Clinical interviewing by a virtual human agent with automatic behavior analysis

A A Rizzo, G Lucas, J Gratch, G Stratou, L-P Morency, R Shilling, A Hartholt, S Scherer

Institute for Creative Technologies, University of Southern California,
12015 East Waterfront Dr., Playa Vista, Los Angeles, California, USA

rizzo@ict.usc.edu, lucas@ict.usc.edu, gratch@ict.usc.edu, stratou@ict.usc.edu, morency@ict.usc.edu, hartholt@ict.usc.edu, scherer@ict.usc.edu

http://ict.usc.edu/

ABSTRACT

SimSensei is a Virtual Human (VH) interviewing platform that uses off-the-shelf sensors (i.e., webcams, Microsoft Kinect and a microphone) to capture and interpret real-time audiovisual behavioral signals from users interacting with the VH system. The system was specifically designed for clinical interviewing and health care support by providing a face-to-face interaction between a user and a VH that can automatically react to the inferred state of the user through analysis of behavioral signals gleaned from the user's facial expressions, body gestures and vocal parameters. Akin to how non-verbal behavioral signals have an impact on human-to-human interaction and communication, SimSensei aims to capture and infer user state from signals generated from user non-verbal communication to improve engagement between a VH and a user and to quantify user state from the data captured across a 20 minute interview. As well, previous research with SimSensei indicates that users engaging with this automated system, have less fear of evaluation and self-disclose more personal information compared to when they believe the VH agent is actually an avatar being operated by a "wizard of oz" human-in-the-loop (Lucas et al., 2014). The current study presents results from a sample of military service members (SMs) who were interviewed within the SimSensei system before and after a deployment to Afghanistan. Results indicate that SMs reveal more PTSD symptoms to the SimSensei VH agent than they self-report on the Post Deployment Health Assessment. Pre/Post deployment facial expression analysis indicated more sad expressions and fewer happy expressions at post deployment.

1. INTRODUCTION

Over the last 20 years, a gradual revolution has taken place in the use of Virtual Reality (VR) simulation technology for clinical purposes. When discussion of the potential use of VR applications for human research and clinical intervention first emerged in the early 1990s, the technology needed to deliver on this "vision" was not in place. Consequently, during these early years VR suffered from a somewhat imbalanced "expectation-to-delivery" ratio, as most who explored VR systems during that time will attest. However, in the last few years the accelerating pace technology development has now led to the creation of VR systems that have caught up with the vision for its clinical use as articulated by scientists in the 1990's. Dramatic advances in the underlying VR enabling technologies (e.g., computational speed, 3D graphics rendering, audio/visual/haptic displays, user interfaces/tracking, voice recognition, intelligent agents, and authoring software, etc.) have supported the creation of low-cost, yet sophisticated VR systems capable of running on commodity level personal computers, and even on mobile devices like the Samsung Gear VR. In part driven by the digital gaming and entertainment sectors, and a near insatiable global demand for mobile and networked consumer products, such advances in technological "prowess" and accessibility have provided the hardware and software platforms needed to produce more usable and hi-fidelity VR scenarios for the conduct of human research and clinical intervention. Thus, evolving behavioral health applications can now usefully leverage the interactive and immersive assets that VR affords as the technology continues to get faster, better and cheaper moving into the 21st Century. While some of this may be due to the commercial enthusiasm generated by the Facebook purchase of Oculus Rift for 2 billion dollars, it is more likely that the highly publicized advances in the enabling technologies for delivering low cost VR simulations has sparked renewed public awareness and enchantment with VR. As well, a solid scientific literature has evolved documenting the value of applying simulation technology to usefully study and address the needs of people with a wide range of clinical health conditions.

©2016 ICDVRAT; ISBN 978-0-7049-1547-3

These advances in the technical and scientific landscape have now set the stage for the next major movement in Clinical Virtual Reality with the "birth" of intelligent virtual human (VH) agents. This has been driven by seminal research and development leading to the creation of highly interactive, artificially intelligent, and natural language capable VHs that can engage real human users in a credible fashion (Swartout et al. 2013). Such intelligent VH agents have been created that control computer generated bodies and can interact with users through natural language speech and gesture in virtual environments (Gratch et al., 2002, 2013; Rizzo, Kenny & Parsons, 2011; Rizzo & Talbot, 2016a; Talbot et al., 2012). Advanced VH agents can engage in rich conversations (Traum et al., 2008), recognize nonverbal cues (Morency et al., 2008; Scherer et al., 2014; Rizzo et al., 2015, 2016ab), improve interactional rapport with users (Park et al., 2013) reason about social and emotional factors (Gratch and Marsella, 2004), and synthesize human communication and nonverbal expressions (Thiebaux et al., 2008). Such efforts have led to the creation of VH systems across a number of fields, including education, clinical training, clinical assessment and providing healthcare guidance. These findings have motivated R&D in our lab focused on the development of VH agent systems that serve as: virtual patients for training novice clinicians (Talbot et al., 2012; Rizzo et al., 2011, 2016a), clinical interviewers to reduce stigma (Rizzo et al., 2015), and as health care guides and clinical support agents (Rizzo et al. 2015).

In the area of clinical assessment, VH's can conduct clinically-oriented interviews within a safe non-judgmental context which may encourage honest disclosure of important information. In a recent study, users reported less concern about being evaluated and disclosed and displayed more sadness in an interview with a VH agent compared to interacting with a VH avatar that they believed was being operated by a human-in-the-loop "Wizard of Oz" controller (Lucas et al., 2014). VH agents have also been shown to promote engagement and longer participation with users asked to answer general questions posed on a mobile application compared to voice only and voice plus static VH image conditions (Kang et al., 2014) and this is in line with other studies indicating higher levels of personal self-discloser when VH agents are used in this form (Kang et al., 2012, 2013). Finally, users interacting with a conversational VH interface to access online health information were more satisfied compared to a conventional Web form-based interface and users with low health literacy were more successful (and satisfied) in their capacity to find information on available clinical trials when supported by a VH agent (Bickmore et al., 2016). Such findings converge to support the idea that users will disclose more and may also have enhanced success in accessing information when interacting with a VH-supported healthcare application.

The incorporation of a VH within the application detailed in the current paper is intended to amplify the effects reported in earlier studies that suggest that computer-mediated interviews are felt to be more anonymous than face-to-face interviews and this resultant anonymity leads to increased disclosure (Weisband & Kiesler, 1996; Baker, 1992; Beckenbach, 1995; Joinson, 2001; Sebestik, Zelon, DeWitt, O'Reilly & McGowan, 1988; van der Heijden, Van Gils, Bouts, & Hox, 2000). The impact of computer-administered VH interviews on the enhancement of disclosure is particularly relevant in health contexts due to the intimate nature of the information revealed. Effects are strongest when the information is illegal, unethical, or culturally stigmatized (e.g., drug use, unsafe sex, suicidal ideation) (Weisband & Kiesler, 1996; van der Heijden, Van Gils, Bouts, & Hox, 2000), which is critical information to disclose in health settings and where a patient's failure to provide honest and adequately detailed responses in medical interviews could lead to serious negative health outcomes when vital information is withheld.

This paper will detail our efforts in the creation of a VH who can serve in the role of a clinical interviewer (i.e., SimSensei) while also using camera and audio sensors to automatically detect behavioral signals from which user state may be inferred. The SimSensei system was specifically designed for clinical interviewing and health care support by providing a face-to-face interaction between a user and a VH that can automatically react to the inferred state of the user through analysis of behavioral signals gleaned from the user's facial expressions, body gestures and vocal parameters. User behavior is captured and quantified using a range of off-the-shelf sensors (i.e., webcams, Microsoft Kinect and a microphone). Akin to how non-verbal behavioral signals have an impact on human-to-human interaction and communication, SimSensei aims to capture and infer user state from signals generated from user non-verbal communication to improve engagement between a VH and a user. The system also can quantify and interpret sensed behavioral signals longitudinally for use to inform diagnostic assessment within a clinical context.

The development of SimSensei required a thorough awareness of the literature on emotional expression and communication. It has long been recognized that facial expression and body gestures play an important role in human communicative signaling (Ekman & Rosenberg, 1997). As well, vocal characteristics (e.g., prosody, pitch variation, etc.) have been reported to provide additive information regarding the "state" of the speaker beyond the actual language content of the speech (Pentland et al., 2009). Pentland (2008) has characterized these elements of behavioral expression as "Honest Signals" and posits that the physical properties of this signaling behavior are constantly activated, not simply as a back channel or complement to our conscious language, but rather as a separate communication network. It is conjectured that these signaling behaviors, perhaps evolved from ancient primate non-verbal communication mechanisms, provide a useful window into our intentions, goals, values and emotional state. From this perspective, an intriguing case can be made for the development of a computer-based sensing system that can capture and quantify such behavior, and using that data, make inferences as to a user's cognitive and emotional state. Inferences from these sensed signals could then be used to supplement information that is garnered exclusively from the literal content of speech.

Recent progress in low cost sensing technologies and computer vision methods has now driven this vision to reality. Indeed, recent widespread availability of low cost sensors (webcams, Microsoft Kinect, microphones) combined with software advances for facial feature tracking, articulated body tracking, and voice analytics (Baltrusaitis et al., 2012; Morency et al., 2008; Whitehill et al. 2009) has opened the door to new applications for automatic nonverbal behavior analysis. This sensing, quantification and inference from nonverbal behavioral cues can serve to provide input to an interactive virtual human interviewer that can respond with follow-up questions that leverage inferred indicators of user distress or anxiety during a short interview. This is the primary concept that underlies the SimSensei interviewing agent (See Figure 1). The SimSensei capability to accomplish this is supported by the "MultiSense" perception system (Morency, 2010; Devault et al., 2014; Scherer et al., 2014), a multimodal system that allows for real-time synchronized capture, tracking, and fusion of behavioral markers of different modalities such as audio as well as visual. MultiSense's fusion enables the analysis of complex behavioral indicators of user states across multiple modalities. Within SimSensei, MultiSense fuses information from a web camera, Microsoft Kinect and audio capture to identify the presence of predetermined nonverbal indicators of psychological distress. Dynamic capture and quantification of behavioral signals are used such as 3D head position and orientation, type, intensity and frequency of facial expressions of emotion (e.g., fear, anger, disgust and joy), fidgeting, slumped body posture, along with a variety of speech parameters (e.g., speaking fraction, latency to respond). These informative behavioral signals serve two purposes. First, they produce the capability of analyzing the occurrence and quantity of behaviors to inform detection of psychological state. Second, they are broadcast to other software components of the SimSensei system to inform the VH interviewer of the state and actions of the participant. This information is then used by the VH to assist with turn taking, rapport building (e.g., utterances, acknowledging gestures/facial expressions), and to drive and deliver follow-on questions.

Figure 1. User with SimSensei virtual clinical interviewer.

SimSensei is one application component developed from the DARPA-funded "Detection and Computational Analysis of Psychological Signals (DCAPS)" project. This DCAPS application has aimed to explore the feasibility of creating "empathic" virtual human health agents for use as clinical interviewers and to aid in mental health screening. The system seeks to combine the advantages of traditional web-based self-administered screening (Scherer et al., 2014), which allows for anonymity, with anthropomorphic interfaces which may foster some of the beneficial social effects of face-to-face interactions (Weisband & Kiesler, 1996). When the SimSensei system is administered in a private kiosk-based setting, it is envisioned to conduct a clinical interview with a patient who may be initially hesitant or resistant to interacting with a live mental health care provider. SimSensei's real time sensing of user behavior aims to identify behaviors associated with anxiety, depression or PTSD. Such behavioral signals are sensed and inferences are made to quantify user state across an interview; that information is also used in real time to update the style and content of the SimSensei follow-up questions. Technical details of the Multisense software as well as the SimSensei dialog management, natural language system, and agent face/body gesture generation methods are beyond the scope of this paper and can be found elsewhere (Devault et al., 2014; Scherer et al., 2014). Instead, we focus on the usefulness of SimSensei in collecting honest health information and behavioral markers of distress during a clinical interview with active duty Service Members (SMs) prior to and immediately following a 9-month deployment to Afghanistan.

2. METHODS AND PROCEDURE

2.1 Participants

Twenty nine (2 female) active duty members of the Colorado National Guard volunteered for this study prior to embarking on a 9-month deployment to Afghanistan. They were a diverse sample regarding age (Mean=41.46, Range=26 to 56) and previous deployments (Number of combat deployments Mean=2.00, Range=1 to 7).

2.2 Assessment Instruments

The study compared the endorsement of Posttraumatic Stress (PTS) symptoms in three formats: 1) standard administration of the Post-Deployment Health Assessment (PDHA) upon return from deployment; 2) an anonymized version of the PDHA; 3) parallel SimSensei interview questions. The PDHA administered upon return from a military deployment is a self-report rating scale designed to assess a service member's current health, mental health or psychosocial issues commonly associated with deployments, special medications taken during the deployment, and possible deployment-related occupational/environmental exposures. Participants signed releases to access their official web-based PDHA that they submitted to the National Guard upon return from this deployment. On the PDHA, participants are asked self-report PTSD-relevant symptom questions "Have you ever had any experience that was so frightening, horrible, or upsetting that, in the past month, you: A) have had nightmares about it or thought about it when you did not want to?, B) tried hard not to think about it or went out of your way to avoid situations that remind you of it?, and C) were constantly on guard, watchful, or easily startled?" These questions assess whether the SM is experiencing the core DSM-4TR diagnostic symptoms for PTSD (intrusive recollections/re-experienceing; avoidance/numbing; hyperarousal). Participants selected "yes" or "no" on each item of the official PDHA and our anonymized version.

The questions SimSensei asked on these topics were worded slightly differently to embed them in the interview without having the VH simply recite the PDHA. Participants were asked: "Can you tell me about an experience you had in the past few months that challenged you on an emotional level?" (trauma event criterion), followed by "Can you tell me about any bad dreams you've had about your experiences, or times when thoughts or memories just keep going through your head when you wish they wouldn't?" (intrusive recollection criterion), "Tell me about any times you found yourself actively trying to avoid thoughts or situations that remind you of past events," (avoidance/numbing criterion) and "Can you tell me about any times recently when you felt jumpy or easily startled?" (hyperarousal). As described in Scherer et al., (2014), MultiSense quantifies facial affect levels (positive, worry/fear), ranging from 0 to 100, where 0 is the absence of the emotion and 100 a strong emotion. Self-reports of these emotions were also elicited: "I am happy" and "I worry too much" were rated on 4 point scales from never to always.

2.3 Procedure

Participants completed our anonymous PDHA and SimSensei measures both before and after deploying (The official PDHA by definition was only administered at post-deployment). After giving consent, participants completed demographic questions as well as a number of measures described elsewhere (DeVault et al., 2014) and not presented in this paper. The confidentiality of all these measures was stressed. Participants then engaged in an interview with the SimSensei VH who conducted a semi-structured screening interview with a user via spoken language. The interview is structured around a series of agent-initiated questions organized into phases: initially there is a rapport-building phase where the agent asks general introductory questions (e.g., "Where are you from originally?"); this is followed by a clinical phase where the agent asks a series of questions about symptoms (e.g., "How easy is it for you to get a good night's sleep?"), which include the naturally embedded PDHA questions; finally, the agent ends with questions designed to return the patient to a more positive mood (e.g., "What are you most proud of?"). At each phase, the agent can ask follow-up questions (e.g., "Can you tell me more about that?"), provide empathetic feedback (e.g., "I'm sorry to hear that"), and produce nonverbal behaviors (e.g., nods, expressions) for active listening. Participants' answers to the three PDHA questions during the interview were coded by two blind coders as to whether the participant was rating symptoms based on an experience in the last month (per criterion of the PDHA questions). These coders had 100% agreement, and codes served as "yes" or "no" answers.

3. RESULTS

To test whether responses to the three versions of the PDHA (official PDHA, Anonymized PDHA, and SimSensei) differed, scores were created for each version by counting the number of "yes" answers to the three questions, which could range from 0 to 3. To compare these scores, we conducted a repeated-measures ANOVA using the 24 participants who successfully completed all three versions. There was a significant effect of assessment type, $F(2, 23) = 4.29$, $p = .02$ (see Figure 2). Follow-up contrasts revealed that participants reported

©2016 ICDVRAT; ISBN 978-0-7049-1547-3

more symptoms of PTSD (responded “yes” on more questions) when asked by SimSensei (M = 0.79, SE = 0.23) than when reporting on the official PDHA (M = 0.25, SE = 0.15)), F(1, 23) = 7.38, p = .01, or even when reporting on our anonymized version of the PDHA (M = 0.33, SE = 0.16), F(1, 23) = 4.84, p = .04. Moreover, unlike a previous study where anonymity increased reporting of symptoms (Warner et al., 2011), our analysis of this sample did not reveal differences between official and anonymized versions of the PDHA, F (1, 23) = 0.19, p = .66, yet did reveal significantly more endorsement of symptoms with SimSensei compared to these forms.

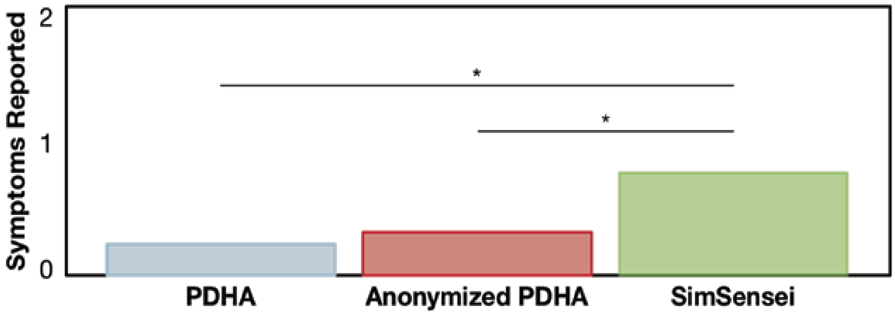

Figure 2. Results of comparison between assessment types.

Additionally, MultiSense facial expression analysis identified pre-to-post reductions in positive affect (M = 0.32, SE = 0.05 to M = 0.15, SE = 0.02, F(1, 22) = 16.33, p = .001) and increases in worry/fear (M = 0.01, SE = 0.002 to M = 0.04, SE = 0.01, F(1, 22) = 8.41, p = .008). This was in contrast to the self-report rating data that showed no difference in self-rated happiness (M = 3.32, SE = 0.13 to M = 3.28, SE = 0.15) and on the worry/fear rating (M = 1.56, SE = 0.12 to M = 1.60, SE = 0.14) pre- to post-deployment (Fs < 0.11 , ps > .74).

4. CONCLUSIONS AND FUTURE WORK

The present study suggests that SMs following a deployment to Afghanistan were more likely to report symptoms of PTSD when interviewed by a VH than on both the official and anonymized versions of the PDHA. This result is in line with our previous work that indicated that users felt less concerned about being evaluated and displayed more sadness in an interview with a VH agent compared to one where they believed a VH avatar was being operated by a human-in-the-loop “Wizard of Oz” controller (Lucas et al., 2014). These results are part of a growing body of research that is suggesting that VH interviews may reduce hesitancy to disclose information (and hence, reduce the fear or experience of stigma) by providing a safe context where users may reveal more honest assessment information in contrast to situations where users are concerned about negative judgments on the part of a human assessors. Additionally, the automatic behavior detection involving facial expression provided a window into the emotional state that differs from self-report in the present study.

We are currently running a replication of this study with a larger sample of U.S. Veterans that aims to investigate whether these results can 1) be replicated, 2) be found within another military population, and 3) are not the product of confounds due to slight wording differences in the assessment questions between the questionnaire and the VH. Specifically, in this study with veterans, the wording is exactly the same across conditions. The data collection is ongoing, but initial results suggest replication and final results from that study will be presented at the conference.

The SimSensei interviewer is also being tested for its effectiveness as a PTSD assessment method within a clinical trial evaluating the use of virtual reality exposure therapy for PTSD due to Military Sexual Trauma. SimSensei interviews are being conducted at Pre-treatment, Mid-treatment, and at Post-Treatment. Data acquired from the capture and analysis of both verbal and non-verbal behavior emitted by the patients during the VH interview process is being compared/correlated with: 1) traditional self-report assessments (Clinician Administered PTSD Scale (CAPS) structured interview and screening measures of PTSD (PCL-M5) and other clinical measures (PHQ-9—Depression, etc.)), and 2) a learning theory-based psychophysiological “startle response” conditioning/extinction protocol. This will allow for a better understanding of the value of VH interview assessment in a situation where stigma due to the nature of sexual trauma may be high and thus, interview questions delivered by a VH may provide a safer context for honest reporting that better reflects the outcomes of treatment.

Acknowledgements. The effort described here is supported by DARPA under contract W911NF-04-D-0005 and the U.S. Army. Any opinion, content or information presented does not necessarily reflect the position or the policy of the United States Government, and no official endorsement should be inferred.

5. REFERENCES

Baker, R.P. (1992). New Technology in Survey Research: Computer-Assisted Personal Interviewing (CAPI). Social Science Computer Review, 10, 145-157.

Baltrusaitis, T, Robinson, P, and Morency, L.-P, (2012), 3D constrained local model for rigid and non-rigid facial tracking, Proceedings of The IEEE Computer Vision and Pattern Recognition, Providence, RI.

Beckenbach, A. (1995). Computer assisted questioning: The new survey methods in the perception of the respondent. Bulletin de Méthodologie Sociologique, 48, 82-100.

Bickmore, T.W., Utami, D., Matsuyama, R., & Paasche-Orlow, M.K. (2016). Improving Access to Online Health Information With Conversational Agents: A Randomized Controlled Experiment. Journal of Medical Internet Research. 18(1):e1) doi:10.2196/jmir.5239

Devault, D, Rizzo, AA, and Morency, L-P, (2014), SimSensei: A Virtual Human Interviewer for Healthcare Decision Support, In the Proceedings of the Thirteenth International Conference on Autonomous Agents and Multiagent Systems (AAMAS).

Ekman, P, and Rosenberg, EL, (1997), What the face reveals: Basic and applied studies of spontaneous expressions using the Facial Action Coding System (FACS), Oxford University Press, New York.

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: 54-61.

Gratch, J. and Marsella, S. (2004). A domain independent framework for modeling emotion. Journal of Cognitive Systems Research. 5(4), 269-306.

Gratch, J., Morency, LP., Scherer, S., Stratou, G., Boberg, J., Koenig, S., Adamson, T. & Rizzo, A.A. (2013). User-State Sensing for Virtual Health Agents and TeleHealth Applications. Studies in Health Technology and Informatics. 191: 151-157.

Joinson, A.N. (2001). Self-disclosure in computer-mediated communication: The role of self-awareness and visual anonymity. European Journal of Social Psychology, 31, 177-192.

Kang, S. & Gratch, J. (2014). Exploring Users' Social Responses to Computer Counseling Interviewers' Behavior. Journal of Computers in Human Behavior, 34C, 120-130

Kang, S. & Watt, J. (2013). The Impact of Avatar Realism and Anonymity on Effective Communication via Mobile Devices. Journal of Computers in Human Behavior 29(3), 1169-1181.

Kang, S. & Gratch, J. (2012). Socially Anxious People Reveal More Personal Information with Virtual Counselors That Talk about Themselves Using Intimate Human Back Stories. The Annual Review of CyberTherapy and Telemedicine, 181: 202-207.

Lucas, G.M., Gratch, J., King, A., and Morency, L.-P. (2014). It's only a computer: Virtual humans increase willingness to disclose. Computers in Human Behavior, 37, 94–100.

Morency, L.-P, Whitehill, J, and Movellan, J, (2008), Generalized adaptive view-based appearance model: Integrated framework for monocular head pose estimation. Proceedings of The 8th IEEE International Conference on Automatic Face Gesture Recognition (FG08), pp. 1–8.

Morency, L.-P., de Kok, I. Gratch, J. (2008). Context-based Recognition during Human Interactions: Automatic Feature Selection and Encoding Dictionary. 10th International Conference on Multimodal Interfaces, Chania, Greece, IEEE.

Morency, L.-P. (2010). Modeling Human Communication Dynamics. IEEE Signal Processing Magazine. 27(6): 112-116.

Park, S., Scherer, S., Gratch, J., Carnevale, P., & Morency, L-P. (2013). In: Affective Computing and Intelligent Interaction (ACII). IEEE. 423-428. DOI: 10.1109/ACII.2013.76

Pentland, A, Lazer, D, Brewer, D, and Heibeck, T, (2009), Using reality mining to improve public health and medicine, Studies in Health Technology and Informatics, 149, pp. 93-102.

Pentland, A, (2008), Honest signals: How they shape our world, MIT Press, Cambridge, MA.

Rizzo, A.A., Kenny, P. & Parsons, T., (2011). Intelligent Virtual Humans for Clinical Training. International Journal of Virtual Reality and Broadcasting, 8(3), Available: http://www.jvrb.org/8.2011/

Rizzo, A.A., Shilling, R., Forbell, E., D., Scherer, S., Gratch, J., & Morency, L-P. (2015). Autonomous Virtual Human Agents for Healthcare Information Support and Clinical Interviewing. In: Luxton, D.D. (Ed). Artifical Intelligence in Mental Healthcare Practice. Academic Press: Oxford. 53-80.

Rizzo, A.A. & Talbot, T. (2016a). Virtual Reality Standardized Patients for Clinical Training. In: Combs, C.D., Sokolowski, J.A. & Banks, C.M. (Eds). The Digital Patient: Advancing Medical Research, Education, and Practice. Wiley: New York. 257-272.

©2016 ICDVRAT; ISBN 978-0-7049-1547-3

Rizzo, A.A., Scherer, S., DeVault, D., Gratch, J., Artstein, R., Hartholt, A., Lucas, G., Marsella, S., Morbini, F., Nazarian, A., Stratou, G., Traum, D., Wood, R., Boberg, J. & Morency, L-P. (2016b-in press). Detection and Computational Analysis of Psychological Signals Using a Virtual Human Interviewing Agent. International Journal on Disability and Human Development, 15(3).

Scherer, S., Stratou, G., Lucas, G., Mahmoud, M., Boberg, J., Gratch, J., Rizzo, A.A. & Morency, LP. (2014). Automatic Audiovisual Behavior Descriptors for Psychological Disorder Analysis. Image and Vision Computing. 32. 648–658.

Sebestik, J., Zelon, H., DeWitt, D., O'Reilly J.M., & McGowan, K. (1988). Initial Experiences with CAPI. Proceedings of the Fourth Annual Research Conference, Arlington, Virginia, 357- 371.

Swartout, B., Artstein, R., Forbell, E., Foutz, S., Lane, H.C., Lange, B., Morie, J., Noren, D., Rizzo, A.A. & Traum, D. (2013). Virtual Humans for Learning. Artificial Intelligence Magazine, 34(4).13-30.

Talbot, T.B., Sagae, K., John, B., Rizzo, A.A. (2012). Sorting out the Virtual Patient: How to exploit artificial intelligence, game technology and sound educational practices to create engaging role-playing simulations. International Journal of Gaming and Computer-Mediated Simulations, 4(3). 1-19.

Thiebaux, M., Marshall, A., Marsella, S., Fast, E., Hill, A., Kallmann, M.,…Lee, J. (2008). SmartBody: Behavior Realization for Embodied Conversational Agents. Proceedings of the International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS). Portugal.

Traum, D., Marsella, S., Gratch, J., Lee, J., Hartholt, A. (2008). Multi-party, Multi-issue, Multi-strategy Negotiation for Multi-modal Virtual Agents. 8th International Conference on Intelligent Virtual Agents. Tokyo, Japan, Springer.

van der Heijden, P. G. M., Van Gils, G., Bouts, J. & Hox, J. (2000). A comparison of randomized response, computer-assisted self-interview and face-to-face direct-questioning. Sociological Methods Research, 28, 505–53

Warner, C.H., Appenzeller, G.N., Grieger, T., Belenkiy, S., Breitbach, J., Parker, J., Warner, C.M. & Hoge, C. (2011). Importance of Anonymity to Encourage Honest Reporting in Mental Health Screening After Combat Deployment. Arch Gen Psychiatry. 68(10), 1065-1071.

Weisband, S. & Kiesler, S. (1996). Self-disclosure on computer forms: Meta-analysis and implications. CHI, 96, 3-10.

Whitehill, J, Littlewort, G, Fasel, I, Bartlett, M, and Movellan, J, (2009), Toward practical smile detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, 31, pp. 2106–2111.