Lessons Learned using Responsive Virtual Humans for Assessing Interaction Skills

Robert C. Hubal
RTI International
RTP, NC
rhubal@rti.org

Diana H. Fishbein
RTI International
Baltimore, MD
dfishbein@rti.org

Mallie J. Paschall
PIRE
Berkeley, CA
paschall@prev.org

ABSTRACT

In this paper we describe lessons learned applying synthetic character vignettes to assess the success of an intervention teaching at-risk adolescents social skills. We designed vignette storylines to target those social skills – engaging in negotiation, maintaining emotional control, and seeking information. We also designed language models so that when participants verbally interacted with the synthetic character, the natural language system would interpret their input as reflecting one or more social skill. Similarly, we designed behavior models so that the character would attempt different strategies to require participants to demonstrate skills. Finally, we devised methods to measure participants' engagement compared to what their engagement would be during role-play. Results from our study are highly encouraging, with adequate measures of construct and criterion validity showing adolescents' behavior while engaged with the synthetic character indeed reflect social skills. Results also show the vignettes are sufficiently realistic to highlight differences in behavior between groups of participants, such as adolescents who are or are not exposed to the intervention. We provide rationale for using synthetic characters in a variety of assessment situations relevant to the military.

ABOUT THE AUTHORS

Rob Hubal is a Senior Research Engineer in RTI's Technology Assisted Learning division. He is co-developer of software that enables synthetic characters to act and behave realistically in controlled learning and assessment contexts. He has applied research results to such everyday domains as medical informed consent and consumer decision-making. Dr. Hubal holds a masters degree in Computer Science and a doctorate in Cognitive Psychology.

Diana Fishbein currently directs the Transdisciplinary Behavioral Science Program for RTI. She was Principal Investigator for the project described in this paper. She consults regularly with federal, state, and local agencies for purposes of expert witnessing in criminal court, training, technical assistance, scientific peer reviews, and development of research protocols. Dr. Fishbein holds a joint doctorate in Criminology and Psychobiology.

MJ Paschall is a Research Scientist at the Pacific Institute for Research and Evaluation, Prevention Research Center. His research focuses on heavy drinking in young adulthood and risk factors for adolescent drug use, and he served as co-PI on the project described in this paper. Dr. Paschall holds a masters degree in Community Health Education and a doctorate in Health Behavior & Health Education.

Lessons Learned using Responsive Virtual Humans for Assessing Interaction Skills

Robert C. Hubal
RTI International
RTP, NC
rhubal@rti.org

Diana H. Fishbein
RTI International
Baltimore, MD
dfishbein@rti.org

Mallie J. Paschall
PIRE
Berkeley, CA
paschall@prev.org

INTRODUCTION

Virtual battlefield simulations are increasingly used for acquisition, practice, and assessment of tactical skills, as an alternative to live troop exercises. Similarly for training interaction skills, where synthetic characters are used as an alternative to role-playing for acquisition and practice of skills such as negotiation, de-escalation, and interrogation (Hubal & Guinn, 2003).

Applications using synthetic characters for assessment of interaction skills, though, compared to training, are more rare because the technology is still advancing. Most researchers, trainers, managers, and other decision makers rely on well-understood methods for assessment, such as questioning and questionnaires, interactive video, and role-plays (see Table 1). We support the use of these methods, but have long argued that synthetic characters offer significant advantages in some contexts (Hubal & Frank, 2001; Hubal, *et al.*, 2000; Kizakevich, *et al.*, 1998).

For instance, post traumatic stress disorder (PTSD) in first responders is highly probable subsequent to response to terrorist attack (Schlenger, *et al.*, 2002), in addition to other precursors, such as combat experience (Kulka, *et al.*, 1990; Southwick, *et al.*, 1995). Left untreated, PTSD can lead to potential harm to self and others, hence applications in which military police and other emergency response personnel are assessed for signs of mental health problems due to recently encountered highly stressful engagements can prove beneficial. Failure to effectively recognize the behaviors exhibited by soldiers suffering from PTSD or co-morbid disorders could lead to increased domestic violence and repeat violent offenses (McFall, *et al.*, 1999; Rosenberg, *et al.*, 2001), impaired job performance, and more distally, an impact on force readiness and attrition. A synthetic character application, used intelligently, could lead to greater understanding of the signs and symptoms associated with PTSD and realization of the benefits of taking appropriate action when PTSD is suspected and the risks of not taking action. An additional potential future audience is military officers who are responsible for their soldiers' mental as well as physical well-being, and who are in prime position to notice behaviors in their soldiers that appear out of line with regular behaviors. The intention would be to train these officers to recognize behaviors suggestive of PTSD or other psychiatric disorders, to assess soldiers' behaviors by having them interact with synthetic characters, and to inform officers and soldiers as necessary of health referral services.

In this paper we present usability and acceptance data that support the use of synthetic characters – also called embodied conversational agents (Cassell, *et al.*, 2000), socially intelligent agents (Dautenhahn, 1997), and responsive virtual humans (Hubal & Frank, 2001) – in general. We then describe results from a study using synthetic character vignettes to assess adolescents' social competency skills in a risky decision-making setting. Specifically, we presented adolescents with vignettes simulating real social encounters that may have adverse consequences, designed to entice the participants to engage in several types of risky behaviors – known to be associated with violent behavior and drug use – that are targeted by many adolescent violence and drug abuse prevention programs (Paschall, *et al.*, in press). Social competency skills assessed include emotional control, information seeking, expressing one's own preferences, negotiation and willingness to compromise, and being non-provocative. The results are highly encouraging, with measures of construct and criterion validity showing decisions made by adolescents that do indeed reflect social competency skills. Results also show that the synthetic character vignettes are sufficiently realistic to highlight differences in behavior between groups of participants, such as those adolescents who do and do not meet diagnostic criteria for conduct disorder.

SYNTHETIC CHARACTER TECHNOLOGY

Synthetic character technology encompasses an interactive 3D environment with natural language processing and behavioral modeling of synthetic characters in the scene. In synthetic character applications, students are engaged by characters who accurately use gaze, gesture, intonation, and body posture as well as verbal feedback during the interaction. In our applications we view the interaction as a virtual role-play (Hubal & Guinn, 2003). Figure 1 shows a representative synthetic character application architecture.

Acceptability of Synthetic Characters

We have conducted several studies involving various user groups in very different domains (Frank, *et al.*, 2002; Guinn, *et al.*, 2004; Kizakevich, *et al.*, 2003). A common finding for the majority of our participants was to suggest users enjoyed using the applications – and/or were observed to be engaged with the synthetic characters – despite occasional technical obstacles, prototype-stage content, and conspicuous presence of the investigators. Some other of our and others' findings include:

- It is critical in synthetic character applications to be able to detect and respond appropriately to 'bad' or inappropriate input. Users often (but not always intentionally) speak utterances that are outside the range of what was expected in the context of the dialog. Not catching these bad responses breaks the participant's flow.
- Because of shortcomings in speech recognition technology, other means of input are often needed to overcome the limitations of large grammars. Developers variously use guided speech, typed input, menu selection, and observer-categorized input.
- Very complex compound sentences, multiple sentence, and even paragraph-long utterances play havoc with language understanding. This phenomenon indicates a need to set user expectations in the training environment prior to their using the system.

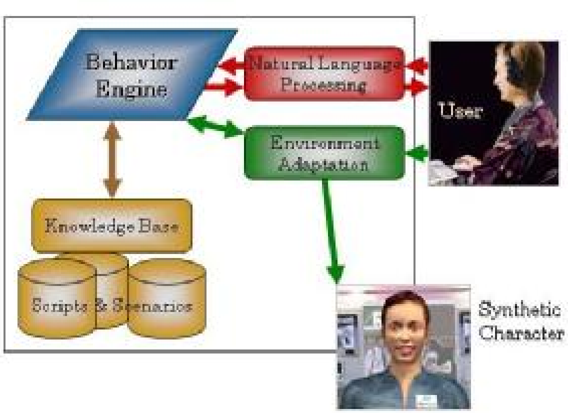

Figure 1. Synthetic Character Architecture

Table 1. Comparison of Assessment Approaches

Role-playing
<u>Advantages</u> Can hire professional actors; student behavior is realistic; role-play skits are relevant; closest to situated learning (Brown, Collins, & Duguid, 1989; Lave & Wenger, 1991). <u>Disadvantages</u> Available time is limited; partner (usually another student) must be available; roles often poorly understood; no variability in age, ethnicity, or gender.
Questioning/Questionnaires
<u>Advantages</u> Easy to implement; can vary scenarios presented; can be used in any setting. <u>Disadvantages</u> Student behavior does not reflect realistic activities; requires student to demonstrate knowledge rather than skills.
Interactive video
<u>Advantages</u> Actions are tracked and can be replayed; enforces standardization of responses; system can guide learning; uses professional actors; can be used in any setting (Fletcher, 1991). <u>Disadvantages</u> Scenarios presented are fixed; scenarios may not be exactly relevant; no variability in age, ethnicity, gender beyond that videotaped; no knowledge of characteristics of interaction partner beyond those given; student behavior does not reflect realistic activities.
Synthetic characters
<u>Advantages</u> Numerous role-play scenarios; actions are tracked and can be replayed; simulation of variable conditions; enforces standardization of responses; different ages, ethnicities, genders, and personalities possible. <u>Disadvantages</u> Start-up costs to usability; scenarios presented are relatively fixed, though flexible; requires decent hardware to allow use in any setting.

- Pre-recorded speech is generally viewed as more acceptable than any currently available speech synthesizer. Still, text-to-speech can lead to some level of engagement.
- The synthetic characters can range from photo-realistic to caricatures. The level of visual realism, in combination with convincing linguistic and emotional models, depends on application content, task demands, and participant characteristics to sufficiently engage participants.

Applications using Synthetic Characters

Applications using synthetic characters for training interaction skills are no longer novel. Researchers have developed such diverse synthetic character applications as: law officers learning to manage encounters with the mentally ill (Frank, *et al.*, 2002); investigative agents gaining experience in listening, asking appropriate questions, identifying signs of deception, developing interrogation themes, and distinguishing between deceptive and truthful responses while interviewing suspects (Olsen, Sellers, & Phillips, 1999); medical interviewing of patients with potential bioterrorist agent exposure (Kizakevich, *et al.*, 2003); primary care clinicians needing to learn protocols to perform interventions for alcohol abuse with their patients (Hayes-Roth, 2003); school resource officers learning to recognize potentially harmful situations for school children, their peers, other individuals, and school property (Sims, 2003); and researchers learning to provide informed consent (Visscher, *et al.*, to appear).

There are fewer applications using synthetic characters for any kind of assessment. One involves informational and emotionally supportive 'face-to-face' interviews between individuals with recent amputations and amputees who have learned to cope with their limb loss (Harless, *et al.*, 2003). Another involves first responders being assessed on how well they follow emergency medical response protocol (Kizakevich, *et al.*, 1998). There are quite a few applications of virtual reality in general for cybertherapy, but few of these applications use synthetic characters (Suler, 2000; Wiederhold & Wiederhold, 2004). One reason for the paucity of applications is the current state of the art: the realism (be it visual, emotional, linguistic, or expressive) is not quite there, or not there in aggregate, so trainers and psychiatrists are understandably hesitant to employ synthetic characters for the critical task of assessing student or patient performance. The particular application of synthetic characters described next, however, was meant not for job performance assessment or for therapy, but for creating a virtual environment for assessing skills acquired from a preventive intervention. Because creating a non-virtual assessment environment for these skills is next to impossible (for ethical reasons), synthetic characters represented a logical potential solution.

CURRENT APPLICATION

We used synthetic character vignettes to assess the success of an intervention teaching at-risk adolescents social skills. By using synthetic character vignettes we observed how participants actually behaved, rather than elicited how they claimed they would behave, in those situations. A synthetic teenager of similar gender and ethnicity to study participants was created to simulate provocative interpersonal social situations in a school setting. The participants spoke naturally to the synthetic character while an observer immediately categorized the input. The character's response, however, was controlled by behavior models underlying the simulation, not by an observer.

Thus, we needed to design vignette storylines to target social skills: engaging in negotiation, maintaining emotional control and being non-provocative, and seeking information. Consequently, we also needed to design language models so that when participants verbally interacted with the synthetic character, the natural language system would interpret their input as reflecting one or more social skill. Similarly, we designed behavior models so that the character would attempt different strategies to require participants to demonstrate skills. Finally, we devised methods to measure participants' engagement compared to what their engagement would be during role-play.

A screenshot of the synthetic character is shown in Figure 2. The participants were African-American male adolescents in 10th grade. Approximately half of the participants met diagnostic criteria for conduct disorder (CD) and half did not have symptoms of CD or any other reported behavioral problem (e.g., truancy, aggression). All of the participants were living in low-income, inner-city neighborhoods.

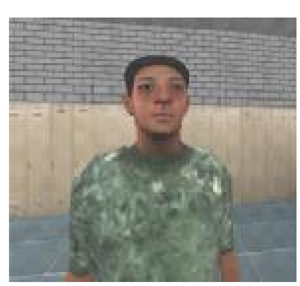

Figure 2. Synthetic Character

The study evaluated a video component of a violence prevention program developed for inner-city African-American youth that provides adolescents with negotiation and conflict-resolution skills training (Yung & Hammond, 1998). Our study focused on short-term effects of the videotape on adolescents for these skills. Participants provided baseline data by running through three synthetic character vignettes, and then approximately two months later a random half of the participants was assigned to watch the 30-minute preventive intervention videotape. Participants who watched the videotape were asked several structured questions during and following the videotape to ensure that they had been adequately exposed to the negotiation skills training information and role playing presented in the videotape. Participants then provided experimental data for behavior and decisions given specific risk scenarios as targeted outcome measures by running through three additional synthetic character vignettes.

The procedure while participants engaged the synthetic character was straightforward. The participant sat in front of a monitor watching the synthetic character and spoke naturally (via a microphone) with the synthetic character. As described below, the natural language processing was disabled, and an observer categorized the participant's speech. Also as described below, an observer rated the participant's ongoing interpersonal and conflict-resolution skills.

Implementation Details

To create the application required visualization, language, and behavioral modeling.

For visualization, a full-body, animated synthetic character was developed to appeal to the target population. A simple school hallway was also created. For the character we generated a base head model for the desired age and ethnicity as well as morph targets for expression and lip synching, and animated and applied motion capture data. We then generated and applied textures and mapping coordinates for both the character and virtual environment.

For language input processing the architecture in general uses the following components: an off-the-shelf speech recognizer; a minimum distance translator that parses the recognizer output, matching to the closest grammatical sentence; dynamically selected grammar files that feed into the speech recognizer and also the parser, categorizing syntactic elements into semantic (i.e., meaningful) categories; and emotional and social tagging that carry information related to emotional or social state variables maintained by the behavior engine. Language output is achieved either by text-to-speech generation (as was the case in this study) or by using pre-recorded files.

Specifically for this application, though, since we had an observer categorize the participant's speech, we bypassed the first two stages of language input processing. (We were concerned with speech recognition, despite the creation and testing of grammar files.) Hence, there was no need to derive a recognized string of words from the participant's input and assign meaning from the string of words; the observer performed these functions. The observer was trained to listen for the following from the participant: agreement, seeking information, demanding compensation or reward, ambivalence, negotiation, disagreement/opposition, or refusal. Every participant utterance was immediately assigned to one of these categories.

For behavior modeling the architecture uses an augmented transition network to map the semantic interpretation to synthetic character behavior. We allowed a fuzzy logic in guiding the synthetic character's behavior; there were often multiple conditional transitions that could be satisfied. That is, our subject-matter experts told us how the synthetic character should behave given the participant's input, but we purposefully allowed for under-specification. Based on the input categorized by the observer, we maintained state variables in the behavior engine that tracked how well the participant was following what social skills were prescribed by the videotape. Variables that we maintained included compromise or negotiation, detachment from the situation, regulation of emotions, seeking of information, understanding the synthetic character's viewpoint, and verbalizing feelings and intentions. As these variables (and others, including the length of the conversation and the appropriateness of each participant response) changed, they caused different transitions to be taken and different behaviors (i.e., replies or requests, gestures) to be exhibited by the synthetic character.

More architectural details can be found in Hubal, Frank, & Guinn (2003).

Application Details

Based on the content of role plays in the prevention program, we developed scripts for the synthetic character in a school hallway to entice study participants to engage in risky decision-making and exhibit impulsive behavior and insensitivity to penalties. The scripts included provocative introductory statements and multiple response options for the participant (i.e., those categorized by the observer). For instance, in one vignette the synthetic character asked the participant to keep a gym bag in his locker without telling the participant what the gym bag contained. If the participant refused, the synthetic character offered money and insisted that the participant owed him a favor. Study participants thus had the opportunity to ask questions and then use refusal and negotiation skills to resolve the situation. We also developed an algorithm by which the synthetic character would initially entice the risky behavior but gradually back off if the participant demonstrated appropriate avoidance and/or de-escalation behavior (i.e., information seeking, negotiation). The algorithm took into account how long the interaction had been going on, how insistent the participant was in refusing to be enticed by the synthetic character, and the content of the conversation. Finally, we modeled the language that the synthetic character used on that used by adolescents in the study sample, though for the present study we used computer-generated speech rather than pre-recorded speech. The synthetic character remained in one position because the entire interaction was based primarily on conversation, so the gestures were not complex, relying mostly on beat gestures and idle motions. However, as the synthetic character became more agitated or aggressive, the gestures became more representational (e.g., pointing, placing the hands on hips), depending on the content of the conversation.

Six virtual confrontational situations – three pairs of two – were developed for the study. Three were presented to participants as a pre-test (before any prevention materials were presented, for those participants in the experimental group), and three were presented to participants as a post-test. In one situation, the synthetic character asked the study participant to keep a gym bag with something inside in his school locker, but did not provide any information about what was in the gym bag. Participants who asked about the contents within the bag were told by the synthetic character not to worry about what was inside the bag and to keep the bag. However, if the participant inquired further, he discovered that either a pair of sneakers or a wallet was in the gym bag. Participants who declined to keep the bag were enticed with a bribe from the synthetic character. Although never explicitly stated by the synthetic character, the sneakers and wallet were stolen goods and participants who agreed to keep the bag were taking a risk by keeping them. In a second situation, the synthetic character invited the participant to join in a known prohibited activity (either coming to a drinking party with girls or going for a joy ride). A request for further information or a rejection of the offer elicited persuasive, high pressure responses from the character: *C'mon, there'll be lots of free booze. These girls are gonna be hot. You don't got no girlfriend.* In a third situation, the synthetic character initiated a confrontation (accusing the participant of bumping into him in the hallway or of messing with his girlfriend), and tried to provoke a fight. If the participant backed off, rejecting offers to fight and defusing the situation, the synthetic character attempted to exacerbate the situation by making cutting remarks: *I'll kick your butt! Don't cry. You must be afraid a your own shadow.* Participants responding to the provocation in an inflammatory manner further aggravated the synthetic character. All of these situations represent actual experiences of urban minority adolescents that are analogous to role plays used in prevention programs (Paschall, *et al.*, in press; Yung & Hammond, 1998). The maximum number of conversational turns for each situation (i.e., vignette) was 10-12, but only if the participant sought all the information that the synthetic character could provide.

LESSONS LEARNED

Tools Used to Gather Modeling Data

We used iterative subject-matter expert guidance and feedback to model synthetic character behavior. The user-character interactions were developed and repeatedly refined based on this input.

We based visualized behavior (e.g., whole body, head, and arms/hands movement, facial expression) on that shown in the preventive intervention videotape as well as that exhibited by similar participants in other studies conducted by our experts. We based linguistic behavior on the kind of language that these adolescents typically encounter, as well as on appropriate generated responses (replies, questions, challenges, denials) to what was interpreted from the participant's input. In other applications, when available, we've also used recordings taken from naturalistic settings, direct observation, and literature on typical behavior characteristics (gesture, facial expression, language, conversational pragmatics).

We based synthetic character cognition and emotion models on the character's intentions in each vignette. We considered what would be known about the world, and how the character should reason about social roles and conventions (what can be stated or asked at a point in the dialog, and how it gets stated or asked), given the participant's input. We updated state based on syntactic analysis (i.e., the input format; command, request, statement, threat), lexical analysis (politeness, degree of compromise, detachment), and semantic content (topic relevance, sensitivity).

The current application fell short in realism in its use of motion capture and text-to-speech. Specifically, motion capture files that we used to derive our character animations were based on adult male movements rather than inner-city adolescent male movements. And the character's verbal output used computer-generated speech instead of pre-recorded speech, hence did not carry the intonation and cadence that these adolescents use.

Results

Measures of decision-making and social competency skills such as information seeking and negotiation were based on ratings of participants' verbal and non-verbal interaction with the synthetic character along seven dimensions: level of engagement, verbalization, emotional control, information seeking, expressing own preferences, compromise/negotiation, and being non-provocative. Demonstrated skill along each dimension ranged from 'very low' to 'moderate' to 'very high'. Two trained researchers independently observed each adolescent's performance on three vignettes. The correlation between the two raters' scores was very high (r = 0.94 to 0.99). We then compared these measures against established methods that measure the same skills (e.g., teacher ratings, text-based scenario responses). Results indicated that exposure to the videotape was positively associated with measures of negotiation skills, and that participant behavior while engaged with the synthetic character had adequate construct and criterion validity (see Paschall, *et al.*, in press).

Results also indicated a significant interaction between participants exposed to the videotape and participants diagnosed with CD, on both conflict-resolution and interpersonal communication skills, with stronger videotape effects observed among participants without CD. Data from the set of experimental participants (27 adolescents diagnosed with CD, 30 controls), as well as a matched control group of 65 participants who were not presented with the intervention, were analyzed for the outcome from the vignettes (see Figure 3). The vignettes were sufficiently engaging that during the interaction – even after possibly having gone through a portion of the prevention program – a sizeable percentage of participants demonstrated potentially risky behavior. For the CD group, on average, without any intervention, the participant agreed to help, go along with, or escalate a confrontation with the synthetic character (all risky behaviors) over half the time. This percentage dropped for the post-test, but not differently for those participants who had been given the intervention compared to those participants who had not been. For the control (i.e., non-CD) group, the participant exhibited risky behavior almost one third of the time on average before any intervention but one sixth of the time after an intervention. The two groups differed from each other significantly, suggesting that the vignettes were realistic enough to be able to discriminate between behavioral groups. That is, participants who were exposed to the preventive materials and adopted their content were significantly more likely to use positive interaction skills during the vignettes than those not exposed or who did not adopt the content, indicating some effectiveness of the intervention but also of the vignettes as a valid assessment.

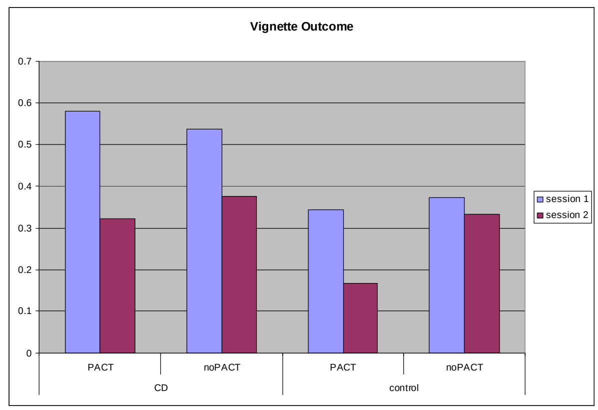

Figure 3. Experiment Results

We also analyzed participant behavior in terms of their engagement with the vignettes, their body language, verbalizations, number of conversational turns, and response time. There were no differences among groups, but the behavioral performance measures reflected the participants' acceptance of the vignettes. For instance, the participants demonstrated a moderate amount of body language while using the vignettes, they performed verbalizations while using the application beyond their inputs to the synthetic character, and they conducted conversations with an average of four conversational turns. (This last finding is important because the language grammars are shallow and narrow; the synthetic character only has about 8-10 different topic responses per vignette, not including a few generic responses.)

We considered vignette outcome and the length of the interaction. For the party vignette, participants who declined took longer, suggesting that they were seeking information. Controlling for interaction length, diagnosis of CD was related to making a bad decision. For the stolen goods and confrontation vignettes, those participants diagnosed with CD tended to make bad decisions.

About three-fourths of the participants were asked if their virtual decision mirrored what would be their real-life decision, if the researcher was not present. Importantly, 90% of participants who were asked felt that this was the case; this was true for both control and CD participants, and was also the case regardless of whether or not the participant agreed with or refused the synthetic character.

FUTURE APPLICATIONS

This approach, using synthetic characters for assessing interaction skills, has application to military training beyond the PTSD example given above, particularly for skills needed in stability and support operations. Hence we conclude by suggesting when and how current synthetic characters might be used for assessment.

First, though, we stress that use of synthetic characters should be considered just one of several valid approaches to assessment. As Table 1 demonstrates, all of the approaches have advantages and disadvantages that should guide their use. For instance, if the need is to quickly gauge soldiers' knowledge about skills, rather than their application of skills, then questionnaires would be cheaper and easier to implement than assessment via synthetic characters. Alternatively, if the need is for high-stakes, highly-realistic, intense simulation, then role-plays embedded into constructive or live environments would be a better choice than virtual role-plays.

Still, we have identified many situations where the need is primarily to assess interaction skills in environments or with individuals that are difficult to replicate in any but a virtual environment. For instance, there are situations where a soldier might come into contact with a child, perhaps as a combat casualty (Lindheim, & Swartout, 2001; Perabo, 2003) or perhaps as a child combatant (Borchini, Lanz, & O'Connell, 2002). In these situations the soldier may need to apply any combination of procedural (e.g., attending to the child medically), strategic (e.g., assessing the need to engage the child soldier in a firefight), and interaction (e.g., calm the child down) skills. It is difficult to hire child actors to portray these roles reliably and consistently, and a 'shoot / don't shoot' device doesn't have the interactivity needed, yet certainly having the soldier just describe what s/he would do, rather than demonstrate skills, is inappropriate. A virtual role-play involving child synthetic characters can overcome some of these issues. Similarly, operations other than war, including manning a checkpoint and crowd control, often involve coming into contact with a range of individuals who vary by age, ethnicity, culture, gender, personality, trauma, and/or mental state. Film clips can convey some of this variability, but virtual role-plays can involve any number of synthetic characters varying along any number of dimensions. Numerous computer-generated forces applications are relevant here (e.g., Mui, *et al.*, 2003; Reece, *et al.*, 2002; Stytz, & Banks, 2003; Wray, *et al.*, 2002; Wray & Laird, 2003).

We are currently attempting to replicate the results we obtained with adolescents with prison inmates. Use of synthetic character applications with dangerous populations, that also include schizophrenics (Frank, *et al.*, 2002) and volatile persons encountered during peace enforcement or terrorist events (Steadman, *et al.*, 2001), can serve as training and assessment for military needs involving skills dealing with some asymmetric threats. Further, assessment using synthetic characters represents a much more realistic measure of tacit knowledge for military leadership than codified surveys (Hedlund, Sternberg, & Psotka, 2000). As the visual, emotional, linguistic, and expressive technologies advance, we envision the use of synthetic characters to move into field and clinical assessment as well.

ACKNOWLEDGEMENTS

We thank the many individuals who contributed significantly to this work. This work was performed under funds from NIDA award #5-R01-DA14813-02 and from NSF award #EIA-0121211. Points of view in this document are those of the authors, and do not necessarily represent the official position of any of the above-listed agencies.

REFERENCES

Borchini, C., Lanz, S., & O'Connell, E. (2002). *Child Soldiers: Implications for U.S. Forces*. Quantico, VA: Center for Emerging Threats and Opportunities, Marine Corps Warfighting Laboratory.

Brown, J.S., Collins, A., & Duguid, P. (1989). Situated Cognition and the Culture of Learning. *Educational Researcher*, *18*, 32-41.

Cassell, J., Sullivan, J., Prevost, S., & Churchill, E. (2000). *Embodied Conversational Agents*. Cambridge, MA: MIT Press.

Dautenhahn, K. (1997). Socially Intelligent Agents. *Working Notes of the AAAI Fall Symposium*. Menlo Park, CA: AAAI Press.

Fletcher, J.D. (1991). Effectiveness and Cost of Interactive Videodisc Instruction in Defense Training and Education. *Multimedia Review, 2*, 33-42.

Frank, G., Guinn, C., Hubal, R., Pope, P., Stanford, M., & Lamm-Weisel, D. (2002). JUST-TALK: An Application of Responsive Virtual Human Technology. *Proceedings of the Interservice/Industry Training, Simulation and Education Conference*, *24*, 773-779.

Guinn, C., Hubal, R., Frank, G., Schwetzke, H., Zimmer, J., Backus, S., Deterding, R., Link, M., Armsby, P., Caspar, R., Flicker, L., Visscher, W., Meehan, A., & Zelon, H. (2004). Usability and Acceptability Studies of Conversational Virtual Human Technology. *Proceedings of the SIGdial Workshop on Discourse and Dialogue*. East Stroudsburg, PA: Association for Computational Linguistics.

Harless, W.G., Zier, M.A., Harless, M.G., & Duncan, R.C. (2003). Virtual Conversations: An Interface to Knowledge. *IEEE Computer Graphics and Applications*, *23*, 46-52.

Hayes-Roth, B. (2003). Online Workshops Training Brief Intervention in Alcohol Abuse. *Proceedings of the CyberTherapy Conference*, *9*, 43-44.

Hedlund, J., Sternberg, R.J., & Psotka, J. (2000). *Identifying the Abilities Involved in the Acquisition of Tacit Knowledge*. Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Hubal, R.C., & Frank, G.A. (2001). Interactive Training Applications using Responsive Virtual Human Technology. *Proceedings of the Interservice/Industry Training, Simulation and Education Conference*, *23*, 1076-1086.

Hubal, R.C., Frank, G.A., & Guinn, C.I. (2003). Lessons Learned in Modeling Schizophrenic and Depressed Responsive Virtual Humans for Training. *Proceedings of the Intelligent User Interface Conference*. New York: ACM Press.

Hubal, R., & Guinn, C. (2003). Interactive Soft Skills Training using Responsive Virtual Human Technology. *Proceedings of the Interactive Technologies Conference*. Arlington, VA: Society for Applied Learning Technology.

Hubal, R.C., Kizakevich, P.N., Guinn, C.I., Merino, K.D., & West, S.L. (2000). The Virtual Standardized Patient – Simulated Patient-Practitioner Dialogue for Patient Interview Training. In J.D. Westwood, H.M. Hoffman, G.T. Mogel, R.A. Robb, & D. Stredney (Eds.), *Envisioning Healing: Interactive Technology and the Patient-Practitioner Dialogue*. Amsterdam: IOS Press.

Kizakevich, P.N., McCartney, M.L., Nissman, D.B., Starko, K., & Smith, N.T. (1998). Virtual Medical Trainer: Patient Assessment and Trauma Care Simulator. In J.D. Westwood, H.M. Hoffman, D. Stredney, & S.J. Weghorst (Eds.), *Art, Science, Technology: Healthcare (R)evolution*. Amsterdam: IOS Press.

Kizakevich, P.N., Lux, L., Duncan, S., Guinn, C., & McCartney, M.L. (2003). Virtual Simulated Patients for Bioterrorism Preparedness Training. In J.D. Westwood, H.M. Hoffman, G.T. Mogel, R. Phillips, R.A. Robb, & D. Stredney (Eds.) *NextMed: Health Horizon*. Amsterdam: IOS Press.

Kulka, R.A., Schlenger, W.E., Fairbank, J.A., Hough, R.L., Jordan. B.K., Marmar, C.R., & Weiss, D.S. (1990). *Trauma and the Vietnam War Generation*. Report of Findings from the National Vietnam Veterans Readjustment Study. New York: Brunner/Mazel.

Lave, J., & Wenger, E. (1991). *Situated Learning*. New York: Cambridge.

Lindheim, R., & Swartout, W. (2001). Forging a New Simulation Technology at the ICT. *Computer*, *34*, 72-79.

McFall, M., Fontana, A., Raskin, M., & Rosenheck, R. (1999). Analysis of Violent Behavior in Vietnam Combat Veteran Psychiatric Inpatients with Posttraumatic Stress Disorder. *Journal of Traumatic Stress*, *12*, 501-517.

Mui, R.C.Y., LaVine, N.D., Bagnall, T., Sargent, R.A., Goodin, J.R., & Ramos, R. (2003). A Method for Incorporating Cultural Effects into a Synthetic Battlespace. *Proceedings of the Behavior Representation in Modeling & Simulation Conference*.

Orlando, FL: Simulation Interoperability Standards Organization.

Olsen, D.E., Sellers, W.A., & Phillips, R.G. (1999). The Simulation of a Human Subject for Interpersonal Skill Training. Proceedings *of the Interservice/Industry Training, Simulation and Education Conference*, *21*.

Paschall, M.J., Fishbein, D.H., Hubal, R.C., & Eldreth, D. (in press). Psychometric Properties of Virtual Reality Vignette Performance Measures: A Novel Approach for Assessing Adolescents' Social Competency Skills. *Health Education Research: Theory and Practice*.

Perabo, B. (2003). *The Innocent Enemy: Children at War and the Boundaries of Combatancy*. Washington, DC: Joint Services Conference on Professional Ethics.

Reece, D., Brett, B., Franceschini, D., & Hursh, S. (2002). ModSAF as a Model of Cognition. *Proceedings of the Conference on Computer Generated Forces and Behavioral Representation*. Orlando, FL: Simulation Interoperability Standards Organization.

Rosenberg, S.D., Mueser, K.T., Friedman, M.J., Gorman, P.G., Drake, R.E., Vidaver, R.M., Torrey, W.C., & Jankowski, M.K. (2001). Developing Effective Treatments for Posttraumatic Disorders Among People With Severe Mental Illness. *Psychiatric Services*, *52*, 1453-1461.

Schlenger, W.E., Caddell, J.M., Ebert, L., Jordan, B.K., Rourke, K.M., Wilson, D., Thalji, L., Dennis, J.M., Fairbank, J.A., & Kulka, R.A. (2002). Psychological Reactions to Terrorist Attacks: Findings from the National Study of Americans' Reactions to September 11. *Journal of the American Medical Association*, *288*, 581-588.

Sims, E. (2003). Using Extensible 3D (X3D) to Repurpose Modeling and Simulation Assets for Advanced Distributed Learning. *Working Notes of the Spring Simulation Interoperability Workshop*. Orlando, FL: Simulation Interoperability Standards Organization.

Southwick, S.M., Morgan, C.A., Darnell, A., Bremner D., Nicholaou, A.L., Nagy, L.M., & Charney, D.S. (1995). Trauma-Related Symptoms in Veterans of Operation Desert Storm: A 2-year Follow-up. *American Journal of Psychiatry*, *152*, 1150-1155.

Steadman, H., Griffin, P., Smith, K., Draine, J., Dupont, R. & Horey, C. (2001). A Specialized Crisis Response Location as a Core Element for Police-based Diversion Programs. *Psychiatric Services*, *52*, 219-222.

Stytz, M.R., & Banks, S.B. (2003). Progress and Prospects for the Development of Computer-generated Actors for Military Simulation: Part I – Introduction and Background. *Presence: Teleoperators and Virtual Environments*, *12*, 311-325.

Suler, J. (2000). Psychotherapy in Cyberspace: A 5-Dimension Model of Online and Computer-mediated Psychotherapy. *CyberPsychology and Behavior*, *3*, 151-160.

Visscher, W.A., Hubal, R.C., Guinn, C.I., Studer, E.J., & Sparrow, D.C. (to appear). A Synthetic Character Application for Informed Consent. *Proceedings of the AAAI Fall Symposium on Dialogue Systems for Health Communication*, 22-24 October, 2004, Arlington, VA.

Wiederhold, B.K., & Wiederhold, M.D. (2004). The future of Cybertherapy: Improved Options with Advanced Technologies. In G. Riva, C. Botella, P. Légeron & G. Optale (Eds.), *Cybertherapy: Internet and Virtual Reality as Assessment and Rehabilitation Tools for Clinical Psychology and Neuroscience*. Amsterdam: IOS Press.

Wray, R.E., Laird, J.E., Nuxoll, A., & Jones, R.M. (2002). Intelligent Opponents for Virtual Reality Trainers. *Proceedings of the Interservice/Industry Training, Simulation and Education Conference*, *24*, 850-858.

Wray, R.E., & Laird, J.E. (2003). Variability in Human Behavior Modeling for Military Simulations. *Proceedings of the Behavior Representation in Modeling & Simulation Conference*. Orlando, FL: Simulation Interoperability Standards Organization.

Yung, B., & Hammond, R. (1998). Breaking the Cycle: A Culturally Sensitive Prevention Program for African American Children and Adolescents. In J. Lutzker (Ed.), *Handbook of Child Abuse Research and Treatment*. New York, NY: Plenum Publishing.