the mean rate of appropriate response in week 2, but overall this figure illustrates the increased rate of appropriate response with the AVP compared with those who did not interact with the AVP, as suggested by the regression analysis. There is no clear increase in rate of appropriate responses over the course of the weeks of the program, as suggested by the regression

Table VI. Matched Pairs Wilcoxon Signed Rank Test for SRS Raw Total and Subscale Scores

Score	V	p-value (1-tailed, post < pre)
Total	8.5	0.20
Awareness	9	0.42
Cognition	11	0.33
Communication	1.5	0.04*
Mannerisms	9.5	0.25
Motivation	9.5	0.75

*$p < 0.05$.

analysis. There is also a drop in the mean rate of appropriate response in week 9. Future research could investigate if this reflects implications of withdrawing the AVP treatment with a longer study employing an appropriate single case experimental design.

7.3.1. Example. Participant 6's (P6) interactions during role play scenarios illustrate appropriate and inappropriate use of components of the social skills lessons. We chose this example because it is a clear illustration of differences. P6 did not use the AVP system during week 1, and her rate of appropriate use of lesson components was 0.762. During the role-play, which asked participants to get to know another person by asking questions, P6 tries to ask P1 questions, but some of the questions are awkward or socially inappropriate questions:

(1) P6: Did you pass kindergarten?
(2) P6: Do you have a mom or a dad?

During week 6, P6 used the AVP to practice having a respectful attitude by using request words and tone. Her rate of appropriate use of lesson components was 0.917 during the role-play scenario, and her exchanges illustrate appropriate requests and responses. For example:

P6: You want to play the new Super Mario Brothers?
P1: . . .oh that would be really nice. I'm a Super Mario fan.

7.4. Social Responsiveness Scale (SRS)

As described earlier, previous studies have found pre-/post- test rating scales are typically insensitive to change over short-term interventions. However, any potential results are promising for future studies and can inform the goals and design of future studies. Comparisons of the group's pre-/post- data for the SRS showed no significant differences for the total score and four of the sub-scales, social awareness, social cognition, social motivation and autistic mannerisms. However, there was a significant difference for the social communication subscale ($p < 0.05$). Table VI summarizes the results for the total SRS score as well as the subscales.

The social communication subscale of the SRS focuses on expressive social communication and includes items such as, "Has difficulty making friends, even when trying his or her best," "Gets frustrated trying to get ideas across in conversation," and "Has difficulty answering questions directly and ends up talking around the subject." Given the current intervention's focus on social communication, including interacting with friends, asking and answering questions, and sustaining a conversation, the significant change in this subscale of the SRS suggests the intervention as a whole may support children in developing social communication skills. While we cannot directly attribute this result to the AVP component of the intervention, and the repetition of Wilcoxon

Table VII. Standard Least Squares Regression by Condition for Parent Skills Ratings

Condition	R^2	F ratio	p-value(1-tailed, post > pre)
Not Covered	0.23	0.51	0.48
Covered, No AVP	0.27	0.97	0.33
Covered & AVP	0.25	0.41	0.52

tests weakens its significance, it is promising for the efficacy of social programs incorporating AVP technology.[5]

7.5. Parent Skills Ratings

Similar to previous intervention studies, comparison of the group's pre-/post- data on the parent skills ratings showed no significant difference for any condition (skills not covered, skills covered but for which children did not interact with the AVP, and skills for which children interacted with the AVP). Table VII summarizes the results.

8. DISCUSSION

This research aims to address the difficulty that children with autism spectrum and related disorders have accessing social interactions with their peers. We evaluated an intervention program targeting reciprocity using an authorable virtual peer as a component of a social skills group. The AVP adds a new instructional approach based on constructionist theory of learning by building artifacts.

We developed an 11-week, technology-integrated intervention program that incorporates the AVP into weekly social skills lessons based on Baker's [2003] approach. The AVP was integrated as an additional step in Baker's four-step structured learning technique. AVP interactions followed didactic instructions and modeling of skill steps and preceded role-play interactions. For the children who interacted with the AVP, this added a constructionist component to their learning: they had the opportunity to build interactive behaviors into a virtual peer similar to those they would then engage in with a peer during role-play.

To evaluate the intervention program as a whole, and more specifically whether authoring virtual peers helps children with ASD employ reciprocal social interactions during interactions with their peers, we asked: *Do AVP interactions help children employ reciprocal social interaction skills in interactions with their peers?* Two hypotheses guided the investigation: (1) after performing the AVP task, children with ASD will use more appropriate reciprocity skills in the dyad role-play task than when they do not first interact with the AVP and (2) over multiple sessions in the group, appropriate use of reciprocity skills will increase in the dyad role-play task. Toward the first hypothesis, our results suggest that appropriate use of general reciprocity skills, such as asking questions, responding, and sharing information, may increase when children first interact with an AVP versus when they do not. In particular, the reciprocity components of giving feedback and responding may be more appropriate after interacting with the AVP. Our results also suggest that appropriate use of specific lesson components, such as the "two-question rule" or "using request words and tone rather than demand words and tone," may be more appropriate when children first interact with an AVP.

[5]Fitting a standard least squares model on multiple responses for the SRS subscales using participant as a random effect yields a marginally significant pre–post factor for the communication subscale, $R^2 = 0.87$; F ratio $= 3.93$; $p = 0.09$, suggesting a trend still exists in the Communication subscale on a multiple regression. Other subscale responses remained insignificant.

Towards the second hypothesis, our results suggest appropriate use of reciprocity components may increase over the weeks of the intervention program. In particular, the reciprocity components of giving feedback and sharing may be more appropriate. Our results also suggest that increased reciprocity may be negatively impacted by language ability; some children's higher language ability may give them less opportunity to improve. Appropriate use of specific lesson components, however, did not increase over the weeks of the intervention program. This is not unexpected because not all lesson components carried over from week to week and difficulty of the lessons increased over the course of the program. These findings suggest efficacy of the intervention program as a whole, though not necessarily the AVP component specifically.

Taken together, these results offer support for using an AVP as a component of a group social skills program for supporting children's reciprocal social communication with their peers. The two biggest drawbacks of the program were (1) the amount of space needed to simultaneously run the group activity, AVP interactions, and role-play scenarios (two small rooms and two large rooms) and (2) the amount of time needed for each child to author the virtual peer; at most two pairs of children could interact with the AVP during each 2-hour intervention session. While the design of the current study required children to author the virtual peer individually and operate it with one partner, future studies could look at how to incorporate the AVP into the group activities so multiple children can benefit at once and not as much space is required.

Efficacy studies, like this study, are just the first step in establishing an evidence-based clinical intervention. Though the social communication subscale yielded significant results in this study, other "impact" [Ozonoff and Miller 1995] measures, including the total score and other subscales of the SRS as well as the parent skills ratings did not yield significant pre-/post- intervention changes. This is a common problem of efficacy studies and could be due to the short length of the intervention program, the small number of participants, or inappropriate and/or insensitive measures. In later phases of developing an intervention, these impact measures become more critical.

This initial evaluation has limitations, in particular the small number of participants and sessions. The regression analyses are based on a small number of observations and are, therefore are not conclusive. However, the scale of the study is appropriate based on where we are in the development of the intervention. From 2002 to 2004, the National Institutes of Health (NIH) sponsored a working group to discuss the methodological challenges of conducting research on interventions for children with ASD [Lord et al. 2005; Smith et al. 2007]. The results of this working group included two papers. One paper discusses the challenges inherent in conducting intervention studies with children with ASD such as participant recruitment, appropriate outcome measures, and putting research interventions into practice [Lord et al. 2005]. The second paper outlines four stages for conducting intervention research: (1) demonstrating efficacy; (2) compiling and pilot-testing an intervention manual; (3) conducting randomized clinical trials (RCTs) that evaluate the intervention under controlled conditions; and (4) conducting studies that evaluate intervention use in community therapeutic practices. Since the AVP is a new technique for working with this population, the first phase of research, which aims to provide a "proof of concept" [Smith et al. 2007] of the technique, is appropriate. While working with small sample sizes at this phase of research limits generalization of results, using small sample sizes addresses some of the challenges of this phase, such as limited resources or procedures that require adjustments and are not standardized. So while future work by clinical researchers can develop and evaluate a standardized intervention program, the current research has taken the first step by identifying a new intervention technique incorporating AVPs and offers evidence of its efficacy on reciprocal social interaction with peers.

ACKNOWLEDGMENTS

The authors would like to thank Autism Speaks, the Spencer Foundation, the Northwestern University Alumnae Association, Gregory Abowd, Andrea Binz, John Borland, Joan Chiao, Darren Gergle, Alberto Gonzalez, Kristina Rodriguez and Carol Rolland.

REFERENCES

American Psychiatric Association. 2013. *DSM-V Diagnostic Statistical Manual of Mental Disorders* (5th ed.). American Psychiatric Association.

J. E. Baker. 2003. *Social Skills Training for Children and Adolescents with Asperger Syndrome and Social-Communication Problems.* Autism Asperger Publishing Co., Shawnee Mission, KS.

C. A. M. Baltaxe. 1977. Pragmatic Deficits in the Language of Autistic Adolescents. *Journal of Pediatric Psychology* 2, 4, 176–180.

S. Baron-Cohen. 1995. *Mindblindness: An essay on autism and theory of mind*. MIT Press, Cambridge, MA.

S. Baron-Cohen. 2002. The Extreme Male Brain Theory of Autism. *Trends in Cognitive Sciences* 6, 248–254.

T. D. Barry, L. G. Klinger, J. M. Lee, N. Palardy, T. Gilmore, and S. D. Bodin. 2003. Examining the Effectiveness of an Outpatient Clinic-Based Social Skills Group for High-Functioning Children with Autism. *Journal of Autism and Developmental Disorders* 33, 685–701.

L. Bartoli, F. Garzotto, M. Gelsomini, L. Oliveto, and M. Valoriani. 2014. Designing and Evaluating Touchless Playful Interaction for ASD Children. In *Proceedings of the Interaction Design and Children (IDC'14).* ACM, New York, 17–26.

M. U. Bers and J. Cassell. 1998. Interactive Storytelling Systems for Children: Using Technology to Explore Language and Identity. *Journal of Interactive Learning Research* 9, 183–215.

A. Bosseler and D. W. Massaro. 2003. Development and Evaluation of a Computer-Animated Tutor for Vocabulary and Language Learning in Children with Autism. *Journal of Autism and Developmental Disorders* 33, 653–672.

B. Brinton, M. Fujiki, and J. M. Powell. 1997. The ability of children with language impairment to manipulate topic in a structured task. *Language, Speech, and Hearing Services in Schools* 28, 3–11.

L. Capps, J. Kehres, and M. Asigman. 1998. Conversational Abilities Among Children with Autism and Children with Developmental Delays. *Autism: The International Journal of Research and Practice* 2, 325–244.

Carnegie Mellon University. 2010. Panda3D - Free 3D Game Engine. Retrieved from https://www.panda3d.org/.

J. Cassell, M. Ananny, A. Basu, T. Bickmore, P. Chong, D. Mellis, K. Ryokai, J. Smith, H. Vilhjálmsson, and H. Yan. 2000. Shared Reality: Physical Collaboration with a Virtual Peer. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'00).* ACM, New York, 259–260.

J. Cassell, K. Geraghty, B. Gonzalez, and J. Borland. 2009. Modeling Culturally Authentic Style Shifting with Virtual Peers. In *Proceedings of the ICMI-MLMI.* ACM, New York, 135–142.

J. Cassell, A. Tartaro, Y. Rankin, V. Oza, and C. Tse. 2007. Virtual Peers for Literacy Learning. *Educational Technology* 47, 39–43.

R. Cole, S. Van Vuuren, B. Pellom, K. Hacioglu, J. Ma, J. Movellan, S. Schwartz, D. Wade-Stein, W. Ward, and J. Yan. 2003. Perceptive Animated Interfaces: First Steps Toward a New Paradigm for Human Computer Interaction. *Proceedings of the IEEE* 91, 9, 1391–1405.

J. N. Constantino. 2005. Social Responsiveness Scale. Retrieved from http://www.wpspublish.com/store/p/2993/social-responsiveness-scale-srs-by-john-n-constantino-md.

M. Cramer, S. E. Hirano, M. Tentori, M. T. Yeganyan, and G. R. Hayes. 2011. Classroom-Based Assistive Technology: Collective Use of Interactive Visual Schedules by Students with Autism. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'11).* ACM, New York, 1–10.

K. Dautenhahn and I. Weery. 2004. Towards Interactive Robots in Autism Therapy. *Pragmatics and Cognition* 12, 1–35.

K. Dautenhahn, I. Werry, J. Rae, P. Dickerson, P. Stribling, and B. Ogden. 2002. Robotic Playmates: Analysing Interactive Competencies of Children with Autism Playing with a Mobile Robot. In *Socially Intelligent Agents—Creating Relationships with Computers and Robots*, K. Dautenhan, A. Bond, L. Canamero, and B. Edmonds (Eds.). Kluwer Academic Publishers, 117–124.

T. Deangelis. 2009. Can Second Life Therapy Help with Autism? *American Psychological Association* 40, 40.

M. Dennis, A. L. Lazenby, and L. Lockyer. 2001. Inferential Language in High-Function Children with Autism. *Journal of Autism and Developmental Disorders* 31, 47–54.

C. A. Disalvo and D. P. Oswald. 2002. Peer-mediated Interventions to Increase the Isocial interaction of children with autism: Consideration of peer expectancies. In *Focus on Autism and Developmental Disorders* 14, 198–207.

L. Escobedo, D. H. Nguyen, L. Boyd, S. H. Hirano, A. Rangel, D. Garcias-Rosas, M. Tentori, and G. R. Hayes. 2012. MOSOCO: A Mobile Assistive Tool to Support Children with Autism Practicing Social Skills in Real-life Situations. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'12).* ACM, 2589–2598.

D. Feil-Seifer, M. Black, M. J. Mataric, and S. Narayanan. 2009. Designing Interactive Technologies for Supporting Research in Autism Spectrum Disorders. In *Proceedings of the International Meeting for Autism Research.*

D. Feil-Seifer and M. J. Mataric. 2010. Using Proxemics to Evaluate Human-Robot Interaction. In *Proceedings of the ACM/IEEE International Conference on Human-Robot Interaction.* ACM, New York, 143–144.

J. Fine, G. Bartolucci, P. Szatmari, and G. Ginsberg. 1994. Cohesive Discourse in Pervasive Developmental Disorders. *Journal of Autism and Developmental Disorders* 24, 315–329.

U. Frith. 2003. *Autism: Explaining the Enigma*. Blackwell Publishing, Malden, MA.

U. Frith, F. Happe, and F. Siddons. 1994. Autism and Theory of Mind in Everyday Life. *Social Development* 3, 108–134.

E. Gal, N. Bauminger, D. Goren-Bar, F. Pianesi, O. Stock, M. Zancanaro, and P. L. T. Weiss. 2009. Enhancing Social Communication of Children with High-Functioning Autism through a Co-Located Interface. *AI & Society* 24, 75–84.

J. E. Gombert. 1992. *Metalinguistic Development*. University of Chicago Press, Chicago.

C. A. Gray and J. D. Garand. 1993. Social Stories: Improving Responses of Students with Autism with Accurate Social Information. *Focus on Autistic Behavior* 8, 1–10.

J. Hailpern, K. Karahalios, and J. Halle. 2009. Creating a Spoken Impact: Encouraging Vocalizations Through Audio Visual Feedback in Children with ASD. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'09).* ACM, New York, 453–462.

C. M. Hale and H. Tager-Flusberg. 2005. The Relationship between Discourse Deficits and Autism Symptomology. *Journal of Autism and Developmental Disorders* 35, 519–524.

F. G. E. Happe. 1993. Communicative Competence and Theory of Mind in Autism: A Test of Relevance Theory. *Cognition* 48, 101–119.

I. Harel and S. Papert. 1991. Situating Constructionism. In *Constructionism*, I. Harel and S. Papert (Eds). Ablex Publishing Corporation.

G. R. Hayes, J. A. Kientz, K. N. Truong, D. R. White, G. D. Abowd, and T. Pering. 2004. Designing Capture Applications to Support the Education of Children with Autism. In *Proceedings of the UbiComp.* 161–178.

A. Hillier, T. Fish, P. Cloppert, and D. Q. Beversdorf. 2007. Outcomes of a Social and Vocational Skills Support Group for Adolescents and Young Adults on the Autism Spectrum. *Focus on Autism and Other Developmental Disabilities* 22, 107–115.

S. H. Hirano, M. T. Yeganyan, G. Marcu, D. H. Nguyen, L. A. Boyd, and G. R. Hayes. 2010. vSked: Evaluation of a System to Support Classroom Activities for Children with Autism. In *Proceedings of the ACM SIGCHI Conference on Human Factors in Computing Systems (CHI'10).* ACM, New York, 1633–1642.

J. P. Hourcade, N. E. Bullock-Rest, and T. E. Hansen. 2012. Multitouch Tablet Applications and Activities to Enhance the Social Skills of Children with Autism Spectrum Disorders. *Personal and Ubiquitous Computing* 16, 157–168.

P. Howlin and P. Yates. 1999. The Potential Effectiveness of Social Skills Groups for Adults with Autism. *Autism* 3, 299–307.

D. A. Jackson, N. F. Jackson, M. L. Bennett, B. M. Bynum, and E. Faryna. 1991. *Learning to Get Along: Social Effectiveness Training for People with Developmental Disabilities.* Research Press, Champaign, IL.

C. Jarrett. 2009. Get a Second Life. *The Psychologist* 22, 490–493.

Y. B. Kafai. 1995. *Minds in Play: Computer Game Design as a Context for Children's Learning.* L. Erlbaum Associates, Hillsdale, NJ.

D. M. Kamps, B. R. Leonard, S. Vernon, E. P. Dugan, J. C. Delquadri, B. Gershon, L. Wade, and L. Folk. 1992. Teaching Social Skills to Students with Autism to Increase Peer Interactions in an Integrated First-Grade Classroom. *Journal of Applied Behavior Analysis* 25, 281–288.

S. J. Kerr, H. R. Neale, and S. V. G. Cobb. 2002. Virtual Environments for Social Skills Training: The Importance of Scaffolding in Practice. In *Proceedings of the International ACM Conference on Assistive Technologies (Assets).* 104–110.

J. A. Kientz, M. S. Goodwin, G. R. Hayes, and G. D. Abowd. 2014. *Interactive Technologies for Autism.* Morgan & Claypool.

J. A. Kientz, G. R. Hayes, G. D. Abowd, and R. E. Gritner. 2006. From the War Room to the Living Room: Decision Support for Home-based Therapy Teams. In *Proceedings of the Conference on Computer Supported Collaborative Work (CSCW'06).* ACM, New York, 209–218.

J. A. Kientz, G. R. Hayes, T. L. Westeyn, T. Starner, and G. D. Abowd. 2007. Pervasive Computing and Autism: Assisting Caregivers of Children with Special Needs. *Pervasive Computing* 6, 1, 28–35.

S. Kopp, B. Kren, S. Marsella, A. N. Marshall, C. Pelachaud, H. Pirker, K. R. Thorisson, and H. Vilhjalmsson. 2006. Towards a Common Framework for Multimodal Generation: The Behavior Markup Language. In *Proceedings of the International Conference on Intelligent Virtual Agents.* Marina Del Rey, CA, 205–217.

L. Krasny, B. J. Williams, S. Provencal, and S. Ozonoff. 2003. Social Skills Interventions for the Autism Spectrum: Essential Ingredients and a Model Curriculum. *Child and Adolescent Psychiatric Clinics* 12, 107–122.

K. A. Kroeger, J. R. Schultz, and C. Newsom. 2007. A Comparison of Two Group-Delivered Social Skills Programs for Young Children with Autism. *Journal of Autism and Developmental Disorders* 37, 808–817.

D. B. Legoff. 2004. Use of LEGO(c) as a Therapeutic Medium for Improving Social Competence. *Journal of Autism and Developmental Disorders* 34, 557–571.

L. Little. 2003. Maternal Perceptions of the Importance of Needs and Resources for Children with Asperger Syndrome and Nonverbal Learning Disorders. *Focus on Autism and Other Developmental Disabilities* 18, 257–266.

C. Lord, A. Wagner, S. Rogers, P. Szatmari, M. Aman, T. Charman, G. Dawson, V. M. Durand, L. Grossman, D. Guthrie, S. Harris, C. Kasari, L. Marcus, S. Murphy, S. Odom, A. Pickles, L. Scahill, E. Shaw, B. Siegel, M. Sigman, W. Stone, T. Smith, and P. Yoder. 2005. Challenges in Evaluating Psychosocial Interventions for Autistic Spectrum Disorders. *Journal of Autism and Developmental Disorders* 35, 695–708.

K. A. Loveland, S. H. Landry, S. O. Hughes, S. K. Hall, and R. E. Mcevoy. 1988. Speech Acts and the Pragmatic Deficits of Autism. *Journal of Speech and Hearing Research* 31, 593–604.

G. Marcu, D. H. Nguyen, and G. R. Hayes. 2009. Use of a Wearable Recording Device in Therapeutic Interventions for Children with Autism. In *Proceedings of the International Meeting for Autism Research*.

S. R. Mcconnell. 2002. Interventions to Facilitate Social Interaction for Young Children with Autism: Review of Available Research and Recommendations for Educational Intervention and Future Research. *Journal of Autism and Developmental Disorders* 32, 351–372.

E. Mcginnis and A. P. Goldstein. 1984. *Skillstreaming the Elementary School Child.* Research Press, Champaign, IL.

E. Mcginnis and A. P. Goldstein. 1990. *Skillstreaming in Early Childhood.* Research Press Company, Champaign, IL.

F. Michaud and C. Theberge-Turmel. 2002. Mobile Robotic Toys and Autism. In *Socially Intelligent Agents: Creating relationships with Computers and Robots*, K. Dautenhahn, A. H. Bond, L. Canamero, and B. Edmonds (Eds). Kluwer Publishers, 125–132.

C. F. Norbury and D. V. M. Bishop. 2002. Inferential Processing and Story Recall in Children with Communication Problems: A Comparison of Specific Language Impairment, Pragmatic Language Impairment and High-Functioning Autism. *International Journal of Language & Communication Disorders* 37, 227–251.

G. I. Orsmond, M. W. Krauss, and M. M. Seltzer. 2004. Peer Relationships and Social and Recreational Activities among Adolescents and Adults with Autism. *Journal of Autism and Developmental Disorders* 34, 245–256.

S. Ozonoff and J. N. Miller. 1995. Teaching Theory of Mind: A New Approach to Social Skills Training for Individuals with Autism. *Journal of Autism and Developmental Disorders* 25, 415–433.

S. Parsons, P. Mitchell, and A. Leonard. 2004. The Use and Understanding of Virtual Environments by Adolescents with Autistic Spectrum Disorders. *Journal of Autism and Developmental Disorders* 34, 449–466.

R. W. Picard. 2009. Future Affective Technology for Autism and Emotion Communication. *Philosophical Transactions of the Royal Society B: Biological Sciences* 364, 3575–3584.

A. M. Piper, E. O'brien, M. R. Morris, and T. Winograd. 2006. SIDES: A Cooperative Tabletop Computer Game for Social Skills Development. In *Proceedings of the Conference on Computer Supported Cooperative Work (CSCW'06)*, ACM, New York, 1–10.

K. Porayska-Pomsta, C. Frauenberger, H. Pain, G. Rajendran, T. Smith, R. Menzies, M. E. Foster, A. Alcorn, S. Wass, S. Bernadini, K. Avramides, W. Keay-Bright, J. Chen, A. Waller, K. Guldberg, J. Good,

and O. Lemon. 2012. Developing Technology for Autism: An Interdisciplinary Approach. *Personal and Ubiquitous Computing* 16, 117–127.

R. Ramloll, C. Trepagnier, M. Sebrechts, and A. Finkelmeyer. 2004. A Gaze Contingent Environment for Fostering Social Attention in Autistic Children. In *Proceedings of the Eye Tracking Research & Applications Symposium on Eye Tracking Research & Applications.* 19–26.

B. Reichow and F. R. Volkmar. 2010. Social Skills Interventions for Individuals with Autism: Evaluation for Evidence-Based Practices within a Best Evidence Synthesis Framework. *Journal of Autism and Developmental Disorders* 40, 149–166.

B. Reichow, F. R. Volkmar, and D. V. Cicchetti. 2008. Development of the Evaluative Method for Evaluating and Determining Evidence-Based Practices in Autism. *Journal of Autism and Developmental Disorders* 38, 1311–1319.

M. Resnick, A. Bruckman, and F. Martin. 1996. Pianos Not Stereos: Creating Computational Construction Kits. *Interactions* 3, 40–50.

J. Robertson and K. Nicholson. 2007. Adventure Author: a learning environment to support creative design. In *Proceedings of the Interaction Design and Children (IDC'07).* ACM, New York, 37–44.

B. Robins, P. Dickerson, P. Stribling, and K. Dautenhahn. 2004. Robot-Mediated Joint Attention in Children with Autism: A Case Study in Robot-Human Interaction. *Interaction Studies* 5, 161–198.

L. Ruble, H. Willis, and V. M. Crabtree. 2008. Social Skills Group Therapy for Autism Spectrum Disorders. *Clinical Case Studies* 7, 287–300.

B. Scassellati. 2005. Quantitative Metrics of Social Response for Autism Diagnosis. In *Proceedings of the 14th International Workshop on Robot and Human Interactive Communication (ROMAN'05).* 585–590.

B. Scassellati. 2005. Using Robots to Study Abnormal Social Development. In *Proceedings of the 5th International Workshop on Epigenetic Robotics (EpiRob'05).*

F. Shic, K. Chawarska, D. Lin, and B. Scassellati. 2007. Measuring Context: The Gaze Patterns of Children with Autism Evaluated from the Bottom-Up. In *Proceedings of the 6th IEEE International Conference on Development and Learning (ICDL'07).*

F. Shic, K. Chawarska, and B. Scassellati. 2008. The Amorphous Fixation Measure Revisited: with Applicaitons to Autism. In *Proceedings of the 30th Annual Meeting of the Cognitive Science Society.*

F. Shic, W. Jones, A. Klin, and B. Scassellati. 2006. Swimming in the Underlying Stream: Computational Models of Gaze in a Comparative Behavioral Analysis of Autism. In *Proceedings of the Cognitive Science.*

T. Smith, L. Scahill, G. Dawson, D. Guthrie, C. Lord, S. Odom, S. Rogers, and A. Wagner. 2007. Designing Research Studies on Psychosocial Interventions in Autism. *Journal of Autism and Developmental Disorders* 37, 354–366.

B. Steiner, N. Kaplan, and S. Moulthrop. 2006. When Play Works: Turning Game-Playing into Learning. In *Proceedings of the Interaction Design and Children (IDC'06)*, ACM, New York, 137–140.

H. Tager-Flusberg. 1994. Dissociations in Form and Function in the Acquisition of Language by Autistic Children. In *Constraints on Language Acquisition: Studies of Atypical Children*, H. Tagerflusberg, Ed. Lawrence Erlbaum Associates, Hillsdale, NJ, 175–194.

H. Tager-Flusberg. 2000. Language and Understanding Minds: Connections in Autism. In *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience*, S. Baron-Cohen, H. Tager-Flusberg, and D. J. Cohen (Eds). Oxford University Press, Oxford, 1–45.

H. Tager-Flusberg and M. Anderson. 1991. The Development of Contingent Discourse Ability in Autistic Children. *Journal of Child Psychology and Psychiatry* 32, 1123–1134.

A. Tartaro. 2011. *Authorable Virtual Peers: Technology as an Intervention for Difficulties with Peer Social Interaction in Autism Spectrum and Related Disorders.* Doctoral Dissertation. Northwestern University.

A. Tartaro and J. Cassell. 2008. Playing with Virtual Peers: Bootstrapping Contingent Discourse in Children with Autism. In *Proceedings of the International Conference of the Learning Sciences.* ACM, New York, 382–389.

A. Tartaro and C. Ratz. 2014. Incorporating Technology into Peer Social Group Programs. In *Technology Tools for Students with Autism: Innovations that Enhance Independence and Learning*, K. I. Boser, M. S. Goodwin, and S. C. Wayland (Ed.). Paul H. Brookes Publishing, Baltimore, MD, 185–200.

M. Tentori and G. R. Hayes. 2010. Designing for Interaction Immediacy to Enhance Social Skills of Children with Autism. In *Proceedings of UbiComp.* 51–60.

M. Thiebaux, A. N. Marshall, S. Marsella, and M. Kallmann. 2008. SmartBody: Behavior Realization for Embodied Conversational Agents. In *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems (AAMAS'08).*

L. L. Travis and M. Sigman. 1998. Social Deficits and Interpersonal Relationships in Autism. *Mental Retardation and Developmental Disabilities Research Reviews* 4, 65–72.

C. Y. Trepagnier, D. E. Olsen, L. Boteler, and C. A. Bell. 2011. Virtual Conversation Partner for Adults with Autism. *Cyberpsychology, Behavior and Social Networking* 14, 1–2, 21–27.

A. Wang and J. Cassell. 2003. Co-authoring, Corroborating, Criticizing: Collaborative Storytelling between Virtual and Real Children. In *Proceedings of the Workshop of Educational Agents: More than Virtual Tutors.*

B. J. Webb, S. P. Miller, T. B. Pierce, S. Strawser, and W. P. Jones. 2004. Effects of Social Skill Instruction for High-Functioning Adolescents with Austism Spectrum Disorders. *Focus on Autism and Other Developmental Disabilities* 19, 53–62.

L. R. Weingart. 1997. How Did They Do That? The Ways and Means of Studying Group Process. *Research in Organizational Behavior* 19, 189–239.

K. M. Wilkinson. 1998. Profiles of Language and Communication Skills in Autism. *Mental Retardation and Developmental Disabilities Research Reviews* 4, 73–79.

B. Wise, R. Cole, S. Van Vuuren, S. Schwartz, L. Snyder, N. Ngampatipatpong, J. Tuantranont, and B. Pellom. 2007. Learning to Read with a Virtual Tutor: Foundations to Literacy. In *Interactive Literacy Education: Facilitating Literacy Learning Environments through Technology.* C. Kinzer and L. Verhoeven (Eds.). Erlbaum, Mahway, NJ.

Received April 2013; revised September 2014; accepted September 2014