A New Generation of Intelligent Virtual Patients for Clinical Training

Albert “Skip” Rizzo¹
Thomas Parsons¹
J. Galen Buckwalter¹
Belinda Lange¹
Patrick Kenny¹

¹Institute for Creative Technologies, University of Southern California

Correspondence:
Albert “Skip” Rizzo
USC Institute for Creative Technologies
12015 Waterfront Drive
Playa Vista, California, USA
90094
arizzo@ict.usc.edu
Abstract

Over the last 15 years, a virtual revolution has taken place in the use of Virtual Reality simulation technology for clinical purposes. Recent shifts in the social and scientific landscape have now set the stage for the next major movement in Clinical Virtual Reality with the “birth” of intelligent virtual humans. Seminal research and development has appeared in the creation of highly interactive, artificially intelligent and natural language capable virtual human agents that can engage real human users in a credible fashion. No longer at the level of a prop to add context or minimal faux interaction in a virtual world, virtual human representations can be designed to perceive and act in a 3D virtual world, engage in face-to-face spoken dialogues with real users (and other virtual humans) and in some cases, they are capable of exhibiting human-like emotional reactions. This paper will present a brief rationale and overview of their use in clinical training and then detail our work developing and evaluating artificially intelligent virtual humans for use as virtual standardized patients in clinical training with novice clinicians. We also discuss a new project that uses a virtual human as an online guide for promoting access to psychological healthcare information and for assisting military personnel and family members in breaking down barriers to initiating care. While we believe that the use of virtual humans to serve the role of virtual therapists is still fraught with both technical and ethical concerns, we have had success in the initial creation of virtual humans that can credibly mimic the content and interaction of a patient with a clinical disorder for training purposes. As technical advances continue, this capability is expected to have a significant impact on how clinical training is conducted in psychology and medicine.
1. Introduction

Over the last 15 years, a virtual revolution has taken place in the use of simulation technology for clinical purposes. Technological advances in the areas of computation speed and power, graphics and image rendering, display systems, tracking, interface technology, haptic devices, authoring software and artificial intelligence have supported the creation of low-cost and usable PC-based Virtual Reality (VR) systems. At the same time, a determined and expanding cadre of researchers and clinicians have not only recognized the potential impact of VR technology, but have now generated a significant research literature that documents the many clinical targets where VR can add value over traditional assessment and intervention approaches (Holden, 2005; Parsons & Rizzo, 2008; Powers & Emmelkamp 2008; Rose et al, 2005; Riva 2005). This convergence of the exponential advances in underlying VR enabling technologies with a growing body of clinical research and experience has fueled the evolution of the discipline of Clinical Virtual Reality. And this state of affairs now stands to transform the vision of future clinical practice and research in the disciplines of psychology, medicine, neuroscience, physical and occupational therapy, and in the many allied health fields that address the therapeutic needs of those with clinical disorders.

A short list of areas where Clinical VR has been usefully applied includes fear reduction with phobic clients (Parsons & Rizzo, 2008; Powers & Emmelkamp, 2008), treatment for Post Traumatic Stress Disorder (Rizzo et al., 2008), stress management in cancer patients (Schneider et al., 2004), acute pain reduction during wound care and physical therapy with burn patients (Hoffman et al., 2004), body image disturbances in patients with eating disorders (Riva, 2005), navigation and spatial training in children and adults with motor impairments (Stanton et al., 1998; Rizzo et al., 2004), functional skill training and motor rehabilitation with patients having central nervous system dysfunction (e.g., stroke, TBI, SCI cerebral palsy, multiple sclerosis, etc.) (Holden, 2005; Weiss et al., in press) and in the assessment (and in some cases, rehabilitation) of attention, memory, spatial skills and executive cognitive functions in both clinical and unimpaired populations (Rose et al., 2005; Rizzo & Kim, 2005; Rizzo et al., 2004). To do this, VR scientists have constructed virtual airplanes, skyscrapers, spiders, battlefields, social settings, beaches, fantasy worlds and the mundane (but highly relevant) functional environments of the schoolroom, office, home, street and supermarket. These efforts are no small feat in light of the technological challenges, scientific climate shifts and funding hurdles that many researchers have faced during the early development of this emerging technology.

Concurrent with the emerging acknowledgement of the unique value of Clinical VR by scientists and clinicians, has come a growing awareness of its potential relevance and impact by the general public. While much of this recognition may be due to the high visibility of digital 3D games, the Nintendo Wii, and massive shared internet-based virtual worlds (World of WarCraft, Halo and 2nd Life), the public consciousness is also routinely exposed to popular media reports on clinical and research VR applications. Whether this should be viewed as “hype” or “help” to a field that has had a storied history of alternating periods of public enchantment and disregard, still remains to be seen. Regardless, growing public awareness coupled with the solid scientific results have brought the field of Clinical VR past the point where skeptics can be taken seriously when they characterize VR as a “fad technology”.

These shifts in the social and scientific landscape have now set the stage for the next major movement in Clinical VR. With advances in the enabling technologies allowing for the
design of ever more believable context-relevant “structural” VR environments (e.g. homes, classrooms, offices, markets, etc.), the next important challenge will involve populating these environments with virtual human (VH) representations that are capable of fostering believable interaction with real VR users. This is not to say that representations of human forms have not usefully appeared in Clinical VR scenarios. In fact, since the mid-1990’s, VR applications have routinely employed VHs to serve as stimulus elements to enhance the realism of a virtual world simply by their static presence.

For example, VR exposure therapy applications have targeted simple phobias such as fear of public speaking and social phobia using virtual social settings inhabited by “still-life” graphics-based characters or 2D photographic sprites (i.e., static full body photo images of a person with background content removed to allow for relatively seamless insertion into a virtual world) (Anderson et al., 2005; Pertaub et al., 2002; Klinger, 2005). By simply adjusting the number and location of these VH representations, the intensity of these anxiety-provoking VR contexts could be systematically manipulated with the aim to gradually habituate phobic patients and improve their functioning in the real world. Other clinical applications have also used animated graphic VHs as stimulus entities to support and train social and safety skills in persons with high functioning autism (Ruten et al., 2003; Padget et al., 2006) and as distracter stimuli for attention assessments conducted in a virtual classroom (Parsons et al., 2007; Rizzo et al., 2006). Additionally, VHs have been used effectively for the conduct of social psychology experiments, essentially replicating and extending findings from studies on social influence, conformity, racial bias and social proxemics conducted with real humans (Blascovich et al., 2002; Bailenson & Beall, 2006; McCall et al., 2009).

In an effort to further increase the pictoral realism of such VHs, Virtually Better Inc. (www.virtuallybetter.com), began incorporating whole video clips of crowds into graphic VR fear of public speaking scenarios. They later advanced the technique by using blue screen captured video sprites of individual humans inserted into graphics-based VR social settings for social phobia and cue exposure substance abuse treatment and research applications. The sprites were drawn from a large library of blue-screen captured videos of actors behaving or speaking with varying degrees of provocation. These video sprites could then be strategically inserted into the scenario with the aim to modulate the emotional state of the patient by fostering encounters with these 2D video VH representations.

The continued quest for even more realistic simulated human interaction contexts led other researchers to the use of panoramic video capture (Macedonio et al., 2007; Rizzo et al., 2003) of a real world office space inhabited by hostile co-workers and supervisors to produce VR scenarios for anger management research. With this approach, the VR scenarios were created using a 360-degree panoramic camera that was placed in the position of a worker at a desk and then actors walked into the workspace, addressed the camera (as if it was the targeted user at work) and proceeded to verbally threaten and abuse the camera, vis-à-vis, the worker. Within such photorealistic scenarios, VH video stimuli could deliver intense emotional expressions and challenges with the aim of the research being to determine if this method would produce emotional reactions in test participants and if it could engage anger management patients to role-play a more appropriate set of coping responses.

However, working with such fixed video content to foster this form of faux interaction or exposure has significant limitations. For example, it requires the capture of a large catalog of possible verbal and behavioral clips that can be tactically presented to the user to meet the requirements of a given therapeutic approach. As well, this fixed content cannot be readily
updated in a dynamic fashion to meet the challenge of creating credible real time interactions with a virtual human, with the exception of only very constrained social interactions. This process can only work for clinical applications where the only requirement is for the VH character to deliver an open-ended statement or question that the user can react to, but is lacking in any truly fluid and believable interchange following a response by the user. Consequently, the absence of dynamic interaction with these virtual representations without a live person behind the “screen” actuating new clips in response the user’s behavior is a significant limiting factor for this approach. This has led some researchers to consider the use of artificially intelligent VH agents as entities for simulating human-to-human interaction in virtual worlds.

Clinical interest in artificially intelligent agents designed for interaction with humans can trace its roots to the work of MIT AI researcher, Joe Weizenbaum. In 1966, he wrote a language analysis program called ELIZA that was designed to imitate a Rogerian therapist. The system allowed a computer user to interact with a virtual therapist by typing simple sentence responses to the computerized therapist’s questions. Weizenbaum reasoned that simulating a non-directional psychotherapist was one of the easiest ways of simulating human verbal interactions and it was a compelling simulation that worked well on teletype computers (and is even instantiated on the internet today: http://www.ai.ijs.si/eliza-cgi-bin/eliza_script). In spite of the fact that the illusion of Eliza’s intelligence soon disappears due to its inability to handle complexity or nuance, Weizenbaum (1976) was reportedly shocked upon learning how seriously people took the ELIZA program (Howell & Muller, 2000). And this led him to conclude that it would be immoral to substitute a computer for human functions that “...involves interpersonal respect, understanding, and love.”

More recently, seminal research and development has appeared in the creation of highly interactive, artificially intelligent (AI) and natural language capable virtual human agents. No longer at the level of a prop to add context or minimal faux interaction in a virtual world, these VH agents are designed to perceive and act in a 3D virtual world, engage in face-to-face spoken dialogues with real users (and other VHs) and in some cases, they are capable of exhibiting human-like emotional reactions. Previous classic work on virtual humans in the computer graphics community focused on perception and action in 3D worlds, but largely ignored dialogue and emotions. This has now changed. Artificially intelligent VH agents can now be created that control computer generated bodies and can interact with users through speech and gesture in virtual environments (Gratch et al., 2002). Advanced virtual humans can engage in rich conversations (Traum et al., 2008), recognize nonverbal cues (Morency et al., 2008), reason about social and emotional factors (Gratch & Marsella, 2004) and synthesize human communication and nonverbal expressions (Thiebaux et al., 2008). Such fully embodied conversational characters have been around since the early 90’s (Bickmore & Cassell, 2005) and there has been much work on full systems to be used for training (Evans et al., 1989; Kenny et al., 2007; Prendinger & Ishizuka, 2004; Rickel et al., 2001), intelligent kiosks (McCauley & D’Mello, 2006), and virtual receptionists (Babu et al., 2006). Both in appearance and behavior, VHs have now passed through “infancy” and are ready for service in a variety of clinical and research applications.

2. Rationale for Intelligent Virtual Human Patients for Clinical Training

An integral part of medical and psychological clinical education involves training in interviewing skills, symptom/ability assessment, diagnosis and interpersonal communication. In
the medical field, students initially learn these skills through a mixture of classroom lectures, observation, and role-playing practice with *standardized patients*—persons recruited and trained to take on the characteristics of a real patient, thereby affording students a realistic opportunity to practice and be evaluated in a simulated clinical environment. Although a valuable tool, there are several limitations with the use of standardized patients that can be mitigated through VR simulation technology. First, standardized patients are expensive. For example, although there are 130 medical schools in the U.S., only five sites provide standardized patient assessments as part of the U.S. Medical Licensing Examination (USMLE) at a cost of several thousand dollars per student ([www.ecfmg.org/usmle/step2cs/centers.html](http://www.ecfmg.org/usmle/step2cs/centers.html)). Second, there is the issue of standardization. Despite the expense of standardized patient programs, the standardized patients themselves are typically low skilled actors making about $10/hr. and administrators face constant turnover resulting in considerable challenges to the consistency of patient portrayals. This limits the value of this approach for producing reliable and valid interactions needed for the psychometric evaluation of clinicians in training. Finally, the diversity of the conditions that standardized patients can characterize is limited by availability of human actors and their skills. This is even a greater problem when the actor needs to be a child, adolescent, elder, or in the mimicking of nuanced or complex symptom presentations.

The situation is even more challenging in the training of clinical psychology students. Rarely are live standardized patients used in such clinical training. Most direct patient interaction skills are acquired via role-playing with supervising clinicians and fellow graduate students, with closely supervised “on-the-job” training providing the brunt of experiential training. While one-way mirrors provide a window for the direct observation of trainees, audio and video recordings are a more common method of providing supervisors with information on the clinical skills of trainees. As well, the imposition of recording has been reported to have demonstrable effects on the therapeutic process that may confound the end goal of clinical training (Bogolub, 1986).

In this regard, Virtual Patients can fulfill the role of standardized patients by simulating diverse varieties of clinical presentations with a high degree of consistency, and sufficient realism (Stevens et al., 2005), as well as being always available for *anytime-anywhere* training. Similar to the compelling case made over the years for Clinical VR generally (Rizzo & Kim, 2005), VP applications can likewise enable the precise stimulus presentation and control (dynamic behavior, conversational dialog and interaction) needed for rigorous laboratory research, yet embedded within the context of an ecologically relevant simulated environment. Toward this end, there is a growing literature on the use of VPs in the testing and training of bioethics, basic patient communication, interactive conversations, history taking, clinical assessment, and clinical decision-making (Bickmore et al., 2006; Bickmore et al., 2007; Lok et al., 2007; Kenny et al., 2007; Parsons et al., 2008). There has also been a significant amount of work creating physical manikins and virtual patients that represent physical problems (Kotranza et al, 2008), such as how to attend to a wound, or interviewing a patient with a stomach problem to assess the condition (Lok et al., in press) and initial results suggest that VPs can provide valid and reliable representations of live patients (Triola et al., 2006; Andrew et al., 2007). However, research into the use of VPs in psychology and related psychosocial clinical training has been limited (Stevens et al., 2005; Bickmore at al., 2007; Lok et al., 2007; Kenny et al., 2007; Parsons et al., 2008; Triola et al., 2006; Andrew et al., 2007; Bernard et al., 2006; Dickerson et al., 2005; Johnsen et al., 2006; Beutler & Harwood, 2004). For example, Beutler and Harwood (2005) describe the development of a VR system for training in psychotherapy (characters primarily with psychological disorders or medical conditions in which a
psychological condition may complicate a straightforward medical diagnosis) and summarize training-relevant research findings. We could find no other references on the use of VPs in clinical psychology training to date, despite online searches through MEDLINE, Ovid, and the psychotherapy literature. From this, there appears to be a gap in research into the design of intelligent VPs that have realistic interaction and communication capabilities for training clinical interviewing, diagnostic assessment and therapy skills in novice clinicians.

The remainder of this paper will detail our work developing and evaluating the use of artificially intelligent VHs designed to serve the role of VPs for training clinical interaction skills in novice clinicians. While we believe that the use of VHs to serve the role of virtual therapists is still fraught with both technical and ethical concerns (Rizzo et al., 2002), we have had success in the initial creation of VHs that can mimic the content and interaction of a patient with a clinical disorder for training purposes and we will briefly describe our initial areas of focus. We will also discuss our related emerging work developing an online VH presence (SimCoach) for providing assistance to military personnel and significant others in the access of relevant psychological health and TBI care information. This project aims to break down barriers to care (e.g. unawareness, stigma, complexity of the military psychological healthcare system, etc.) and assist users in the process of initiating a first contact with a live human healthcare provider.

3. Virtual Patient Projects

The art and science of evaluating interviewing skills using VPs is still a young discipline with many challenges. One formative approach is to compare performances obtained during interviews with both live standardized patients and with VPs, and then to conduct correlational analyses of metrics of interest. This information can then be evaluated relative to an Objective Structured Clinical Examination (OSCE) (Hardin et al., 1975; Walters et al., 2005). Such tests typically take from 20-30 minutes and require a faculty member to watch the student perform a clinical interview while being videotaped. The evaluation consists of a self-assessment rating along with faculty assessment and a review of the videotape. This practice is common, although is applied variably, based on the actors, available faculty members and space and time constraints at the training site. A general complication involved in teaching general interviewing skills is that there are multiple theoretical orientations and techniques to choose from and the challenge will be to determine what commonality exists across these methods for the creation of usable and believable VPs that are adaptable to all clinical orientations. To minimize this problem in our initial efforts, we have concentrated on assessing the skills required to diagnose very specific mental disorders (i.e., conduct disorder, PTSD, depression, etc.). We also use the setting of an initial intake interview to constrain the test setting for acquiring comprehensible data to drive future research. In our test protocols clinicians are typically provided some knowledge as to why the patient is there (i.e., a referral question), but need to ask the patient strategic questions to obtain a detailed history useful for specifying a clinical condition in support of coming to a differential diagnosis and for formulating a treatment plan. In this manner, the system is designed to allow novice clinicians the opportunity to practice asking interview questions that eventually lead to the narrowing down of the alternative diagnostic options, leading to the arrival of a working diagnosis based on the VP meeting the criteria for a specific DSM diagnosis (or not!).
3.1 Adolescent Male with a Conduct Disorder

Our initial project in this area involved the creation of a virtual patient, named “Justin”. Justin portrays a 16-year old male with a conduct disorder who is being forced to participate in therapy by his family (see Figure 1). The system was designed to allow novice clinicians to practice asking interview questions, to attempt to create a positive therapeutic alliance and to gather clinical information from this very challenging VP. Justin was designed as a first step in our research. At the time, the project was unfunded and thus required our lab to take the economically inspired route of recycling a virtual character from a military negotiation-training scenario to play the part of Justin. The research group agreed that this sort of patient was one that could be convincingly created within the limits of the technology (and funding) available to us at the time. For example, such resistant patients typically respond slowly to therapist questions and often use a limited and highly stereotyped vocabulary. This allowed us to create a believable VP within limited resources for dialog development. As well, novice clinicians have been typically observed to have a difficult time learning the value of “waiting out” periods of silence and non-participation with these patients. We initially collected user interaction and dialog data from a small sample of psychiatric residents and psychology graduate students as part of our iterative design process to evolve this application area. The project produced a successful proof of concept demonstrator, which then led to the acquisition of funding that currently supports our research in this area.

3.2 Adolescent Female with PTSD due to Sexual Assault

Following our successful Justin proof of concept, our 2nd VP project involved the creation of a female sexual assault victim, “Justina” (see Figure 2). The aim of this work was two fold: 1. Explore the potential for creating a system for use as a clinical interview trainer for promoting sensitive and effective clinical interviewing skills with a VP that had experienced significant personal trauma; and 2. Create a system whereby the dialog content could be manipulated to create multiple versions of Justina to provide a test of whether novice clinicians would ask the appropriate questions to assess whether Justina met the criteria for the DSM-4r diagnosis of PTSD based on symptoms reported during the clinical interview.

For the PTSD content domain, 459 questions were created that mapped roughly 4 to 1 to a set of 116 responses. The aim was to build an initial language domain corpus generated from subject matter experts and then capture novel questions from a pilot group of users (psychiatry residents) during interviews with Justina. The novel questions that were generated could then be fed into the system in order to iteratively build the language corpus. We also focused on how well subjects asked questions that covered the six major symptom clusters that can characterize PTSD following a traumatic event. While this approach did not give the Justina character a lot of depth, it did provide more breadth for PTSD-related responses, which for initial testing seemed prudent for generating a wide variety of questions for the next Justina iteration. (This 2nd iteration is currently in progress.)

In the initial test, a total of 15 Psychiatry residents (6 females, 9 males; mean age = 29.80, SD 3.67) participated in the study and were asked to perform a 15-minute interaction with the VP to take an initial history and determine a preliminary diagnosis based on this brief interaction with the character. The participants were asked to talk normally, as they would to a standardized patient, but were informed that the system was a research prototype that uses an experimental speech recognition system that would sometimes not understand them. They were
instructed that they were free to ask any kind of question and the system would try to respond appropriately, but if it didn’t, they could ask the same question in a different way.

From post questionnaire ratings on a 7-point Likert scale, the average subject rating for believability of the system was 4.5. Subjects reported their ability to understand the patient at an average of 5.1, but rated the system at 5.3 as frustrating to talk to, due to speech recognition problems, out of domain answers or inappropriate responses. However most of the participants left favorable comments that they thought this technology will be useful in the future, and that they enjoyed the experience of trying different ways to talk to the character in order to elicit an relevant response to a complex question. When the patient responded back appropriately to a question, test subjects informally reported that the experience was very satisfying.

Analysis of concordance between user questions and VP response pairs indicated moderate effects sizes for Trauma inquiries (r = 0.45), Re-experiencing symptoms (r = 0.55), Avoidance (r = 0.35), and in the non-PTSD general communication category (r = 0.56), but only small effects were found for Arousal/Hypervigilance (r = 0.13) and Life impact (r = 0.13). These relationships between questions asked by a novice clinician and concordant replies from the VP suggest that a fluid interaction was sometimes present in terms of rapport, discussion of the traumatic event, the experience of intrusive recollections and discussion related to the issue of avoidance. Low concordance rates on the arousal and life impact criteria indicated that a larger domain of possible questions and answers for these areas was not adequately modeled in this pilot effort and this is now being addressed in the 2nd iteration of Justina.

3.3 Military VPs for Training in Depression/Suicide Issues and for PTSD Exposure Therapy Training

The 2nd iteration of Justina is underway, now informed by both the quantitative and qualitative results from this user test, and this work has now formed the basis for a new project that will further modify the Justin and Justina VP characters for military clinical training. In one project, both Justina (see Figure 3) and new male characters (see Figure 4-5) will appear as military personnel who are depressed and possibly contemplating suicide, for use as a training tool for teaching clinicians and other military personnel how to recognize the potential for this tragic event to occur. A related component of this project focuses on the military version of Justina with the aim to develop a training tool that clinicians can practice sensitive interviewing skills for addressing the growing problem of sexual assault within military ranks. The system is also being designed for use by command staff to foster better skills for recognizing the signs of sexual assault in subordinates under their command and for improving the provision of support and care.

We are also intending to use both military VP versions to serve in the role of a patient who is undergoing both Imaginal and Virtual Reality-delivered exposure therapy (see Figure 6) for PTSD. In the imaginal exposure version, the VPs will be programmed with a variety of trauma narrative content and trainees will have to opportunity to practice the skills that are required for appropriately fostering emotional engagement with the trauma narrative as is needed to promote optimal therapeutic habituation during exposure therapy sessions. In the VR exposure system, the VP will appear in a simulation of a therapy room wearing a VR head mounted display. In this version the clinician will be given training in how to use the Virtual Iraq/Afghanistan interface controls and practice the skills that are required to use this technological enhancement for exposure therapy in a safe and effective fashion. This simulation
of a patient experiencing VR exposure therapy uses the Virtual Iraq/Afghanistan PTSD system (Rizzo et al., 2008) as the VR context, and the training methodology is based on the Therapist’s Manual created for that VR application by Rothbaum, Difede and Rizzo (2008). We believe the “simulation of an activity that occurs within a simulation” is a novel concept that has not been reported previously in the VR literature! This R&D effort is in collaboration with the USC School of Social Work as part of their efforts to improve experiential training content in a new program that confers a master’s degree in Military Social Work. For this project, we are also creating mixed reality natural human scale projection set-ups (see Figure 7).

3.4 Sick Call – A Military Medical Training System

The U.S. Army has recently funded our group to develop a VP project entitled Sick Call. This project focuses on using virtual patients for training differential diagnosis skills, particularly in situations where psychological factors may complicate the assessment of medical symptoms. The project will provide a gallery of VP’s to allow health practitioners the opportunity to practice interviewing skills, clinical assessments, diagnosis, and interpersonal communications within the context of a simulation of the standard daily review of military personnel requiring medical evaluation (i.e., Sick Call) in a military hospital ward. Designed to enhance the skills and experience of military medical professionals (e.g., physicians, medics, physicians assistants), Sick Call will present trainees with patients having two medical conditions that are often confounded. The accurate assessment of the patient will require both the physical measurement of symptoms and verbal interaction. The overall patient evaluation will follow the “SOAP” format (subjective, objective, assessment, and plan) and include taking a history and conducting a limited physical examination, to train assessment skills in the formation of a treatment plan.

While the specific VPs and the presentation of their conditions are still being designed at the time of this writing, the aim is to simulate accurate physical behaviors relevant to a presenting health condition. In addition to verbalizations, these will include realistic gestures, facial expressions, and reactions consistent with pain that will provide emotional and physical cues for training health professional to distinguish diagnostic conditions. Such cues will include (but are not limited to): 1) sweating at varying levels, from mild to extreme; 2) skin discoloration that could indicate heat stroke, allergic reaction, trouble breathing, burns, etc.; 3) facial expressions indicating varying degrees of anxiety; 4) facial expressions indicating pain, from mild to extreme; 5) Physiological response to intervention, as indicated by changes in heart rate or breathing or other appropriate responses. 6) Verbalizations and other auditory responses (grunts, screams, crying, etc).

The Sick Call system will have the ability to track and record what the health professional is doing throughout the diagnostic process. Advanced natural language processing capabilities that will allow the health professional to communicate with the VPs by voice and a variety of interactional behaviors will be captured and logged. The effectiveness of the dialog that takes place between the health professional and the virtual patient will be assessed (e.g., is the health professional able to discover additional information from the discussions with the patient that helps with diagnosis; can they calm distressed patients; how well can they answer the patient’s questions and concerns in an empathetic manner). Another design feature of Sick Call is to provide an assessment and immediate feedback to the health professional regarding clinical performance. The feasibility of automatically identifying key points and missed opportunities in the interaction along with an instant replay of these events (i.e., after action review) to the health
professional will be investigated. The *Sick Call* project aims to create VPs that present a complex mix of physical and psychological symptoms to enhance medical decision-making in situations that commonly challenge a simple and straightforward diagnosis. In spite of the *Sick Call* medical setting and emphasis, it will be possible to investigate clinical decision-making issues relevant to psychological practice by modifying the virtual context once this system is completed.

3.5 SimCoach - An Online Virtual Human Guide for Promoting Access to Psychological Care Content and for Support in Initiating Live Care

In response to the challenges that the conflicts in Iraq and Afghanistan have placed on the burgeoning population of service members and their families, the Department of Defense (DoD) has supported the creation of the Defense Centers of Excellence for Psychological Health and Traumatic Brain Injury (DCoE). Its primary mission to assess, validate, oversee and facilitate sharing of critical information relative to the areas of injury prevention, resilience, identification, treatment, outreach, rehabilitation, and reintegration programs for psychological health and traumatic brain injury. In line with this mission, DCoE has now funded the development of an online virtual human presence to serve as a guide for service members, Veterans and their families seeking behavioral health information and advice. This is in response to the fact that in spite of a Herculean effort on the part of the DoD to produce and disseminate behavioral health programs for military personnel and their families, the complexity of the problem continues to challenge the best efforts of military mental health care experts, administrators and providers. Since 2004, numerous blue ribbon panels of experts have attempted to assess the current DoD and VA healthcare delivery system and provide recommendations for improvement (DoD Mental Health Task Force Report, 2007; Institute of Medicine of the Academies of Science, 2007; Dol-Shalala Commission, 2007; Tanielian et al., 2008; American Psychological Association, 2007). Many of these reports cite a need for the identification and implementation of ways to enhance the healthcare dissemination/delivery system. For example, the American Psychological Association Presidential Task Force on Military Deployment Services for Youth, Families and Service Members (American Psychological Association, 2007) presented their preliminary report in February of 2007 that flatly stated that they were, “…not able to find any evidence of a well-coordinated or well-disseminated approach to providing behavioral health care to service members and their families.” The APA report also went on to describe three primary barriers to military mental health treatment for service members and families: *availability, acceptability and accessibility*. The overarching goal reported from this and other reports is to provide better awareness and access to existing care while concurrently reducing the complexity and stigma in seeking psychological help. In essence, new methods are needed to reduce such barriers to care.

To address these barriers to care, DCoE has recently funded our group to develop an intelligent, interactive virtual human program currently referred to as SimCoach. The SimCoach VH experience is designed to attract and engage service members and Veterans (and their families/significant others) to assist them in accessing healthcare information. As well, SimCoach is being designed to support users in making a decision as to whether they will take the first step toward initiating psychological care with a live provider. It is not the goal of SimCoach to breakdown all of the barriers to care or to provide diagnostic or therapeutic services that are best delivered by a real clinical provider. Rather, SimCoach will foster comfort and
confidence by promoting users’ efforts to understand their situations better, to explore available options and initiate treatment when appropriate. Coordinating this experience will be a VH, selected by the user from a variety of archetypic characters (see Figures 8-10), who will answer direct questions and/or guide the user through a sequence of user-specific questions, exercises and assessments. This interaction between the VH and the user will provide the system with the information needed to guide users to the appropriate next step of engagement with the system or to initiate contact with a live provider. However, the SimCoach project is not conceived to deliver treatment or diagnosis or as a replacement for human providers and experts. Instead, SimCoach will aim to start the process of engaging the service member and/or their family by providing support and encouragement, increasing awareness of treatment options, and in assisting individuals, who may otherwise be initially uncomfortable talking to a “live” care provider, in their efforts to initiate care. The following is a use-case of how SimCoach will interact with a potential user.

Maria was the 23-year old wife of Juan, an OIF veteran who had completed two deployments before leaving the service. After his return, she noticed something different. He had become distant, never discussed his experiences in Iraq, and when asked, he would answer, “that was then, this is now, case closed”. He also wasn’t as involved with their two children (the 2nd one was born while he was in Iraq), only playing with their oldest boy after hours of begging. For the most part, Juan stayed home and had yet begun to look for a civilian job. He didn’t sleep much and when he did manage to fall asleep, he would often wake up after an hour, highly agitated claiming that he heard someone trying to get in the bedroom window. When this happened, he would sometimes sit till dawn, peering through slits in the closed blinds, watching for the “imaginary” intruder to return. He seemed jumpy when not drinking and watching TV. He drank heavily during the day and Maria would often find him asleep or passed out on the couch when she got home from her job after picking the kids up from her mother’s house. She told her mother that it felt like she was living with a ghost, but that she still loved him. She just wanted the “old Juan” back. However, each day things got worse and she was feeling like she couldn’t live like this much more. She felt guilty for the increasing resentment that she felt but didn’t know how (or was afraid) to talk to Juan about what she was feeling. Juan also kept a pistol in the house and one time she had moved it off the dresser while cleaning and when Juan couldn’t find it, he went ballistic and ran frantically around the house, screaming, “how am I gonna protect my family without my weapon!”

Maria was at a loss as to what to do when her mother mentioned hearing on Oprah about a way to find help for these kinds of problems on the internet with a thing called “SimCoach”. Maria had only occasionally “played” on the AOL games site before and she didn’t own a computer, but her older sister’s son was a “computer nut” and agreed to let her come over to use his computer and try out SimCoach. She couldn’t understand how a computer could help her, but she was desperate for any help she could get. So her nephew showed her how to type in the address for www.simcoach.mil on his computer and then went out with his friends to a movie.

Maria was intrigued when the screen lit up and created the illusion of standing in front of a “craftsman”-like building with the sign above it reading, “DCoE Helpcenter”. Immediately the “virtual” director of the center walked out onto the porch and beckoned
her to come in. The director stated that “we are here to understand your needs and get you started on the path to help” and showed Maria a poster just inside the door that had images and short biographies of the staff. Pointing towards the poster, she said, “here is our staff, have a look and click on the picture of who you would feel comfortable meeting with.” Maria paused when she noticed a staff member that reminded her of a teacher she had in high school who was always helpful and kind to her. She clicked on this picture and was then guided through the hallway of the center that actually looked quite warm and peaceful with virtual people in the hall smiling and talking to each other softly. The program whisked her into a room where Dr. Hartkis, sitting in a thick fabric chair next to a fireplace, smiled, and softly asked her how he could help. Maria knew that this was just a virtual human, but for some reason she felt comforted by his soft voice and kind facial expressions. She had never been to a clinician before for this kind of help and was surprised by how safe and comfortable she felt. Not knowing what to expect, she described how her husband, Juan, was having problems ever since he came back from the war. She was surprised when the doctor said in a reassuring voice, “If you want to tell me more about it, I think I can help you and your family.” After requesting some basic information, Dr. Hartkis then asked Maria some questions that seemed like he really might “understand” some of what she was going through. Eventually, after answering a series of thoughtful questions, Dr. Hartkis reassuringly smiled and then pointed to a wall in the room and said, “Here are some websites that have information that is available to help folks that are going through what you are feeling. We can pull up one of them and take a look at what is available or I can find a care provider in or near your zip code that we can make an appointment with right now so you can begin to find the help that both you and Juan could benefit from. Or, if you’re not ready for that yet, we can still talk more about what you’re going through now.” Maria couldn’t believe that this computer character seemed so genuine in his face and his manner, and that she felt like she wanted to tell him more. Perhaps he might really be able to get her started on the road to help both her and Juan? Suddenly she realized that she had been online for an hour and needed to go home. As she was leaving, she wondered aloud if she could think about the options that she learned about today and then come back to make a decision on what to do. Dr. Hartkis smiled and said, “Of course we can meet again...you see, I will always be here to guide you to the help you need, whenever you’re ready.”

While the use-case presented above is fictional, it illustrates one of a myriad of forms of confidential interaction that a tireless and always-available virtual human can foster. A fundamental challenge of the SimCoach project will be to better understand the diverse needs of the user base such that appropriate individual user experiences can be designed to promote better healthcare access. At the most basic level, there are immense differences in the needs of service members and their families. Further, there are likely large differences in the level of awareness users will have of existing resources and in their own need/desire to engage such resources. Within the service member population itself there is a high likelihood that individual users will have had very diverse combat experiences, help-seeking histories and consequent impact on significant others. The net result of attempting to engage such a diverse user base is that the system will need to be able to employ a variety of general strategies and tactics to be relevant to each individual user.
In this regard, the SimCoach project is employing a variety of techniques to design the user experience. One relevant clinical model is the PLISSIT clinical framework (Permission, Limited Information, Specific Suggestions, and Intensive Therapy) (Annon, 1976), which provides an established model for encouraging help-seeking behaviors in persons who may feel stigma and insecurity regarding a clinical condition. In the SimCoach project, the aim is to address the “PLISS” components, leaving the intensive therapy component to live professionals to which users in need of this level of care can be referred. Another source of knowledge is social work practice. Such models take a case management approach, serving both as an advocate and a guide. Another source of knowledge is the entertainment/gaming industry. While this knowledge from this community is not typically applied towards healthcare, they do focus explicitly on attracting and engaging individuals. As we work to develop this web-based VH interactive system we are working closely with experts in all three of these models to achieve our goal of engaging and focusing this unique user base on the steps to initiate care as needed. Additionally, all interactions will be consistent with findings that suggest that interventions with individuals with PTSD and other psychosocial difficulties achieve the following: 1) promotion of perceptions of self-efficacy and control 2) encouragement of the acceptance of change; 3) encouragement of positive appraisals; and 4) an increase in the usage of adaptive coping strategies (Whealin et al., 2008). These principles of intervention will be implicit in all of the interactions of SimCoach and its users.

4. Conclusions

The systematic use of artificially intelligent virtual humans in clinical virtual reality applications is still clearly in its infancy. But the days of limited use of VH’s as simple props or static elements to add realism or context to a clinical VR application are clearly in the past. In this article we have presented examples of the creation and use of VH characters to serve the role of digital “standardized patients” for training clinical skills, in both psychological and medical care domains. These initial projects have lead to new opportunities for exploring the use of VHS to serve as online mental healthcare guides or coaches. This work is focused on breaking down barriers to care (stigma, unawareness, complexity, etc.) by providing military service members, Veterans, and their families with confidential help in exploring and accessing psychological healthcare content and for promoting the initiation of care with a live provider if needed. These projects are nascent efforts in this area, yet in spite of the current limits of the technology, it is our view that the clinical targets selected can still be usefully addressed. The capacity to conduct clinical training within simulations that provide access to credible virtual patients where novice clinicians can gain exposure to the presentation of a variety of clinical conditions will soon provide a safe and effective means for learning skills before actual training with real patients and for supplementing continuing education throughout the professional lifespan. And as the underlying enabling technologies continue to advance, significant opportunities will emerge that will reshape the clinical training landscape.

If this exploratory work continues to show promise, we intend to address a longer-term vision—that of creating a comprehensive DSM diagnostic trainer that has a diverse library of VPs modeled after each diagnostic category. The VPs would be created to represent a wide range of age, gender and ethnic backgrounds and could be interchangeably loaded with the language and emotional models defined by the criteria specified in any of the DSM disorders. We believe this vision will also afford many research opportunities for investigating the functional and
ethical issues involved in the process of creating and interacting with virtual humans and patients. While ethical challenges may be more intuitively appreciated in cases where the target user is a patient with a clinical condition seeking a virtual clinician, the training of clinicians with VPs will also require a full appreciation of how this form of training impacts clinical performance with real patients. These are not trivial concerns and will require careful ethical and scientific consideration. But as computing power continues to develop at exponential rates, the creation of highly interactive, intelligent VHs is not only possible, but probable. The birth of this field has already happened; the next step is to insure that it has a healthy upbringing.
References


**Figures**

Figure 1. Justin

Figure 2. Civilian Justina

Figure 3. Military Justina
Figure 4. Military Male Character

Figure 5. New Male Characters

Figure 6. Justina in VR Exposure Therapy
Figure 7. Full Scale Mixed Reality Projection Set-up

Figures 8-10. SimCoach Archetypes – Female Aviator, Battle Buddy, Retired Sergeant Major