Modeling Believable Virtual Humans for Interpersonal Communication

Edward M. Sims, Ph.D.
Vcom3D, Inc.
Orlando, FL
EdS@vcom3d.com

Glenn Taylor
Soar Technology
Ann Arbor, MI
glenn@soartech.com

ABSTRACT

Human Social, Cultural, and Behavioral Modeling are gaining significant importance in current and future Warfighter training and operational requirements. Peacekeeping and humanitarian missions require the ability to identify civilians' needs and intentions, and to successfully influence or direct their actions. In order to perform successful interviews, rapport building, and negotiations, it is important to recognize non-verbal cues of gesture, facial expression, and body language. However, when interacting with persons of different cultures, we often see behaviors that we think we understand, or use behaviors that we think are understood, but which really have a very different meaning in the other culture. Furthermore, the outward expression of emotion may be suppressed in accordance with varying cultural norms or to avoid the retaliation or censure of others in the community. Ignorance or misunderstanding of these non-verbal cues can make the difference between successful and failed missions. As part of an OSD-Army Research Institute Small Business Innovation Research (SBIR) project, the authors have integrated Virtual Humans with a Cultural Cognitive Architecture to model the culturally-influenced physical behaviors that a Warfighter might experience in face-to-face interactions with persons from non-Western cultures. Using videos of actual discussions and interviews, ethnographic research, and research-based models of culturally-influenced behavior, we developed a cognitive model of an Arab sheikh and the corresponding outward behavior he would likely take during interaction with a Soldier attempting to establish a rapport, obtain information, and influence the sheikh's actions. In this paper, we review methodologies that were tested and used and discuss the methods developed to model the visual cues that were found to be essential to achieve believable and accurate results.

ABOUT THE AUTHORS

Ed Sims is a co-founder of Vcom3D, an Orlando-based company that develops interactive visualizations for Web-based learning. As Chief Technology Officer for the Company, he directs a team of software developers, linguists, and educational technologists, in the development of virtual human applications for education, training, and accessibility. Prior to co-founding Vcom3D, he held positions as Chief Scientist and Technical Director at Lockheed Martin Information Systems Company. Dr. Sims holds a B.S. in Mathematics from the College of William and Mary and the M.S. and Ph.D. degrees in Systems Engineering from Rensselaer Polytechnic Institute. He has been awarded five patents in the area of real-time visual simulation. He is a member of the Association for Computing Machinery (ACM), the American Council on the Teaching of Foreign Languages (ACTFL), the International Society for Technology in Education (ISTE), and the Phi Beta Kappa honorary scholastic society.

Glenn Taylor is a Senior Scientist at Soar Technology, and has been researching and developing human, social-cultural behavior models (HSCBM) for the last 10 years. He has been principle investigator for a variety of DoD-sponsored R&D projects, and oversees all of Soar Technology's work in the HSCBM area. He has a BS and MS from the University of Michigan in computer science and artificial intelligence.

Modeling Believable Virtual Humans for Interpersonal Communication

Edward M. Sims, Ph.D.
Vcom3D, Inc.
Orlando, FL
EdS@vcom3d.com

Glenn Taylor
Soar Technology, Inc.
Ann Arbor, MI
glenn@soartech.com

INTRODUCTION

In the conduct of Sustainment, Stability, Transition, and Reconstruction (SSTR) operations, the success of a mission frequently depends on the ability to communicate with the local population. However, even when an interpreter is available, misunderstandings can arise from lack of knowledge of local customs or from the misinterpretation of non-verbal behaviors. Researchers such as Mehrabian (1971) have reported that as much as 93 percent of the impact of face-to-face interaction comes from non-verbal cues. When communicating with persons of our own culture, we use intonation, prosody, eye-gaze, facial expression, and other non-verbal factors to help identify another's emotions, level of attention, and underlying intent. We also recognize and practice standard courtesies and social protocols. However, when we deal with persons of other cultures, the rules that we have learned by experience may not apply.

Despite the importance of learning how to communicate with persons of different cultures, there are very few opportunities for troops, support contractors, and others being deployed overseas to practice such interactions in a safe learning environment. The use of Virtual Human actors (Sims, 2005; Sims and Pike, 2004) can potentially fill this need. However, few of the game engines and virtual worlds that are being used for experiential training provide the necessary visual cues for face-to-face interaction. Even when detailed facial expressions and body motions can be displayed, there are few cognitive models that provide sufficient information to drive these observable behaviors.

Since 2006, Vcom3D and Soar Technology have been collaborating under a Small Business Innovation Research (SBIR) contract funded by the Office of the Secretary of Defense and managed by the Army Research Institute (ARI) to develop integrated, reusable cognitive and observable behavior models that reflect cultural differences.

NEEDS IN CROSS-CULTURAL TRAINING

Face-to-face interactions are most effective when we successfully infer the other person's inner, private thoughts, intentions, and emotions from their outer, observable manifestations. As with spoken language, social protocols and nonverbal behaviors are not readily acquired through lecture and demonstration; they are best mastered by practicing receptive and productive skills in natural contexts. Our goal is to create interactive, believable cultural avatars as the centerpiece of immersive training environments. To do this, we must include the unobservable thoughts and emotions of a computer generated character and depict their effects on a character's actions and demeanor.

Social Protocols

All cultures have protocols that provide guidance regarding a wide range of social interactions, including such norms as:

- When is it appropriate to accept or decline hospitality or gifts?
- How much rapport building is appropriate before getting down to business?
- What are appropriate discussion topics for getting to know someone? What topics are private or off limits?

Greeting someone of high status is typically different from greeting strangers on the street. For example, greeting an important community leader with "How's it going?" might be deemed inappropriate and immediately put off the addressee. Business meetings are often conducted differently than informal gatherings among friends and family. For example, business meetings in Iraq typically begin with tea and conversation as a way to develop rapport, and only later turn to key business exchanges. To hurried Americans, tea and conversation might seem like an unwelcome distraction.

These kinds of interactions, and the protocols involved in them, tend to be culture-specific. Those people within the culture learn how to behave appropriately in these situations through observation and practice while interacting with others in that culture. To be effective as an outsider, learning these protocols and how to act within them is critical.

Non-verbal communication

It is also important as a trainee to learn to recognize the non-verbal aspects of communication that can also be culture-specific. Although there is strong evidence that the interpretation of facial expressions of emotion is universal throughout all cultures (Ekman, 1971), most gestures are learned from one's culture. Furthermore, even though the meaning of facial expressions is universal, the "display rules" for when it is appropriate to show or suppress these emotions are quite variable (Matsumoto et al., 1998) Learning when to use particular gestures or when to exhibit emotion, and learning to recognize these aspects of communication, are key elements of mastering effective cross-cultural communication.

Non-verbal cues used in the Arab world are often quite different from those used in America or in Western Europe. In the 1970s Barakat (1973) documented 247 gestures used in the Arab world. We found that many of these are understood by persons from a widespread area of the Middle East today; but few are understood by Americans. In order to build a more current database of nonverbal signals used in conversation today, we identified over 200 gestures from actual interviews with citizens in Baghdad. These included not only "emblem" gestures (gestures that have a verbal translation that is understood by persons within a cultural group), but also gestures used for regulating dialog, such as the turn-taking gestures (from Antoon et al., 2005) shown in **Figure 1**. The gesture by the man on the right, which indicates his desire to interrupt, is so powerful that the man on the left feels the need to forcibly push the first man's hand down to stop the interruption.

Figure 1: Gestures in conversational Iraqi Arabic.

Other non-verbal language is used in conjunction with the spoken word and can help to remove intended or unintended ambiguity. An example from working with Soldiers involved in operations in Iraq is the use of "insha'allah" in Arabic when speaking of one's own future acts. Literally this means "God willing" and is used whenever speaking of events that might happen in the future. For example, "I will provide the supplies you need tomorrow at 9 AM, God willing." The actual meaning may range from "if you're really lucky" to "with God's help, we will make it happen". The phrase may or may not be intentionally ambiguous. In either case, clues to the real intent of the speaker must be found in other elements of the communication, such as the tone of the utterance, in facial expression, or in the accompanying gestures. The same skills are often required of non-Americans when an American says something like "Yeah, right!", a phrase that can be changed from an emphatic "yes" to a sarcastic "no" with a little smirk and pitch modulation.

Eye motion provides another important cue in face-to-face interaction. The most obvious significance of eye motion is an indication of a person's focus of attention. During communication between two individuals, each party tends to look at the other's face to obtain visual cues that supplement speech. However, even in an intent conversation, eye gaze does not remain fixed on the other party. The eyes frequently dart away and return. Many simulations do not provide any eye motion. Those that do are likely to use a gaze model in which conversing characters occasionally scan the environment, using a saccadic eye motion that jumps to various background objects, and then returns to the character with whom they were conversing. In our experience, this gives an appearance that they were disinterested in the conversation – even rude. On the other hand, uninterrupted staring at the speaking party gives an equally unnatural appearance. The listener appears to be "wooden", even when they exhibit facial expression and gesture.

As described in (Sims, 2005), we vary eye motion based on cognitive activity, based on research by Lewis (1998), Lewis and Pucelik (1990), and others. During a process of recall or constructive thought, we reduce visual stimulation by looking away from sources of activity. As cognitive load increases, the eyes become less observant of external activity. As reported by Lewis (1998), research has shown that the eyes move in different ways when a person is accessing remembered or internally constructed images, sounds, dialog, or emotions.

Whereas gestures and facial displays provide streams of information, attitude refers to a more persistent cognitive/emotive state that is observable whether the subject is in an expressive or a receptive state. In the context of communication, visual indicators of general attitude are of most interest for the "back channel" that they provide to someone who is speaking. A person's level of attentiveness, trust, and agreeability are reflected by their posture and level of “fidgeting”, as well as by their expression and eye movements.

CULTURAL COGNITIVE ARCHITECTURE

In order to exhibit believable, accurate, and culturally differentiated behavior, it is necessary to have a cognitive model that not only models mental processes but also has the granularity to connect these to specific, observable behaviors. To this end, Soar Technology has been developing a framework for building rich cultural models for driving interactive characters. This framework, called the Cultural Cognitive Architecture (CCA; Taylor et al., 2007), is based on theories of human cognition (Newell, 1990; Schank and Abelson, 1977) and culture, specifically Cultural Schema Theory (D’Andrade and Strauss, 1992; Shore, 1996).

Cultural Schema Theory itself builds on theories of cognition, and takes a knowledge-based approach to culture in which differences in culture can be explained in terms of the knowledge learned as being part of a culture. This knowledge is encoded as *schemas* that represent the relationships between important concepts in the environment, the expectations about how situations might play out, and the goals different characters in the situation might have. Schemas provide both representations of these concepts and a process for connecting elements in the environment to their meaning. This process allows cultural situations to be recognized and understood. It also allows for appropriate goals to be generated within the context of the situation.

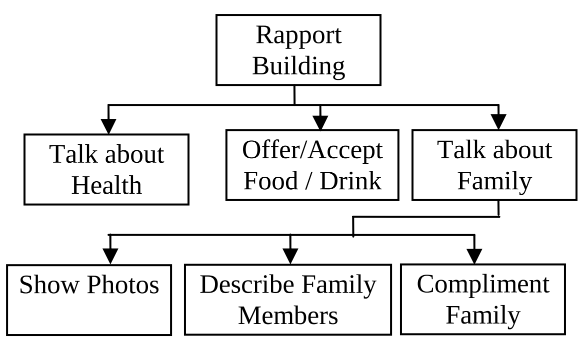

Figure 2: An example cultural schema in CCA

An example interaction schema for building rapport is shown in Figure 2. The schema is hierarchical – there are multiple ways to accomplish rapport building (e.g., talk about family or health). The schema also contains a sub-protocol for talking about family. At this level of description, this schema could probably be applied to many cultures around the globe. However, finer detail in each of the leaves shown here might have culture-specific rules. For example, among two men talking in Arab cultures, it may be improper to speak or ask specifically about the women in either family. Likewise, an American showing a picture of his family at the beach might be seen as breaking rules of social propriety. These cultural variations can be represented as further sub-schema.

CCA also incorporates theories of human emotion, specifically based on appraisal theories of emotion (Scherer, 1997). In CCA, perceived events are appraised along a number of dimensions such as how well the event fits with the avatar’s goals (*goal conduciveness*), how well the event fits within established norms *(internal and external standards)*, and how surprising or novel the event is. We implement a subset of Scherer’s appraisal and emotion dimensions. The outputs of the appraisal-based emotion system include an immediate emotional appraisal based on each perceived event, and an updated running average emotional state (updated with every event) that allows for a more coherent basis for generating behavior. It is this appraisal process that accounts for the interpretation of actions as rude or improper. For example, if the next step in the “rapport building schema” above is for the trainee to “compliment family” and the trainee does not do this, the act of breaking the schema generates a negative emotional state in the avatar.

Building a cultural model consists of encoding schema drawn from the culture in question, in particular the sequences of behavior that constitute a social protocol, as well as the culture-specific appraisal of perceived events or situations. For the purposes of cultural training, a cultural character includes knowledge of the culturally correct interaction patterns relevant to a particular training situation. A trainee interacts with this cultural character by choosing from a set of actions (for example, things to say or do in the situation – asking about family, removing hat and glasses). The cultural avatar processes these trainee actions against its own expectations about the interaction, represented as schema and appraisals. In this process, the cultural character generates an emotional response to the trainee’s actions. The avatar’s response is selected based on a combination of what the current situation demands and the computed emotional state of the agent as a result of observing the trainee’s actions.

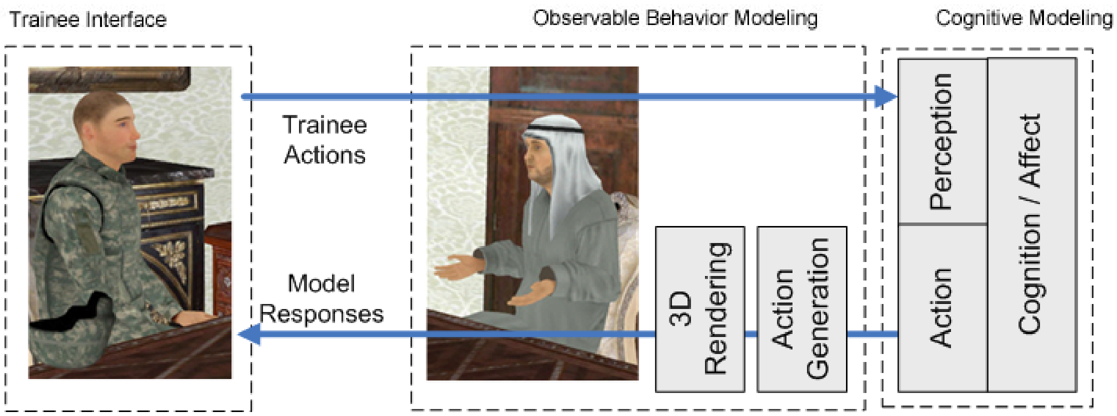

Figure 3. Integration of Physical and Cognitive Model into a Training Environment

SYSTEM ARCHITECTURE

Figure 3 shows our integration of the cognitive and observable behavior models. For our implementation, the cognitive system is in charge of high-level goal-setting and high-level behavior generation. As described earlier, cognitive behavior generation occurs through the matching of situation-dependent schema and the generation of emotional state, which itself may trigger schema. Where a schema indicates the 3D cultural avatar should take an action, the cognitive model exports to the physical model the selected action as well as the immediate appraisal and the emotional state of the character at that point in time. Actions can be selected based on responses to external events, or based on emotional states of the agent, or both.

The physical model receives the action to perform and the emotional state, and uses the emotional content to vary some of the parameters in the generated physical action, such as the emphasis that is placed on some gestures. The culture-specific gestures are pre-scripted, but some of their parameters can be modified on the fly during the scenario. For example, a gesture to gain control of the conversation with a low anger rating would be somewhat muted, whereas with a high anger rating would be much more punctuated physically. The emotional state might also influence ambient behaviors, such as facial expressions (e.g., surprise by raising eyebrows), eye gaze and blinking rates, or shifting weight between feet. All of these contribute to the overall physical manifestation of the emotional state of the agent. It is through this physical appearance of the emotion that the trainee can learn to pick up on non-verbal cues.

CREATING A CULTURAL KNOWLEDGE BASE

Both the physical and cognitive models shown in Figure 3 require extensive, culture-specific knowledge bases in order to produce meaningful results. Different types of information are required in order to model the cultural schema, appraisals, and nonverbal behaviors required for the cultural avatars. However, readily available data on different cultures often consists of such data as:

- Characterization of the culture along a number of cultural dimensions such as individuality vs. collectivism, power distance, and uncertainty avoidance (e.g., Hofstede, 2001).
- Lists of cultural “do’s and taboos”.
- Lists of common gestures and their meanings.

Seldom do these sources of information by themselves provide enough data to build cultural avatars.

In order to build satisfactory models of nonverbal behaviors, we acquired over 13 hours of videos of interviews and group discussions with Iraqis; with most of these being conducted in Iraq. As described in (Sims, 2005), we transcribed and translated the dialog (if this had not already been done) and then annotated gestures and other nonverbal behaviors of interest using Max Planck Institute’s Eudico Linguistic Annotator (ELAN; Hellwig, 2006), shown at the top center of Figure 4. These annotations include the formation, apparent meaning, and context of each identified behavior. By searching, for example, for all instances of dialog turn-taking, we were able to identify standard protocols and gestures associated with this behavior. Likewise, we were able to identify facial expressions and eye movements associated with reconstructing incidents from memory. In all, we were able to identify about 200 nonverbal behaviors used in face-to-face interactions in Iraqi culture.

We developed a second tool, shown in the lower right corner of Figure 4 for rapidly modeling the kinematic formation and visual appearance of gestures. This tool provides an interface whereby the user creates gestures

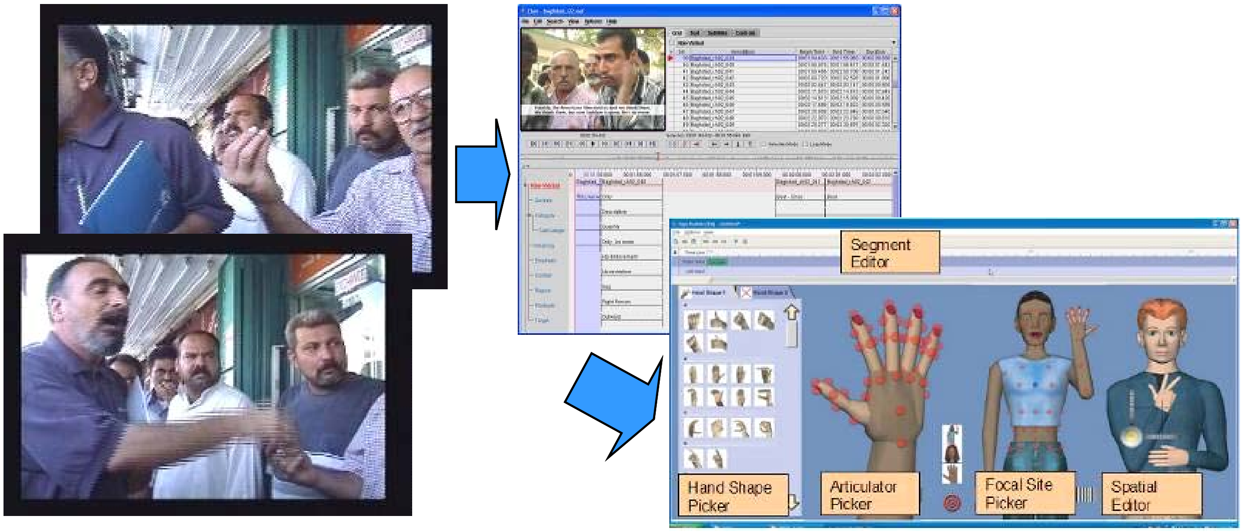

Figure 4. ELAN and Gesture Builder tools for identifying, annotating and modeling non-verbal behavior.

as a sequence of segments, where each segment is formed by a handshape, an active articulator (pointer), and a focal site. This concept for representing gestures is based on the work in sign language phonetics developed by S. Liddell and R. Johnson (1995). To the best of our knowledge, this system for *analyzing* signs and gestures has never before been used to *synthesize* gestures.

The Gesture Builder tool is shown in Figure 4. To create a gesture, the author adds a (right or left hand) segment along the time-line of the Segment Editor. Each segment consists of a movement and a hold period, each of which may be set to a desired duration. (If the hold period is set to zero, the arm/hand motion simply passes through that pose without stopping.) For each segment, the author selects a hand shape from a list of over 100 shapes, using the Hand Shape Picker. Next, the author selects an “active articulator”: the site on the hand that will touch or point to another point on the body, or a point in space. Finally, the author selects a “focal site”: a point on the body that will be pointed to or touched. Alternatively, the Spatial Editor may be used to select a point in space, an offset from a focal site. These basic options are supplemented by additional tools that allow the selection of oscillating hand shapes and contact types (brush or tap, as well as touch). Segments may also be copied and pasted.

When a gesture is completed, it is exported as an Extensible Markup Language (XML) Gesture File that is stored in a database that can be accessed by the Vcommunicator scenario authoring tool. As compared to general-purpose character animation tools, the Gesture Builder proved to be highly efficient. Whereas using tools such as 3D Studio Max and Character Studio, we were only able to animate and characterize about one gesture per hour, with the Gesture Builder, we were able to animate six to ten gestures per hour. Graphic designers were also able to learn to use the tool much more rapidly. Much more importantly, since each gesture was defined by such descriptive features as “active articulator” and “focal site”, and was synthesized at run-time, it was possible to develop gestures that would work with any character, independent of that characters dimensions or range of motion. Finally, the gestures could be spatially inflected in real-time, by identifying focal sites as variables, whose value is supplied dynamically at run-time. Additional nonverbal behaviors were identified from such references as Barakat (1973), Kavanaugh (2000), and Morris et al. (1980). These behaviors were tested with recent immigrants from Iraq to verify that their formation, usage, and meaning were still current.

We have also been developing a cultural schema editor that allows a modeler to encode the schema that will drive the behavior of the avatar. A schema defined in the editor connects the lowest level behaviors observable in the environment (e.g., “offer tea”) with higher-order organizational schema (“rapport building”). The lowest level schema are also connected to physical behaviors developed in the Gesture Builder. A cultural cognitive model is potentially composed a multitude of disconnected schema covering a range of situations and protocols relevant to the culture. The schema editor also serves as a framework for constructing the training scenarios, allowing the instructor to construct meaningful sequences of activity for a particular training exercise.

Figure 5. Student-controlled American Captain, building rapport with Sheikh Ghazi and Samir.

the need to act promptly in order to keep the interaction from diverting from their own agenda.

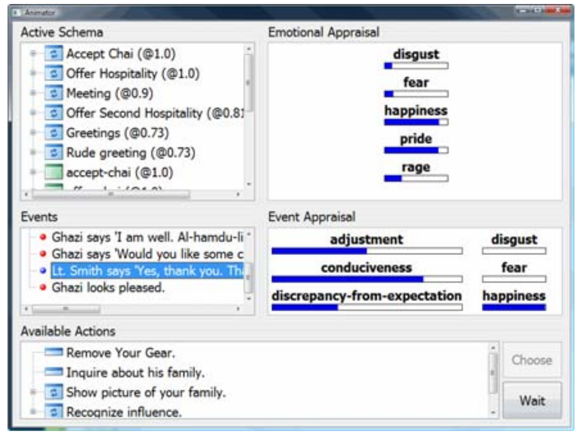

Figure 6. Interface for Soar Cultural Cognitive Agent.

EXAMPLE TRAINING SCENARIO

In order to demonstrate and evaluate the integrated system, we developed a Rapport Building Scenario in which a student takes on the role of a Captain who is meeting an Iraqi sheikh for the first time after taking over operations in the area. The sheikh's actions are controlled by a Cultural Cognitive Agent, as are those of his confidant, Samir. The student's goal is to get to know the sheikh and win his cooperation in reducing the number of IED assaults on the Captain's men. The sheikh's goals include finding employment for his constituents and maintaining their respect and allegiance.

Figure 5 illustrates the interactive 3D visualization of the scenario, while Figure 6 is a snapshot of the cognitive model that controls the sheikh's actions. At any time during the scenario, several actions are available to the Captain, depending on the active schema. These may be gestures, utterances, physical actions such as removing gear, or combinations of multiple types of actions, as shown in the bottom window of Figure 6. As the student selects actions, the sheikh forms an appraisal and responds with a next course of action. The sheikh also initiates actions based on his own goals, as well as responding to the Captain's action. Providing this type of "mixed initiative" dialog was found to be a key in achieving realism and in "suspending disbelief". In a version without sheikh-initiated actions, users commented that they lost the sense of needing to make prompt decisions about what to do next. On the other hand, when the sheikh acted on his own initiative, they felt

As the interaction progresses, the overall emotional appraisal formed by the sheikh changes in accordance with his cumulative reactions to the Captain's sequence of actions, as shown in the upper right window of Figure 6. In an actual training scenario, the student does not see the schema, event appraisals, and emotional states as shown in Figure 6. Instead he/she must infer the sheikh's mental state based on observable behaviors, as shown in Figure 5.

The scenario progresses through three stages, as shown in Figure 7: greetings, rapport building, and negotiation. In accordance with his cultural norms, the sheikh provides minimal direct feedback, but tries to use indirect language to guide the Captain through appropriate rapport building to a successful conclusion of the negotiations. For example, if the Captain fails to remove his helmet and sunglasses, the sheikh eventually remarks "Is it too bright in here?" The sheikh also lets the Captain know about the availability of men to work at the Army base without specifically requesting that the Captain find them jobs. At each step of the interaction, the Soldier must use cues from the sheikh's intonation and nonverbal behaviors to help determine how the discussion is going, what the real intent behind indirect utterances may be, and how he may need to adjust his approach. At the end of the meeting, when the Captain asks for assistance, the sheikh responds "insha'allah", but with very different gestures and body language in accordance with the level of skill the Captain has shown in building a rapport, leaving it up to the student to judge how well he thinks the meeting went.

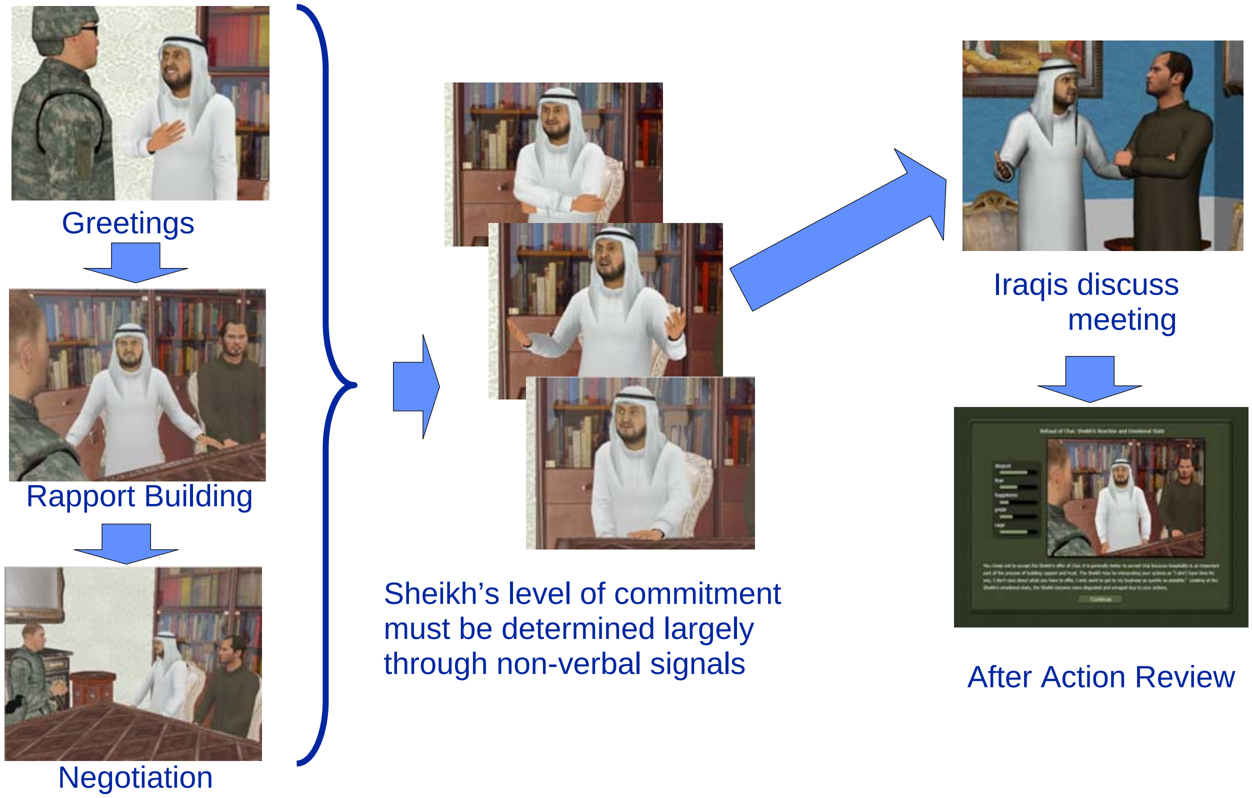

Figure 7. Rapport Building Scenario

At the end of the meeting, in order for the student to see the results of his decisions during the scenario, he is shown a brief follow-on meeting between the sheikh and his cohort Samir, where they discuss their impressions of the Captain and their own planned course of action. Depending on whether the student has earned their trust, they may decide to work with the Captain, to wait and see, to request another meeting, or to work against him.

EVALUATION AND FUTURE PLANS

Design decisions that were made in the course of this Phase II project are based in large part on a pilot evaluation that was conducted with Soldiers of an earlier "Knock and Talk" scenario that used the same level of physical behavior simulation. In an evaluation with U.S. Army Soldiers, users scored the usefulness of the nonverbal behaviors 4.8 on average on a scale from 0 to 5; and scored the realism of the characters and their behaviors 4.5 on the same scale. Pre-test and post-test evaluation showed positive training transfer. A series of evaluations of the current Rapport Building Scenario is planned beginning in August 2009.

While the software tools and behavior libraries we have described in this paper serve to reduce cost and increase reusability of cultural avatars and behavior models, we are taking additional steps to increase reusability across a wide range of gaming, Web, virtual world, and mobile platforms. Originally developed to run in a Web plug-in, we have used Humanoid Animation (H-Anim, 2004) and Collaborative Design Activity (COLLADA) standards to provide exports of the avatars and behaviors to Forterra's Online Interactive Environment (OLIVE), Emergent's Gamebryo, Unity Technologies' Unity game engine, and Vcom3D's xPosition engine for Apple's iPhone and iPod Touch. Our overall goal is to provide "plug and play cultural avatars" that can be developed once and imported into any of the many training environments being used by the DoD.

ACKNOWLEDGEMENTS

This work was performed under US Government contract W91WAW-08-C-0052, monitored by the Army Research Institute. Opinions expressed in this paper are those of the authors and do not necessarily represent the views of these sponsoring organizations. On-going development of a Plug and Play version of the cultural avatars is being performed under US Government contract W91CRB-09-C-0004, administered by the Combating Terrorism Technical Support Office.

REFERENCES

Antoon, S., S. Shapiro, S. Salamy, B. Haddad, and M. Mikdashi, *About Baghdad*. 2005, AFD Studio. (DVD: 90 minutes).

Barakat, R., *Arabic Gestures.* Journal of Popular Culture, 1973: p. 750-794.

D'Andrade, R. and C. Strauss, eds. (1992). *Human Motives and Cultural Models*. 1992, Cambridge University Press: Cambridge, UK.

Ekman, P. (1971). *Universals and cultural differences in facial expressions of emotions*. in *Nebraska symposium on motivation*. 1971. Lincoln, NB: University of Nebraska Press.

Gibson, J.J. (1979). *The Ecological Approach to Visual Perception,* Houghton Mifflin, Boston.

Gratch, J., Rickel, J., André, E., Cassell, J., Petajan, E., and Badler, N. (2002). Creating Interactive Virtual Humans: Some Assembly Required. IEEE Intelligent Systems Vol. 17, No. 4, 54-63.

Hellwig, B. (2006). Eudico Linguistic Annotator (ELAN) version 2.6 User's Manual. Retrieved from www.mpi.nl/tools, June 14, 2009.

Hofstede, Geert (2001). *Culture's Consequences: comparing values, behaviors, institutions, and organizations across nations* (2nd ed.). Thousand Oaks, CA: SAGE Publications.

Humanoid Animation Working Group (2004). *ISO/IEC FCD 19774:200x, Humanoid animation (H-Anim).* Retrieved June 24, 2004, from http://www.h-anim.org.

Itti, L., Dhavale, N., and Pighin, F. (2003). Realistic Avatar Eye and Head Animation Using a Neurobiological Model of Visual Attention. Institute for Creative Technologies report. Retrieved June 24, 2005, from http://www.ict.usc.edu.

Kavanaugh, J. (2000). *Worldwide gestures: A traveler's introduction to common international gestures and symbols.* Blaine, WA: Waterford Press.

Lewis, B., and Pucelik, F. (1990). Magic of NLP Demystified, Metamorphosis Press, Oregon.

Lewis, H. (1998) *Body Language: A Guide for Professionals*. Response Books, New Delhi.

S. Liddell and R.E. Johnson, American Sign Language: The Phonological Base in C. Valli and C. Lucas, *Linguistics of American Sign Language*, Gallaudet University Press, Washington, D.C., 1995.

Matsumoto, D., S. Takeuchi, S. Andayani, N. Kouznetsova, and D. Krupp, *The contribution of individualism-collectivism to cross-national differences in display rules.* Asian Journal of Social Psychology, 1998. **1**: p. 147-165.

Mehrabian, A. (1971). Silent Messages. Wadsworth Publishing, Belmont, CA.

Morris, D., Collett, P., Marsh, P., & O'Shaughnessy, M. 1980 *Gestures, their Origins and Distribution.*

Newell, A., *Unified Theories of Cognition*. 1990, Cambridge, MA: Harvard University Press.

Rehfeld, S., Alderson, R. M., Jentsch, F. G., Fiore, S. M., Bowers, C. A., and Salas, E. (2003). *Research for an Augmented Learning System for the Interpretation of Non-Verbal Cues,* Team Performance Laboratory, Department of Psychology, University of Central Florida.

Schank, R. and Abelson, R. (1977). *Scripts, Plans, Goals, and Understanding: An enquiry into human knowledge structures*. 1977, Hillsdale, NJ: Earlbaum Assoc.

Scherer, K.R. (1997). *The Role of Culture in Emotion-Antecedent Appraisal.* Journal of Personality and Social Psychology, **73**(5): p. 902-922.

Sims, E., and D. Silverglate (2002). Interactive 3D Characters for Web-based Learning and Accessibility, ACM SIGGRPAPH, San Antonio, 2002.

Sims, E., and Pike, W. (2004). Reusable, Lifelike Virtual Humans for Mentoring and Role-Playing. *Proceedings of IITSEC 2004.*

Sims, E. (2005). *Simulating Believable, Context-aware and Culture-specific Human Behaviors*. in *2005 Interservice/Industry Training Simulation & Education Conference*. 2005. Orlando, FL.

Sims, E. (2007). *Tools for Creating Objects and Behaviors for Distributed Simulations with a Cultural Context.* Submitted to U.S. Army Research Institute for the Behavioral and Social Sciences, Orlando Unit.

Taylor, G., M. Quist, S. Furtwangler, and K. Knudsen. *Toward a Hybrid Cultural Cognitive Architecture*. in *CogSci Workshop on Culture and Cognition*. 2007. Nashville, TN: Cognitive Science Society.