

# FRAMEWORK AND ASSESSMENT OF CONVERSATIONAL VIRTUAL HUMANS AS ROLE-PLAYERS IN SIMULATED SOCIAL ENCOUNTERS WITH PEOPLE

**John L. Hart<sup>1\*</sup>, Michael D. Proctor<sup>2</sup>**

<sup>\*1</sup>*U.S. Army Research Laboratory, USA*

<sup>2</sup>*Department of Industrial Engineering and Management Systems, University of Central Florida, USA*

<sup>\*1</sup>*Email: [john.l.hart24.civ@mail.mil](mailto:john.l.hart24.civ@mail.mil), <sup>2</sup>[Michael.Proctor@ucf.edu](mailto:Michael.Proctor@ucf.edu)*

**\*Corresponding Author: -**

*Email: [john.l.hart24.civ@mail.mil](mailto:john.l.hart24.civ@mail.mil)*

---

## **Abstract: -**

*As technology advances, people increasingly converse with computer conversational agents serving in a personal assistant role such as Siri or Cortona. Beyond the personal assistant role, a virtual humans may substitute for a person in another role during a social encounter in a simulated environment. From an experiential or pedantic training perspective, the resulting simulated social encounter may provide a forum by which real people may gain social skills. This article reviews advances in technology and science of virtual humans in face-to-face social encounters with people. An assessment framework is proposed based on a modification of the Schroeder, Adesope, & Gilbert framework [1]. Conversational attributes include verbal content and expression and nonverbal behavior such as head position or inclination, micro and macro facial or body emotional expression, eye gaze and direction, and other facial and other bodily gestures. Using that framework, ten instantiations of virtual humans are analyzed with respect to simulating a person in a bi-directional conversation. Findings and discussion address the question: can a virtual human truly be a social partner with a person for the purpose of social skills training? Within a social skills learning setting, analysis raises concern about the impact of textual data concurrently displayed with verbal message content and expression, and nonverbal behavior. Suggested future advanced technology research includes virtual human eye tracking technology in order to a better understand the cognitive science and behavioral responses associated with periods of engagement and boredom during interactions with a person.*

**Keywords: -** *Virtual human communications, virtual role-players, simulated social interaction, communication, conversational model*



**I. INTRODUCTION**

With the development of agents like Siri and Cortona, people now converse with computer conversational agents on a daily basis. Visually embodied and intelligently-enhanced conversational agents called virtual humans assume various roles across a wide spectrum of virtual environments from games to education including experiential and pedantic social communication skills training of people [2-8]. As “software entities [that] look and act like people and can engage in conversation and collaborative tasks,” [9], virtual humans have advantages over people as social skills role-players in that virtual humans provide consistent performances unaffected by fatigue, illness, or personal issues [10]. Examples of virtual humans as role- players in training environments include law enforcement personnel [11, 12], medical patients [13, 14], and military personnel [10, 15, 16]. The capability to simulate face-to-face conversations between a virtual human and a person holds great promise for social skills training and development [17, 18].

**II. Conversations With Virtual Humans**

Despite recent successes, conversational agent implementation failures abound. Eliza, an early computer-based communications agent, simply responded to users based on key words and only reflected the user’s input [20] making a long conversations difficult. Microsoft’s Clippy failed to engage users [21]. To better understand the source of failure or success, Stamp [19] identified seventeen different scientific categories of research that contribute to the capability to successfully simulate a conversation between a computer agent and a person. Later, scientific research demonstrated that adding human-like attributes, such as a face, made the interaction with people more comfortable and natural [22, 23] since people accept and may even prefer communication agents with human faces [22, 24, 25], social behaviors [26], basic emotions [27, 28], and nonverbal behavioral feedback [29, 30].

In terms of advanced technology, people may accept conversational virtual humans as if they are real [25, 31-34] and perceive them as social partners, progressing science from a research novelty [33, 35-38] to instantiated advanced technology as actual role-players in real-world training applications [39-42]. Recent research indicates virtual human role-players may improve a person’s job interview skills [42] and may improve social behaviors of children with autism spectrum disorder [41].

This article synthesizes literature within the last ten years on the topic of virtual humans participating in simulated conversations and concludes with a discussion of the current challenges and possible future scientific and advanced technology directions.

**III. Conversational Virtual Human Feature Framework**

Human conversation often involves more than intelligent content and voice cadence, tone, and accents of an audible message. In face-to-face social encounters, macro and micro facial expressions, hand and body gestures, and eye contact are nonverbal communicative behaviors humans use to place emphasis on message content as well as show attention and interest [17, 18, 43, 44]. Social skills training for face-to-face social settings intends to develop social or communication skills by understanding partner messages and if and how to exhibit appropriate social behavior [17, 18].

Schroeder, Adesope, and Gilbert proposed a framework for pedagogical agents where the agents are defined as “on-screen characters that facilitate instruction to the learner” who by definition are not intelligent, but may be combined with intelligent tutoring systems [1]. While the original framework [1] did not include intelligence, Table 1 presents a modified framework adding intelligence [15, 39, 47, 50, 61, 64, 76]. The framework accounts for the full range of on screen virtual characters they may be potentially engaged with a person in a simulated, face-to-face social encounter. Touch, smell, and touch features and physically embodied virtual humans are beyond this research scope.

**TABLE 22: SOCIAL SKILLS TRAINING APPLICATIONS UTILIZING VIRTUAL ROLE-PLAYERS (2005-2016)**

Features	Characteristic Levels	Communication Function
<b>Form or Appearance</b>	<ul style="list-style-type: none"> <li>• Humanoid</li> <li>• Non-humanoid</li> <li>• Video-quality human -</li> <li>• Mix</li> </ul>	Visual appearance of virtual human with face and eyes providing a reference for the visual system and eye gaze provides attention signals
<b>Voice Type</b>	<ul style="list-style-type: none"> <li>• No voice</li> <li>• Computer-edited</li> <li>• Computer generated/Synthetic</li> <li>• Human</li> </ul>	Provides verbal communication channel with auxiliary information from tone, inflections, stress, rhythm, prosody and emotion
<b>Animation Effect and Physical Effect</b>	<ul style="list-style-type: none"> <li>• Animated Gestures</li> <li>• Animated Facial Expressions (displays of emotion)</li> <li>• Static</li> <li>• Mixed agent types</li> <li>• Self Propulsion</li> <li>• Object manipulation</li> </ul>	Micro and macro facial expressions, and bold or subtle appendage or bodily gestures provide nonverbal communication channels and signals for conversational control and flow
<b>Artificial Intelligence</b>	<ul style="list-style-type: none"> <li>• Sensing &amp; Perception</li> <li>• Knowledge-base with Short &amp; Long term Memory</li> <li>• Cognition</li> <li>• Communications</li> <li>• Believes, desires, and intentions</li> </ul>	Provides cognitive functions for: <ul style="list-style-type: none"> <li>• Understanding and perception of verbal and nonverbal behaviors of conversational partner using both audio and visual inputs</li> <li>• Situational awareness and understanding for decision making and planning of responses</li> </ul>

**IV. Representative Instantiated Conversational Virtual Humans**

Past instantiated virtual humans provide scientific and technology lessons learned that may improve future virtual human designs. The interface is the basis for the conversational model allowing multiple methods of single or bi-directional communications. Some of the earliest virtual humans or conversational agents were human-like in both their appearance and behavioral animations for their body, arms, and face [33, 35-38].

Many of the early instantiations used computerized generated voices [33, 35, 36, 38] where [38] used recorded human responses for increased realism. User interfaces vary from automated speech recognition [33, 35, 38] to menu [37] or keyboard inputs [36]. Often interfaces lack a means to perceive the real world environment of the user or the user’s attention but some agent system introduced cameras to sense the presence of the human user [33, 35]. A visual system or camera may provide a virtual human with essential information or cues beyond verbal feedback cues that a social partner may use as part of meta-analysis of the state or progress of a face-to-face conversation.

Face-to-face social skills training systems often utilize virtual role-players in simulated face-to-face conversations and report varying levels of success [6, 11, 39, 41, 42, 45]. Table 2 contains a subset of applications identified in articles retrieved from structured searches including Google Scholar and multiple search criteria including “virtual human,” “role-player,” “social skills,” “training,” and “conversation”.

Applications selected for Table 2 are actual or representative training applications where people or persons interact in a simulated social encounter with role-playing virtual humans who primarily focus on aiding in the development of various social skills in people. As opposed to virtual humans, avatars, wherein a human adds intelligence and/or voice to an onscreen animation, are not included in the list below.

**TABLE 22: SOCIAL SKILLS TRAINING APPLICATIONS UTILIZING VIRTUAL ROLE-PLAYERS (2005-2016)**

System Name	Reference	Training Area	Person Communication Channels to Virtual Human	Virtual Human Communications Channels to Human
SimSensi “Ellie” (R&D prototype)	[46, 47]	Healthcare Interviewer (no training)	<ul style="list-style-type: none"> <li>• Speech recognition</li> <li>• Facial expressions</li> <li>• Head and body movements</li> <li>• Camera and Kinect with MultiSense for detection of smile intensity, facial expressions, head/body movement, and gaze direction</li> </ul>	<ul style="list-style-type: none"> <li>• Pre-recorded voice responses</li> <li>• Virtual human representation displayed on large screen monitor (&gt; 40 inches) with animated behaviors synchronized to speech</li> </ul>
Virtual Reality Job Interview Trainer (VR-JIT) (Commercial product)	[42, 49]	Job interview training	<ul style="list-style-type: none"> <li>• Speech recognition of specific user responses are provided</li> <li>• No visual system</li> </ul>	<ul style="list-style-type: none"> <li>• Recorded video clips of human actor controlled by non-branching logic</li> </ul>
Authorable Virtual Peer (AVP) (R&D prototype)	[41]	Social interactions of children with autism spectrum disorder (ASD)	<ul style="list-style-type: none"> <li>• Speech recognition and text files for users to author new content</li> <li>• User nonverbal behaviors observed via camera for human observer feedback and control</li> </ul>	<ul style="list-style-type: none"> <li>• Projected life-sized virtual human</li> <li>• Pre-recorded human voice with synchronized animated facial expressions and behaviors</li> <li>• GUI for selecting virtual peers social behaviors</li> </ul>
BiLat (Government product)	[39]	Negotiation in cultural context	<ul style="list-style-type: none"> <li>• Menu selection of user statements</li> <li>• No visual system</li> </ul>	<ul style="list-style-type: none"> <li>• Virtual human displayed on desktop or laptop monitor with synchronized animations</li> <li>• Computer generated voice responses</li> </ul>
Tactical Language and Culture	[6, 50]	Foreign language and cultural training	<ul style="list-style-type: none"> <li>• Speech recognition inputs</li> </ul>	<ul style="list-style-type: none"> <li>• Virtual human with synchronized animations in virtual environment on desktop monitor or laptop display</li> </ul>
Training System (TLCTS) (Commercial product)			<ul style="list-style-type: none"> <li>• Mouse selection of cultural gestures</li> <li>• No visual system</li> </ul>	<ul style="list-style-type: none"> <li>• Recorded voice responses</li> </ul>
INOTS & ELITE (Government product)	[10, 45]	Military person and performance counseling	<ul style="list-style-type: none"> <li>• Speech recognition of multiple choice user statements</li> </ul>	<ul style="list-style-type: none"> <li>• Virtual human life-size screen display</li> <li>• Pre-recorded VH voice responses</li> <li>• No visual system</li> </ul>
Job Interview Simulator (R&D prototype)	[51]	Job interview skills (no training)	<ul style="list-style-type: none"> <li>• Speech recognition inputs</li> <li>• Hand and body positions and movements</li> <li>• Facial expressions – smile</li> <li>• Voice activity</li> <li>• Head position</li> <li>• Camera and Kinect inputs for social cue recognition and nonverbal behavior analysis</li> </ul>	<ul style="list-style-type: none"> <li>• Virtual human displayed on large screen monitor with synchronized behaviors</li> </ul>
Public Speaking Simulator (R&D prototype)	[52]	Public speaking skills (no training)	<ul style="list-style-type: none"> <li>• Facial expression</li> <li>• Body positions</li> <li>• Human speech inputs</li> <li>• Camera for facial expression recordings</li> <li>• Kinect for captures of speaker’s body position – provided information and control of audience behaviors</li> </ul>	<ul style="list-style-type: none"> <li>• Interactive virtual audience displayed on multiple large screens with varied attentive and non-attentive behaviors</li> </ul>
My Automated Conversation Coach (MACH) (R&D prototype)	[53]	Job interview skills (no training)	<ul style="list-style-type: none"> <li>• Speech recognition</li> <li>• Facial expression – smiles</li> <li>• Head positions</li> <li>• Voice prosody</li> <li>• Camera for analysis of video to determine and interpret user’s nonverbal behavior</li> </ul>	<ul style="list-style-type: none"> <li>• Virtual human displayed on large screen monitor with synchronized behaviors</li> <li>• Synthesized VH voice</li> </ul>
Automated Social Skills Trainer (R&D prototype)	[54]	Anxieties associated with social interactions	<ul style="list-style-type: none"> <li>• Speech recognition</li> <li>• Voice volume, rate, pauses, and quality</li> <li>• Camera for recording of user’s interaction for playback during after action review</li> </ul>	<ul style="list-style-type: none"> <li>• Virtual human displayed on monitor with synchronized behaviors</li> </ul>

## V. Assessment And Discussion About Conversational Virtual Humans

Using the ten instantiated conversational virtual human systems identified in Table 2, Table 3 reports an assessment of each system's ability to develop or improve a particular social skill within a given social encounter, acceptance of the virtual human as a role-player by the person or people engaged in the social encounter, and usability concerns with the system.

**TABLE 33: ASSESSMENT OF REPORTED CONVERSATIONAL VIRTUAL HUMAN SYSTEMS**

System Name	Social Skills Objectives and Performance Improvement	Acceptance of the Virtual Human	Usability Concerns with System
SimSensi "Ellie" [75]	VH conducts interview and healthcare screening; found willingness to disclose information without feeling of judgement (N=239; ages 18-65 yrs)	Acceptance reported through anecdotal remarks	No usability issues were reported Continuing to conduct R&D with goal of prototype kiosk use in hospital
VR-JIT [42]	Job interview training; demonstrated trend-level improvements interview content and performance (N=26; ages 18-31 showing ASD tendencies)	90% attendance of multiple sessions/interactions and increased feeling of confidence	Reported as easy to use, helpful, and enjoyable
AVP [41]	Use of appropriate reciprocity skills during social interactions may increase with the use of AVP (N=7; children with ASD part of 11 week intervention program)	Anecdotal comments suggest the AVP was engaging and a relevant component of the program	No reported issues with usability
BiLat [39]	Bilateral meeting and negotiation skills in a cultural context showing best results with users that had no previous negotiation experience (N=31; Army soldiers)	Release for Army wide training in 2009; while focus is on middle eastern cultures, it is still available for use	Anecdotal comments stating that user's would prefer to speak or type responses opposed to menu selections
TLCTS [4]	4 hours of language and cultural training including VR practice environment; self reported benefit to taking training prior to deployment (N=21; Marines who completed both pre-deployment training and post-deployment surveys)	Self report general positive reactions; aspects incorporated into new company products centered on learning languages and cultures within virtual role-playing environments	No issues reported
INOTS & ELITE [45]	Interpersonal communication skills focused on personal and performance counseling; reported no measurable physiological differences between interactions with live and virtual role-player (N=21, Naval Reserve Officer Training Corps at UCLA)	Virtual role-player slightly less expressive than real human actor; Lack of ability to sustain eye contact with user potentially impacted the immersive experience	No issues reported
Job Interview Simulator [51]	Interaction with virtual job recruiter to prepare for job interviews; participants stated in post-hoc interviews that system was helpful in preparing them for job interviews (N=6, students with avg age of 28.83)	Most participants felt that virtual recruiter could sense their feelings and behaviors and half felt the recruiter was interested in them. Reported limitations and future improvements include detection of social cues from voice features, eye gaze, and laughter	Positive attitude towards system
Public Speaking Simulator [52]	Practicing public speaking skills; system showed improved skills or reduced anxiety from self, expert, and objective assessment (N=47; avg age reported 37; results from 3 perspectives user, expert, and objective assessment)	Users were eager to use the system	System was easy to use and non-threatening
MACH [53]	Job interview training; Interacting with the virtual role-player resulted in statistically significant performance improvement (N=90 college juniors)	Reported to likely use the system again	Participants reported that the system is easy to use
Automated Social Skills Trainer [54]	General social skills training aimed at reducing anxiety and discomfort in social situations; results showed improvement in narrative skills after interaction with virtual human (N=30 Japanese graduated students)	Character's behaviors appeared natural and human-like	Rated as easy to use and participants reported they would use the system frequently

Table 3 indicates that virtual humans may be suitable role-players for various social skills training encounters. [45, 4, 36] indicate that people accept virtual humans in the roles, but virtual human technology designs often do not integrate or fail to address the bi-directional nature of the person-virtual human interaction. More generally, virtual human role-player limitations include inability to assess the user's behaviors and adapt appropriately [51, 45]. During typical human face-to-face communication, participants use message content and bi-directional behaviors such as eye contact, facial expressions, and hand and body gestures [43, 55, 18] as cues to maintain and adapt their responses within a conversation. Eye contact is an important element of many social skills [44] and may indicate a person's level of attention that they have during a conversation [56]. Less than half of the applications in Table 3 include a virtual human with some form of visual perception of their social partner.

With the lack of sensory inputs to provide the virtual human a perception of the visual communications channels, designers of social skills training systems need to realize the limitations of the virtual human presented to the person engaged in the simulated social interaction. Applications without even a simple visual system or camera to determine if the human user is looking at the virtual human reflect designer's choice to ignore the importance of visual attention or eye contact in a person-to-person, face-to-face encounter.

Additionally training applications often include with the virtual human role-player textual and supporting instructional information that would not necessarily be found in the real world. Previous research related to multimedia approaches for training and education provide insight into the impact of nonverbal information on training with virtual humans. Mayer concluded from his research using multimedia learning techniques that "people learn more deeply from words and pictures than words alone" [57]. Schroeder et al.

[1] concluded that agents in pedagogical roles that communicate with learners via on-screen text are more effective. One might then assume that the additional textual display of the dialogue is beneficial to the learner in the case of virtual human training applications [58]. In the case of social skills training where behaviors such as eye contact are important,

textual display may create a conflict for the learner in that the display contains two different visual focal points and potentially competitive visual sources of information. For social skills training, the presentation of an animated and vocalized virtual human with on-screen dialogue text may split the learners' attention or confound focus of attention on text at the expense of the virtual human. Since verbal and visual attention to a person being conversed with is important to understanding spoken and unspoken messages within a face-to-face conversation, one might question the value of the on-screen text which adds redundancy and could potentially lower performance [58-60]. For example, TLCTS, BiLat, INOTS/ELITE, and VR-JIT interfaces display the dialogue between the virtual human and human in text form in addition to the virtual human's audible voice and animated nonverbal behaviors. While the text may accommodate those with a reader bias, the issue raised is whether the additional information distracts from or possibly undermines learning the intended objective of interacting with a virtual role-player in a simulation of a real-world environment where text is not present. If the purpose of the training is to provide the learner with the skills necessary to interact effectively in a real-world social situation, then gathering meaning from spoken words and unspoken "body language" of the virtual person are what are important to determining response. The level of importance of observing virtual human nonverbal behavior in training applications remains an empirical question. Further research is needed to explore whether during a simulated social encounter the addition of a textual information not found in a real world encounter will enhance or detract from learning during a simulation social interaction.

## VI. Conclusions, Limitations And Future Directions

This research expands the Schroeder, Adesope, Gilbert Feature Framework for Virtual Humans to virtual humans in role playing and social interaction skills training by adding artificial intelligence with a focus on the virtual human's ability to see and perceive world around them. Analysis of findings in Table 2 and 3 indicates the impact in a social skills setting of virtual humans conversational attributes such as message content, verbal expression, and nonverbal behavior such as head position or inclination, micro and macro facial or body emotional expression, eye gaze and direction, and other facial and other bodily gestures. Discussion addresses the question: can a virtual human truly be a social partner with a person for the purpose of social skills training? In response to that question, virtual humans have progressed significantly from ELIZA [20] to SimSensi and MACH [46, 53]. Virtual humans have gone from being a novelty to being an integral part of a training applications as instructors, mentors, coaches [61, 62] to participating in simulated role-playing exercises for training social skills [4, 10, 11, 39]. Based on the role playing virtual humans analyzed, this research identified a visual system as an important virtual human attribute to support real-time conversational dialog with the ability to handle interruptions as well as synchronized and bi-directional handling of head position, emotions, eye gaze direction, and facial and other bodily gestures. One concern identified in the research is that for social skills training, the presentation of an animated and vocalized virtual human with on- screen dialogue text may split the learners' attention or confound focus of attention on text at the expense of the virtual human. Limiting this assessment is lack of empirical clarity when the virtual human is part of a learning system containing significant supporting information such as text boxes. The level of contribution of virtual human micro-expressions [63] and verbal tones [64] remain current topics in human-virtual human engagement in role playing research. Virtual human research with the previously mentioned SimSensi and MACH, wearable agents [65, 66], virtual peers and companions [67, 68] and virtual patients [69, 70] may provide further insights.

From the perspective of bi-directional conversation, none of the virtual humans discussed above are cloud-based. Cloud based agents Watson, Siri, and Cortana raise the possibility of overcoming limitations of local hardware solutions to the challenge of maintaining intelligent conversational content and dialog by accessing the computational power of the cloud in real time. To raise Watson to the level of a Jeopardy grand champion also required extensive training involving scores of scenarios [71, 72]. Beyond these concerns virtual assistants Siri and Cortana, not only lack a human face possibly necessary for social skills training, they are challenged by the user's current situation such as the user's physical and emotional states. The awareness from physiological sensors (blood pressure, heart rate), and current user task or situation (walking, sleeping, driving, at home, at work) may aid future virtual assistant to adopt social skills in a meaningful and relevant conversation [66]. This requires full utilization of current and new sensors and applications create a new level of awareness for virtual humans that only begins to model human conversation. Whether selected role playing virtual humans will be able to garner the resources to achieve such conversational capabilities may likely be based on market demand.

Limiting the above analysis is insight into how humans visually attend to the virtual human or what might cause disengagement. Chen and Wu [73] suggest that an effective approach to understanding attention, cognitive load, and overall learner performance is to study the visual attention towards an area of interest.

D'Mello, Olney, Williams & Hays [74] identified periods of engagement and boredom during interactions with an intelligent tutor from the use of eye trackers. They concluded that strategies could be employed to increase the attention of the trainee but more research is needed to evaluate different strategies. A better understanding of visual area interest and periods of engagement and boredom during interaction may provide a greater insight into how humans fully interact with virtual humans and if a virtual human can truly be a social partner.

## References

- [1]. Schroeder, N. L., Adesope, O. O., & Gilbert, R. B. (2013). How Effective are Pedagogical Agents for Learning? A Meta-Analytic Review. *Journal of Educational Computing Research*, 49(1), 1-39.
- [2]. Holsbrink-Engels, G. A. (2001). Using a computer learning environment for initial training in dealing with social-

- communicative problems. *British Journal of Educational Technology*, 32(1), 53–67.
- [3]. Irish, J. E. N. (2013). Can I sit here? A review of the literature supporting the use of single-user virtual environments to help adolescents with autism learn appropriate social communication skills. *Computers in Human Behavior*, 29(5), A17–A24.
- [4]. Johnson, W. L. (2014). Using Virtual Role-Play to Prepare for Cross-Cultural Communication. In T. Ahram, W. Karwowski, & T. Marek (Eds.), *Proceedings of the 5th International Conference on Applied Human Factors and Ergonomics AHFE 2014*. Krakow, Poland.
- [5]. Lane, H. C. (2010). Individualized Virtual Humans for Social Skills Training 2 Towards Individualized Practice with Virtual Humans. *AI Magazine*, 273–274.
- [6]. Ogan, A., Aleven, V., Kim, J., & Jones, C. (2010). Intercultural negotiation with virtual humans: The effect of social goals on gameplay and learning. In *Intelligent Tutoring Systems* (pp. 174–183). Springer Berlin Heidelberg.
- [7]. Parsons, S., & Mitchell, P. (2002). The potential of virtual reality in social skills training for people with autistic spectrum disorders. *Journal of Intellectual Disability Research*, 46(5), 430–43.
- [8]. Tartaro, A., Cassell, J., Ratz, C., Lira, J., & Nanclares-Nogués. (2014). Accessing Peer Social Interaction: Using Authorable Virtual Peer Technology as a Component of a Group Social Skills Intervention Program. *ACM Transactions on Accessible Computing*, 6(1), 2–29.
- [9]. Gratch, J., Rickel, J., Andre, E., Cassell, J., Petajan, E., & Badler, N. (2002). Creating Interactive Virtual Humans: Some Assembly Required. *IEEE Intelligent Systems*, (July/August 2002).
- [10]. Campbell, J., Hays, M. J., Core, M., Birch, M., Bosack, M., & Clark, R. E. (2011). Interpersonal and Leadership Skills: Using Virtual Humans to Teach New Officers. In *Interservice/Industry Training, Simulation, and Education Conference (IITSEC) 2011*. Orlando, FL.
- [11]. Frank, G., Guinn, C., Hubal, R., Carolina, N., Pope, P., Stanford, M., & Lamm-weisel, D. (2002). JUST- TALK: An Application of Responsive Virtual Human Technology. In *Proceedings of Interservice/Industry Training, Simulation and Education Conference*. Orlando, FL.
- [12]. Mykoniatis, K., Angelopoulou, A., Proctor, M. D., & Karwowski, W. (2014). Virtual Humans for Interpersonal and Communication Skills ' Training in Crime Investigations. In *Virtual, Augmented and Mixed Reality. Designing and Developing Virtual and Augmented Environments* (pp. 282–292). Switzerland: Springer International Publishing.
- [13]. Sotomayor, T. M., & Proctor, M. D. (2009). Assessing Combat Medic Knowledge and Transfer Effects Resulting from Alternative Training Treatments. *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, 6(3), 121–134.
- [14]. Talbot, T. B., Sagae, K., John, B., & Rizzo, A. (2012). Designing Useful Virtual Standardized Patient Encounters. In *Interservice/Industry Training, Simulation, and Education Conference (IITSEC) 2012*.
- [15]. Johnson, W. L. (2010). Serious Use of a Serious Game for Language Learning. *International Journal of Artificial Intelligence in Education*, 20(2), 175–195.
- [16]. Proctor, M., Cardona, G., & Hollander, C. (2008). Using Simulation to investigate a Non-Anthropomorphic Framework for Communications within a Human-Agent War-Fighting Team. *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, 5(2), 122–138.
- [17]. McFall, R. M. (1982). A review and reformation of the concept of social skills. *Behavioral Assessment*, 4(1), 1–33.
- [18]. Segrin, C., & Givertz, M. (2003). Methods of Social Skills Training and Development. In J. O. Greene & B. R. Burlesono (Eds.), *Handbook of Communication and Social Interaction Skills*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- [19]. Stamp, G. H. (1999). A qualitatively constructed interpersonal communication model - A grounded theory analysis. *Human Communication Research*, 25(4), 531–547.
- [20]. Weizenbaum, J. (1966). ELIZA--A Computer Program For the Study of Natural Language Communication Between Man and Machine. *Communications of the ACM*, 9(1), 36–45.
- [21]. Whitworth, B. (2005). Polite computing. *Behaviour & Information Technology*, 24(5), 353–363.
- [22]. Sproull, L., Subramani, M., Kiesler, S., Walker, J. H., & Waters, K. (1996). When the Interface Is a Face. *Human-Computer Interaction*, 11, 97–124.
- [23]. Takeuchi, A., & Naito, T. (1995). Situated Facial Displays: Towards Social Interaction. In I. Katz, R. Mack, L. Marks, M. B. Rosson, & J. Nielsen (Eds.), *Human Factors in Computing Systems: CHI'95 Electronic Conference Proceedings*. ACM.
- [24]. Graesser, A. C., Lu, S., Jackson, G. T., Mitchell, H. H., Ventura, M., Olney, A., & Louwerse, M. M. (2004). AutoTutor: a tutor with dialogue in natural language. *Behavior Research Methods, Instruments, & Computers : A Journal of the Psychonomic Society, Inc*, 36(2), 180–192.
- [25]. Yee, N., Bailenson, J. N., & Rickertsen, K. (2007). A meta-analysis of the impact of the inclusion and realism of human-like faces on user experiences in interfaces. In *Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '07* (p. 1). San Jose, CA: ACM Press.
- [26]. Garau, M., Slater, M., Vinayagamoorthy, V., Brogni, A., Steed, A., & Sasse, M. A. (2003). The impact of avatar realism and eye gaze control on perceived quality of communication in a shared immersive virtual environment. In *Proceedings of the conference on Human factors in computing systems - CHI '03* (p. 529). New York, New York, USA: ACM Press.
- [27]. Beale, R., & Creed, C. (2009). Affective interaction: How emotional agents affect users. *International Journal of Human-Computer Studies*, 67(9), 755–776.

- [28]. De Melo, C. M., Carnevale, P., & Gratch, J. (2010). The Influence of Emotions in Embodied Agents on Human Decision-Making. In *Intelligent Virtual Agents 2010* (pp. 357–370).
- [29]. Bickmore, T., & Cassell, J. (2001). Relational Agents: A Model and Implementation of Building User Trust. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 396–403). New York.
- [30]. Gratch, J., Okhmatovskaia, A., Lamothe, F., Marsella, S., Morales, M., Werf, R. J. Van Der, & Morency, L. (2006). Virtual Rapport. In *6th International Conference on Intelligent Virtual Agents*. Marina del Rey, CA.
- [31]. Bailenson, J. N., Blascovich, J., Beall, A. C., & Loomis, J. M. (2003). Interpersonal Distance in Immersive Virtual Environments. *Personality and Social Psychology Bulletin*, 29(7), 1–15.
- [32]. Bickmore, T., Gruber, A., & Picard, R. (2005). Establishing the computer-patient working alliance in automated health behavior change interventions. *Patient Education and Counseling*, 59(1), 21–30.
- [33]. Cassell, J., Stocky, T., Bickmore, T., Gao, Y., Nakano, Y., Ryokai, K., ... Vilhjálmsón, H. (2002). MACK: Media lab Autonomous Conversational Kiosk. In *Proceedings of IMAGINA '02*. Monte Carlo.
- [34]. Gratch, J., Wang, N., Gerten, J., Fast, E., & Duffy, R. (2007). Creating Rapport with Virtual Agents. In *International Conference on Intelligent Virtual Agents*. Paris, France.
- [35]. Cassell, J., Bickmore, T., Billingham, M., Leaper, C., Chang, K., Vilhjálmsón, H., & Yan, H. (1999). Embodiment in Conversational Interfaces: Rea. In *SIGCHI conference on Human Factors in Computing Systems* (pp. 520–527).
- [36]. Kopp, S., Gesellensetter, L., Krämer, N. C., & Wachsmuth, I. (2005). A conversational agent as museum guide - Design and evaluation of a real-world application. In *Intelligent Virtual Agents 2005* (pp. 329–343).
- [37]. Bickmore, T. W., Vardoulakis, L. M. P., & Schulman, D. (2013). Tinker: A relational agent museum guide. *Autonomous Agents and Multi-Agent Systems*, 27(2), 254–276.
- [38]. Swartout, W., Traum, D., Artstein, R., Noren, D., Debevec, P., Bronnenkant, K., ... White, K. (2010). Ada and Grace: Toward Realistic and Engaging Virtual Museum Guides. In *Intelligent Virtual Agents 2010* (pp. 286–300).
- [39]. Kim, J. M., Hill, R. W., Technologies, C., Lane, H. C., Forbell, E., Core, M., ... Hart, J. (2009). BiLAT: A Game-Based Environment for Practicing Negotiation in a Cultural Context. *International Journal of Artificial Intelligence in Education*, 19, 289–308.
- [40]. Baur, T., Damian, I., Gebhard, P., Porayska-Pomsta, K., & Andre, E. (2013). A Job Interview Simulation: Social Cue-based Interaction with a Virtual Character. In *2013 ASE/IEEE International Conference on Social Computing (Socialcom 2013)*.
- [41]. Tartaro, A., Cassell, J., Ratz, C., Lira, J., & Nanclares-Nogués. (2014). Accessing Peer Social Interaction: Using Authorable Virtual Peer Technology as a Component of a Group Social Skills Intervention Program. *ACM Transactions on Accessible Computing*, 6(1), 2–29.
- [42]. Smith, M. J., Ginger, E. J., Wright, M., Wright, K., Boteler Humm, L., Olsen, D., ... Nanclares-Nogués. (2014). Virtual Reality Job Interview Training for Individuals With Psychiatric Disabilities. *The Journal of Nervous and Mental Disease*, 202(5), 659–667.
- [43]. Argyle, M., & Cook, M. (1976). *Gaze and Mutual Gaze*. Oxford, England: Cambridge University Press.
- [44]. Argyle, M., & Dean, J. (1965). Eye-Contact, Distance and Affiliation. *Sociometry*, 28(3), 289–304.
- [45]. Hays, M. J., Campbell, J. C., Trimmer, M. A., Poore, J. C., & Webb, A. K. (2012). Can Role-Play with Virtual Humans Teach Interpersonal Skills? In *Interservice/Industry Training, Simulation, and Education Conference (IITSEC) 2012*. Orlando, FL.
- [46]. Devault, D., Artstein, R., Benn, G., Dey, T., Fast, E., Gainer, A., ... Morency, L. (2014). SimSensei Kiosk: A Virtual Human Interviewer for Healthcare Decision Support. In *13th International Conference on Autonomous Agents and Multi-Agent Systems, International Foundation for Autonomous Agents and Multiagent Systems* (pp. 1061–1068).
- [47]. Morency, L., Stratou, G., Devault, D., Hartholt, A., Lhommet, M., Lucas, G., ... Rizzo, A. (2015). SimSensei Demonstration: A Perceptive Virtual Human Interviewer for Healthcare Applications. In *Twenty-Ninth AAAI Conference on Artificial Intelligence*.
- [48]. Qu, C., Brinkman, W. P., Ling, Y., Wiggers, P., & Heynderickx, I. (2014). Conversations with a virtual human: Synthetic emotions and human responses. *Computers in Human Behavior*, 34, 58–68.
- [49]. Smith, M. J., Ginger, E. J., Wright, K., Wright, M. A., Taylor, J. L., Humm, L. B., ... Fleming, M. F. (2014). Virtual reality job interview training in adults with autism spectrum disorder. *Journal of Autism and Developmental Disorders*, 44(10), 2450–2463.
- [50]. Johnson, W. L., & Valente, A. (2008). Tactical Language and Culture Training Systems: Using Artificial Intelligence to Teach Foreign Languages and Cultures. In *AAAI* (pp. 1632–1639).
- [51]. Baur, T., Damian, I., Gebhard, P., Porayska-Pomsta, K., & Andre, E. (2013). A Job Interview Simulation: Social Cue-based Interaction with a Virtual Character. In *2013 ASE/IEEE International Conference on Social Computing (Socialcom 2013)*.
- [52]. Chollet, M., Morency, L., Shapiro, A., Scherer, S., & Angeles, L. (2015). Exploring Feedback Strategies to Improve Public Speaking: An Interactive Virtual Audience Framework, 1143–1154.
- [53]. Hoque, M., Courgeon, M., & Martin, J. (2013). Mach: My automated conversation coach. In *Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing* (pp. 697–706).
- [54]. Tanaka, H., Sakti, S., Neubig, G., Toda, T., Negoro, H., Iwasaka, H., & Nakamura, S. (2015). Automated Social Skills Trainer. In *Proceedings of the 20th International Conference on Intelligent User Interfaces* (pp. 17–27).
- [55]. Cassell, J. (2000). Nudge Nudge Wink Wink: Elements of Face-toFace Conversation for Embodied Agents. In J.

- Cassell, J. Sullivan, S. Prevost, & E. Churchill (Eds.), *Embodied Conversational Agents*. Cambridge, MA: MIT Press.
- [56]. Frischen, A., Bayliss, A. P., & Tipper, S. P. (2007). Gaze cueing of attention: visual attention, social cognition, and individual differences. *Psychological Bulletin*, 133(4), 694–724.
- [57]. Mayer, R. E. (2005). Cognitive Theory of Multimedia Learning. In *The Cambridge Handbook of Multimedia Learning* (pp. 31–48). Cambridge, MA: Cambridge University Press.
- [58]. Mayer, R. E., & Johnson, C. I. (2008). Revising the redundancy principle in multimedia learning. *Journal of Educational Psychology*, 100(2), 380–386.
- [59]. Mayer, R. E., Heiser, J., & Lonn, S. (2001). Cognitive constraints on multimedia learning: When presenting more material results in less understanding. *Journal of Educational Psychology*, 93(1), 187–198.
- [60]. Kalyuga, S., Chandler, P., & Sweller, J. (1998). Levels of Instructional Design. *Human Factors*, 40(1), 1–17.
- [61]. Baylor, A. L., & Kim, Y. (2005). Simulating Instructional Roles through Pedagogical Agents. *International Journal of Artificial Intelligence in Education*, 15(5), 95–115.
- [62]. Veletsianos, G., & Russell, G. S. (2014). Pedagogical Agents. In *Handbook of research on educational communications and technology* (pp. 759–769). New York: Springer.
- [63]. Queiroz, R. B., Musse, S. R., & Badler, N. I. (2014). Investigating Macroexpressions and Microexpressions in Computer Graphics Animated. *Presence*, 23(2), 191–208.
- [64]. Kim, Y., & Baylor, A. L. (2015). Research-Based Design of Pedagogical Agent Roles: a Review, Progress, and Recommendations. *International Journal of Artificial Intelligence in Education*, 25.
- [65]. Kang, S., Feng, A. W., Leuski, A., Casas, D., & Shapiro, A. (2015). Smart Mobile Virtual Humans: “ Chat with Me !” In *Intelligent Virtual Agents 2015*.
- [66]. Riccardi, G. (2014). Towards Healthcare Personal Agents. In *RFMIR '14* (pp. 53–56). Istanbul, Turkey.
- [67]. Bickmore, T., Asadi, R., Ehyaei, A., Fell, H., Henault, L., Intille, S., ... Paasche-orlow, M. K. (2015). Context-Awareness in a Persistent Hospital Companion Agent. In *Intelligent Virtual Agents 2015*.
- [68]. Finkelstein, S., Ogan, A., Vaughn, C., & Cassell, J. (2013). Alex: A Virtual Peer that Identifies Student Dialect. In *Proceedings of Workshop on Culturally-Aware Technology Enhanced Learning*. Paphos, Cyprus.
- [69]. Bartley, E. J., Boissoneault, J., Vargovich, A. M., Wandner, L. D., Hirsh, A. T., Lok, B. C., ... Robinson, M. E. (2015). The Influence of Health Care Professional Characteristics on Pain Management Decisions, 99–111.
- [70]. Kleinsmith, A., Rivera-Gutierrez, D., Finney, G., Cendan, J., & Lok, B. (2015). Understanding empathy training with virtual patients. *Computers in Human Behavior*, 52, 151–158.
- [71]. Ferrucci, D., Brown, E., Chu-Carroll, J., Fan, J., Gondek, D., Kalyanpur, A. a., ... Welty, C. (2010). Building Watson: An Overview of the DeepQA Project. *AI Magazine*, 31(3), 59–79.
- [72]. Shah, H. (2011). Turing’s misunderstood imitation game and IBM’s Watson success. In *AISB 2011 Convention* (pp. 1–5). University of York.
- [73]. Chen, C.-M., & Wu, C.-H. (2015). Effects of different video lecture types on sustained attention, emotion, cognitive load, and learning performance. *Computers & Education*, 80, 108–121.
- [74]. D’Mello, S., Olney, A., Williams, C., & Hays, P. (2012). Gaze tutor: A gaze-reactive intelligent tutoring system. *International Journal of Human Computer Studies*, 70(5), 377–398.
- [75]. Lucas, G. M., Gratch, J., King, A., & Morency, L.-P. P. (2014). It’s only a computer: Virtual humans increase willingness to disclose. *Computers in Human Behavior*, 37, 94–100.
- [76]. Wooldrige, M. (2002). *An Introduction to Multiagent Systems*. New York: John Wiley & Sons, Inc.