

Virtual Humans for Interpersonal and Communication Skills' Training in Crime Investigations

Konstantinos Mykoniatis^{1,2}, Anastasia Angelopoulou^{1,2}, Michael D. Proctor²,
and Waldemar Karwowski²

¹ Modeling & Simulation Graduate Program, University of Central Florida,
Orlando, FL 32816, USA

{kmykoniatis, aangelopoulou}@knights.ucf.edu

² Institute for Advance System Engineering,
Department of Industrial Engineering and Management Systems,
University of Central Florida, Orlando, FL 32816, USA

{Michael.Proctor, wkar}@ucf.edu

Abstract. Virtual Humans (VHs) have been employed in multidisciplinary fields to advance interpersonal skills critical to many professional, including law enforcement agents, military personnel, managers, doctors, lawyers and other professionals. Law enforcement agencies in particular have faced a growing need to develop human to human interpersonal training to increase interviewing and interrogation skills. In this paper, we present a prototype VE that has been developed to provide law enforcement agents with effective interview and interrogation training and experiential learning. The virtual training environment will need to be tested and formally evaluated to verify the benefits compared to live exercises and traditional training techniques.

Keywords: Virtual Human, Training, Law enforcement agents, Interpersonal Skills, Virtual Environment.

1 Introduction

Virtual Humans (VHs) are the subjects of research in multidisciplinary fields, ranging from computer sciences (computer generated forces), healthcare and education industries to military and entertainment applications. Even though VH research has evolved rapidly over the last years (Gratch et al. 2002; Rickel et al. 2002; Swartout 2006, Swartout 2010), few studies have focused on how to employ VHs to advance interpersonal skills through training in interviewing, negotiating, leadership, cultural awareness, tactical questioning, or eliciting information (Hubal et al. 2001; Kenny et al. 2007). These interpersonal skills are critical to many professional, including law enforcement agents, military personnel, lawyers, doctors, managers, supervisors and other professionals.

Law enforcement agencies in particular have faced a growing need to develop human to human interpersonal training to increase interviewing and interrogation skills.

This kind of training is traditionally conducted through live exercises, “talk and chalk” lectures, role playing, group discussions or non-experiential media learning. These techniques have led to a critical training gap since students are provided with limited practice time and limited variety of scenarios (Hubal et al. 2001). Regarding interrogation training, agencies typically use a variety of techniques and tools to train investigators in developing skills to identify deception indicators, ranging from verbal communication to non-verbal cues. This kind of training is time consuming as it is accomplished through practical interrogation with known suspects of committed crimes (Luciew et al. 2011). Although some existing Virtual Environments (VE) provide physical skills' training, team training, or even strategy and tactics, their overall limitation is that human user's interactions with the VHS are usually restricted to shooting activities (Kenny et al. 2007).

The purpose of this paper is to help fill the gap in the interpersonal skills training in the field of law enforcement. To this purpose, a prototype VE has been developed to provide law enforcement agents with effective interview and interrogation training and experiential learning. The prototype VE was built in Unity3D Pro, using, Maya 2013, 3DsMax, SonicAcid Pro and the VH toolkit which was provided by the University of Southern California (USC) Institute for Creative Technologies (ICT). The developed VE simulates a variety of aspects, from the physics of objects to realistic human behavior. The VHS have three attributes: they are believable, i.e. able to mimic human behavior and express realistic emotions in order to draw the human user into the scenario; they are responsive, meaning they are capable of interacting with trainees and other VHS via text or spoken dialogues; and they are interpretable, so that the user will be able to interpret the verbal and nonverbal cues in their responses.

2 Existing Virtual Environments for Training Interpersonal Skills

In this section, we review existing VEs in the field of crime investigation. If properly used, VEs can be useful training tools. VEs have been used for intercultural communication training (Kim et. al 2009), clinical interviewing (Kenny et al. 2008), and police officer training (Hubal et al. 2003), to name a few. Some examples of VEs developed for training interpersonal skills are briefly described below.

The JUST-TALK training system (Fig. 1) was developed to assist law enforcement agents in managing situations involving mentally ill people and responding appropriately. The overall objective of the JUST-TALK project was to improve law enforcement training using Natural Language Processing and Virtual Reality technology (Frank et al. 2002).

Another VE training example is the Tactical-Questioning system (Kenny et al., 2007), which allows trainees to interview a suspect of a bombing incident (Fig. 2).



Fig. 1. The JUST-TALK Virtual Environment (Source: Frank et al., 2002)



Fig. 2. Tactical Questioning System (Source: Kenny et al., 2007)

Finally, an immersive criminal investigator prototype developed for training investigations of child physical and sexual abuse is presented (Luciew et al. 2011). This system has two modes: the "interview training" mode, which enables investigators to experience a virtual interview with a victim of child abuse (Fig. 3); and the "interrogation training" mode, which allows investigators to virtually interrogate a suspect of sexual assault (Fig. 4).



Fig. 3. A virtual interview subject (Source: Luciew et al., 2011)



Fig. 4. A virtual interrogation subject (Source: Luciew et al., 2011)

3 Attributes of the VE Training System Prototype

Our goal is to build an interactive VE and VHS that will improve interpersonal skills training to benefit Law Enforcement Agents, i.e. advance their interviewing and interrogation techniques, critical skills for the success in the current working environment. We propose an interactive training tool to teach investigators how to conduct effective interviews. The proposed tool allows for natural and interactive dialog between the trainee and the VH.

The system has two modes: the "pre-interview mode" and the "training segment" mode. In the "pre-interview mode", the trainee is first introduced to a simulated scenario. The scenario is like a short video clip, which has the purpose of providing information and evidence to the trainee about the criminal case. The simulated scenario of the prototype follows below. The owner of a house has been murdered during a Halloween Party. Then, the trainee is presented with a list of possible suspects and applicable "police" profiles, a list of evidence from the scene and other relevant information and events occurred before and after the murder, as depicted in Fig.5 and Fig. 6.

After reviewing potential suspects and taking appropriate notes, the trainee is presented with a target "training segment." In the "training segment", the trainee can interact with the VHS and interview one VH suspect at a time until a conclusion is made.

The *VE Training System Prototype* allows for the following interactions:

- The trainee can navigate within the VE to collect data, crime evidence and cues (Fig. 7).
- The trainee can interview VHS about the crime, i.e. the guests of the house, in order to determine their involvement. Each VH is capable of answering a number of questions that are relevant to the investigation case.
- The trainee can observe verbal and non-verbal behaviors, and deception cues by interacting within the VE and VHS. The ability to detect deception cues is a critical skill for law enforcement agents. The proposed training system is capable of augmenting this type of skill.

POLICE PROFILE

	NAME	John Morgan
	DOB	January 13, 1972
	HEIGHT	1 meter, 80 cm
	WEIGHT	112 kg
	HAIR	Grey
	EYES	Brown
	Writes	Left hand

Notes: **John Morgan** is Brad's three years older cousin. He was M&S Inc. employee since 2000. John is the Chief Human Resources Officer of M&S Inc. He hired **Jenny Miller** six months ago. In his prior report he claimed that Brad had problems with his wife, who was jealous of his new secretary.

POLICE PROFILE

	NAME	Jenny Miller
	DOB	July 20, 1986
	HEIGHT	1 meter, 68 cm
	WEIGHT	60 kg
	HAIR	Brown
	EYES	Green
	Writes	Right Hand

Notes: **Jenny Miller** was born in Pennsylvania. **Jenny** had been **Brad's** private secretary for six months. She was hired to M&S Inc. by **John Morgan**, from Human Resources Department. In her prior reportshe said : "I can't believe my boss was murdered!"

POLICE PROFILE

	NAME	Rachel Anderson Morgan
	DOB	February 8, 1980
	HEIGHT	1 meter, 65 cm
	WEIGHT	67 kg
	HAIR	Blond
	EYES	green
	Writes	right hand

Notes: **Rachel Anderson** had been married to Brad for 3 years. Rachel has inherited M&S Inc. from her father. Rachel found her husband dead and called 911 at 21:45. In her prior report she was grieving... " This can't be real! I can't believe my love is really dead! She did it, Jenny did it!"

POLICE PROFILE

	NAME	Jack Morris
	DOB	January 13, 1955
	HEIGHT	1 meter, 85 cm
	WEIGHT	112 kg
	HAIR	Grey
	EYES	Brown
	Writes	Left hand

Notes: **Jack Morris** was a friend of **Rachel's** father, and the Chief Financial Officer (CFO) of M&S Inc. In his prior report he said: "I was enjoying the music on the living room, when the thunderbolt stroke and then...blackout! While I was searching for a lighter, I heard two gun shots, BANG! BANG! Oh my God! I can't believe it! Brad my boy! You can't be dead! Please officer, forgive me I am shocked..."

Fig. 5. "Police Profiles" (the VH suspects) in the Prototype

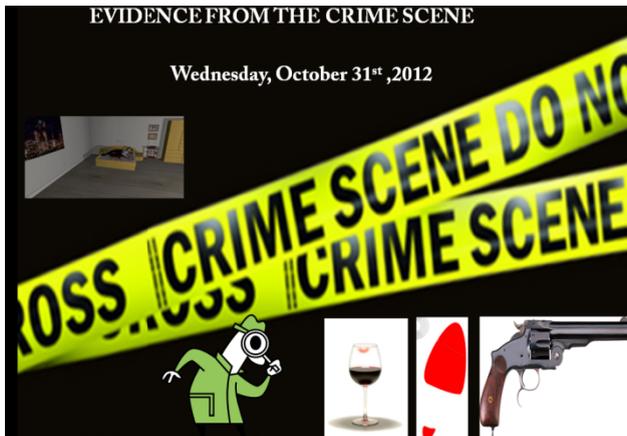


Fig. 6. Collected Evidence Screenshot in the Prototype



Fig. 7. Scenes taken from the VE Prototype

A stochastic model produces responses to the trainee's questions, which are selected from a pre-scripted list of possible questions. The VE enables the trainee to gain experience in asking proper questions and distinguishing between deceptive and truthful responses.

4 Prototype Focus on Deception Cues

In the developed prototype interviewing skills training is a major component. The opportunity for a realistic interaction with subjects capable of demonstrating cues related to truthfulness and deceit is a proven benefit. Research indicates that the virtual training environment can improve trainees' ability to detect deception by helping them identify the right cues (Lane et al. 2010). The trainee should ask the proper questions and identify common cues of truth-telling and deceit. Cues to deception are commonly divided into the three categories: verbal, nonverbal, and vocal. Verbal cues come from the content of the speaker's statement. Examples include declaration of lack of memory or self-references. Nonverbal cues include eye contact, posture, hand movement and can be observed individually by the suspect's behavior. Vocal cues can be a sound or the tone of the suspect's voice used to send a particular message and not an actual word. Vocal cues are significant, since they may help the trainee understand the message behind the words by either confirming or discrediting the message.

A person's body language is a form of mental and physical ability of human non-verbal communication, consisting of body posture, gestures, facial expressions, and eye movements, which can tell whether a suspect may lie (DePaulo, et al. 2003). In the prototype the VH characters can perform some of the natural human characteristics via the integrated SmartBody system within the VHToolkit. Deception indicators include self-fidgeting, posture shifts, response length, and verbal and vocal uncertainty. The simulated VHs may have a nervous habit, such as fidgeting their hair or wringing their hands, which can indicate when they might lie. Other VHs may exaggerate their emotions and then suddenly return to normal. A VH suspect who is lying is more likely to take a defensive posture, such as crossing his/her arms and remaining stiff. Fig.8 illustrates a sample of different postures, gestures, facial expressions, and eye movements of a VH character.



Fig. 8. Examples of VH's nonverbal behavior in the VE Prototype

The first step in creating the VHs' profiles was to decide on incidents that would provide the environment for the interviews. The law enforcement scenario that was chosen for the simulation was a murder. Four VH characters were used as the main suspects. In this version of the system, the VHs use basic nonverbal behaviors in their speech, but they are not high fidelity modeled to deceive. For instance, the VHs do not have facial expressions or eye movement based on their emotional state.

The trainee can interact with the VHs by asking questions via either an integrated typed interface or an automated speech recognition system that provides animated and oral responses from the VHs. The training focus is primarily on helping the trainee to:

- Ask the appropriate questions (Socratic Method, Improve skill of investigative interviewing or interrogation of non-cooperative suspects).
- Identify deception cues (Verbal, Non Verbal, and Vocal).

5 Intelligent Tutoring Capability in Prototype System

The proposed system is capable of supporting an intelligent tutoring system. The *first phase* of the developed prototype focused on the construction of the VE, VHs, and interactive dialogues.

The *second phase* would focus on the effective implementation of intelligent tutor's elements, such as:

- Levels of interactions (novice, intermediate, expert). Different simulated scenarios would be provided for different level of training experience
- Feedback on the accomplished tasks
- Hints

- Applying Socratic instructional theory of learning to improve questioning skills
- Cumulative scoring and analysis of trainees' solutions to determine trainees' progress.

The training process would be based on the Socratic Learning Theory (Woolf, 2001). This learning theory is a teaching strategy derived from Socrates (Greece 469-399 B.C) and it is based on the belief that each person contains the essential ideas and answers to solve problems. By applying the Socratic Method trainees would improve their questioning skills and seek to discover the subject of matter, since their goal is to know what they do not know.

The intelligent tutor would assist the trainees in finding out and correcting their own misconceptions by asking the trainees instructive questions. The tutor would also track the trainee's performance and record, analyze and evaluate the questions, decisions and actions without interfering with the trainee's efforts. The tutor would be designed to motivate and provide an After Action Review (AAR) of the trainee's performance after the completion of the scenario to validate the skills the trainee has acquired. The tutor would also be able to provide oral feedback to the student after the interview is completed, by making suggestions to the trainee for future actions.

Finally, the *third phase* would include a pilot testing and evaluation of the effectiveness of the intelligent tutoring system. The evaluation findings will be judged by law enforcement key stakeholders and this information will be reused to improve the existing system.

6 Audio in the VE Prototype

As the visuals of the simulated scenario unfold, audio and sound effects can convey emotions such as excitement, and fear. According to Brown (1994), "*it is the combination of the visuals with music that makes the viewers feel those emotions and affects the users' subjective sense of presence or "being there".*" For the audio production of this project, we used the Sony ACID Pro, which is a professional digital audio workstation software program. Audio tracks were produced for:

- VE background music
- VHs realistic voices and vocal cues
- Natural Phenomena sound effects, such as raining and thunder strike sound effects
- Audio sounds visible within the VE, such as VHs' voices and vocal cues, footsteps, door-opening, gunshot, CD/radio music player.

"*Music is used in two modes within films: diegetic and non-diegetic*" (Robertson et al., 1998). Diegesis is a Greek term for "recounted story". In other words, the diegesis is the total world of the story action. Diegetic music can come from a sound source visible on the screen or a source implied to be occurred within the narrative of the virtual scenario. Examples of diegetic sounds in the developed VE are the VHs' voices, sounds made by VH's or other objects in the story, and music attributable to some source coming from instruments within the VE such as Hi-Fi CD/radio music player or jukebox.

Non-diegetic music comes from a sound source which is neither part of the narrative, nor attributable to a source in the virtual scenario. This type of music comes from outside the space of the simulated story events. Non-diegetic sounds can be sound effects, such as narrator's commentary, added to underline a dramatic effect. Non-diegetic music is usually used to affect the emotions of the audience. It can be used to cue the audiences to feel uncomfortable; to build up tension; to point out an approaching disaster or to indicate a love affair. Examples of non-diegetic sounds in the developed VE are the sound effects added to indicate an upcoming argument between VHs or to prepare the trainees for the crime to be committed.

7 Expected Benefits of the Proposed VE

The overall training system aims to help train law enforcement agents to conduct effective interviews and interrogations, to interpret the verbal and non-verbal communication of their subjects, to improve their abilities, such as cognitive, sensory, psychomotor skills (Fleishman et al, 1991) and to cultivate and improve analytic and decision making skills. The system will allow the simulation of a variety of VE scenarios, increasing the trainee's situational awareness, as situations can be repeated over and over with no further training costs.

Trainees will be provided with more practice time, increased access to training and consistent training experience, leading to improvements of their problem-solving abilities. Benefits will also include saving time and reducing training costs, individualized tutoring, realistic and engaging experience to provide valuable law enforcement training.

8 Discussion and Future Work

In this paper, we have discussed the need for training law enforcement agents to conduct effective interviews and interrogations. We reviewed existing VEs and proposed a virtual training environment that could fill the gap in this kind of training using VH suspects of criminal cases. This training system could be used to supplement real-world training, enhancing interviewing skills and reducing training cost, when compared to traditional interpersonal skills training. The weakness of using traditional methods, such as recorded videos or "chalk and talk" lectures when training interpersonal skills, is that traditional methods are not able to provide experiential learning and evaluate the trainee's ability to conduct investigative interviews, even though they can evaluate recognition skills.

Artwork of the VHs, the VE and any background music are important parts in a simulation because they contribute in sustaining the trainee's attention, interest, and motivation. However, these elements are usually underestimated (Kenny et al. 2007). An effort was made to implement a simulated scenario of animated VHs with realistic appearance, natural behavior, believable sound effects and atmospheric background music. Efforts were also made for the simulation of the effects of rain, water, candle fire and lighting.

Future work will focus on the advancement of the prototype system to an intelligent tutoring system. The trainee will be able to experience a variety of real-life training scenarios and will be provided with the time needed to develop the necessary skills, while receiving analytic feedback. The virtual training environment will need to be tested and formally evaluated to verify the benefits compared to live exercises and traditional training techniques.

References

1. Brown, R.S.: *Overtones and Undertones*. University of California Press, Berkeley and Los Angeles (1994)
2. DePaulo, B.M., Lindsay, J.J., Malone, B.E., Muhlenbruck, L., Charlton, K., Cooper, H.: Cues to deception. *Psychological Bulletin* 129(1), 74 (2003)
3. Fleishman, E.A., Mumford, M.D., Zaccaro, S.J., Levin, K.Y., Korotkin, A.L., Hein, M.B.: Taxonomic efforts in the description of leader behavior: A synthesis and functional interpretation. *Leadership Quarterly* 2(4), 245–287 (1991)
4. Frank, G., Guinn, C., Hubal, R.: JUST-TALK: An application of responsive virtual human technology. In: *The Interservice/Industry Training, Simulation & Education Conference (IITSEC)*, vol. 1. National Training Systems Association (January 2002)
5. Gratch, J., Rickel, J., André, E., Badler, N., Cassell, J., Petajan, E.: Creating Interactive Virtual Humans: Some Assembly Required. *IEEE Intelligent Systems* 54–63 (July/August 2002)
6. Hubal, R.C., Frank, G.A.: Interactive training applications using responsive virtual human technology. *Children* 21, 25 (2001)
7. Hubal, R.C., Frank, G.A., Guinn, C.I.: Lessons learned in modeling schizophrenic and depressed responsive virtual humans for training. In: *Proceedings of the 8th International Conference on Intelligent User Interfaces*, pp. 85–92. ACM (January 2003)
8. Kenny, P., Hartholt, A., Gratch, J., Swartout, W., Traum, D., Marsella, S., Piepol, D.: Building interactive virtual humans for training environments. In: *The Interservice/Industry Training, Simulation & Education Conference (IITSEC)*, vol. 2007(1). National Training Systems Association (January 2007)
9. Kenny, P., Parsons, T., Gratch, J., Rizzo, A.: Virtual humans for assisted health care. In: *Proceedings of the 1st International Conference on Pervasive Technologies Related to Assistive Environments*, vol. 6. ACM (July 2008)
10. Kim, J.M., Hill, J.R.W., Durlach, P.J., Lane, H.C., Forbell, E., Core, M., Hart, J.: BiLAT: A game-based environment for practicing negotiation in a cultural context. *International Journal of Artificial Intelligence in Education* 19(3), 289–308 (2009)
11. Lane, H.C., Schneider, M., Michael, S.W., Albrechtsen, J.S., Meissner, C.A.: Virtual humans with secrets: Learning to detect verbal cues to deception. In: Aleven, V., Kay, J., Mostow, J. (eds.) *ITS 2010, Part II*. LNCS, vol. 6095, pp. 144–154. Springer, Heidelberg (2010)
12. Luciew, D., Mulkern, J., Punako, R.: Finding the Truth: Interview and Interrogation Training Simulations. In: *The Interservice/Industry Training, Simulation & Education Conference (IITSEC)*, vol. 2011(1). National Training Systems Association (January 2011)
13. Rickel, J., Marsella, S., Gratch, J., Hill, R., Traum, D., Swartout, W.: Toward a new generation of virtual humans for interactive experiences. *IEEE Intelligent Systems* 32–38 (2002)

14. Robertson, J., de Quincey, A., Stapleford, T., Wiggins, G.: Real-time music generation for a virtual environment. In: Proceedings of ECAI 1998 Workshop on AI/Alife and Entertainment (August 1998)
15. Swartout, W.: Virtual Humans. In: Twenty-First National Conference on Artificial Intelligence (AAAI-06) (Senior Paper), Boston, MA (2006)
16. Swartout, W., Gratch, J., Hill, R., Hovy, E., Marsella, S., Rickel, J., Traum, D.: Toward Virtual Humans. *AI Magazine* 27(1) (2006)
17. Swartout, W.: Lessons Learned from Virtual Humans. *AI Magazine* 31(1) (2010)
18. Woolf, B.P.: Building intelligent interactive tutors: Student-centered strategies for revolutionizing e-learning. Morgan Kaufmann (2010)