The First Person Cultural Trainer

Marjorie A. Zielke, Ph.D.
The University of Texas at Dallas
Richardson, Texas
margez@utdallas.edu

Thomas E. Linehan, Ph.D.
The University of Texas at Dallas
Richardson, Texas
Thomas.linehan@utdallas.edu

ABSTRACT

This paper will describe the First Person Cultural Trainer (FPCT), sponsored by TRADOC G2 Intelligence Support Activity. FPCT is a 3D interactive simulation that trains soldiers on the values and norms of a specific culture in order to facilitate missions. The environment acts in a nonlinear way, as a Middle-Eastern geographic area would. The interactive simulation utilizes unique technology that gives soldiers the ability to read non-verbal communications of the non-player characters (NPCs) in the game. The project is currently focused on Iraq and Afghanistan, but has applications in many other cultural and geographic situations. FPCT includes four clearly defined stages which lead to the successful completion of missions. The project has adopted a challenge to develop extremely high-fidelity representations, using the living-world construct, to create an environment that serves as a training tool for cultural training before or during actual deployment.

ABOUT THE AUTHORS

Marjorie A. Zielke, Ph.D., is an assistant professor of Arts and Technology at the University of Texas at Dallas and is the principal investigator on First Person Cultural Trainer. Dr. Zielke has served as the project manager on a series of culturally related military-funded simulation projects over the last several years. Her areas of research are cyberpsychology and hyper-realistic simulations and she works in the intelligence, health and marketing sectors. Dr. Zielke received her Ph.D. from the University of Texas at Dallas, and also has an MBA and a masters in international business.

Thomas E. Linehan, Ph.D., is professor of Arts and Technology and director of the Institute for Interactive Arts and Engineering (IIAE) at the University of Texas at Dallas. Dr. Linehan is the co-principal investigator on First Person Cultural Trainer and began the cultural simulation research program at IIAE. Dr. Linehan received his Ph.D. from The Ohio State University.
The First Person Cultural Trainer (FPCT) is a 3D interactive training simulation, sponsored by TRADOC G2 Intelligence Support Activity, which teaches the values and norms of a specific culture in order to facilitate intelligence missions. The game is currently focused on Iraq and Afghanistan and missions related to uncovering improvised explosive device (IED) network information. FPCT includes four clearly defined stages that lead to the successful completion of missions:

- Step One: Establish a presence in the community
- Step Two: Understand community problems
- Step Three: Develop an action plan
- Step Four: Gather intelligence like information about IED networks.

The project has adopted an ongoing challenge to develop extremely high-fidelity representations to create an environment that serves as a cultural training tool before, or during, actual deployment. The game must act in a nonlinear way, as a Middle-Eastern geographic area would. Further, players must have the ability to read nonverbal communications of the non-player characters (NPCs) in the game.

Our game inserts the player into rural, semirural, and urban settings within Iraq and Afghanistan. The player’s goals are to explore the social and political settings through conversation, establish himself within the area, provide humanitarian aid, and eventually use the social status he has established to gather information about IEDs and the IED network.

In a high-fidelity simulated environment like FPCT offers, players can attempt to achieve realistic objectives in the context of accurate cultural representations. The technology is much more flexible than staged, full-scale model villages with live endemic actors — one way this type of training has been done to date. A high-fidelity simulation such as FPCT allows any number of players, who may be scattered around the globe, opportunities to quickly experience a wide variety of situations and specific cultures.

Conversation with the NPCs, otherwise known as virtual agents, is the primary method of interacting with the world. Tactfulness, emotional and cultural sensitivity, and problem resolution will bring respect among area inhabitants. Agents may become more comfortable with the player over time, and become increasingly willing to talk about issues in the area. The player is expected to rate moods and reliability, and gather information through each of the four stages described above. At the end of gameplay, the player is assessed based upon the accuracy of his ratings of the populace and the quality of information gathered.

To achieve these design requirements, the FPCT development team has developed a living-world construct to replicate in-theatre cognitive complexity. Param-
eters on which FPCT simulates a face-to-face experience include modeling psychology, culture, behavior and emotion; developing virtual humans and populations; modeling highly accurate environments; dynamically generating gameplay; and assessing player performance. This paper discusses the design strategies utilized to develop FPCT with the living-world construct.

**THE CULTURAL MODELING SPACE**

Due to the challenges of the Middle-Eastern theatres that have dominated twenty-first-century U.S. military priorities, there are several games in the same cultural awareness training domain as FPCT. Among them are ELECT BiLAT, Virtual Cultural Awareness Trainer (VCAT), Vcommunicator Mobile Translator, and the Tactical Language & Culture Training System (TLCTS).

The ELECT BiLAT game is primarily a bilateral negotiation trainer. The player researches key agents, building appropriate relationships with them through a series of set-piece encounters to successfully conclude the negotiations. Cultural awareness is a skill the player must refine in order to be an effective negotiator (USC Institute for Creative Technologies, 2009). VCAT is a story-driven cultural awareness trainer that provides the player with various intercultural scenarios to play through (Team Orlando, 2009). Vcommunicator Mobile is a tactical language trainer based on the Apple iPod platform and briefs the user on culturally-appropriate phrases and gestures (Vcom3D Inc., 2007). TLCTS is somewhat of a synthesis of VCAT and Communicator Mobile and teaches cultural awareness and tactical language skills simultaneously (Alelo Inc., 2008).

All of these programs have merit in the contribution they make to the cultural training tools available. The research emphasis of FPCT is on the living-world construct and corresponding architecture. Another main emphasis of FPCT is high-fidelity NPCs that can display culturally specific nonverbal communication. Finally, a major component of the FPCT research is the ability to link visualization to the cultural models, and in effect, drive the entire synthetic experience from data and models.

**MAKING CULTURE**

Our model uses visual, auditory, cultural, and behavioral components for immersive cultural training using the living-world construct. Living worlds offer nonlinear, unscripted processes for experiencing and safely learning the cognitive complexity and nuance of culture through emergent high-fidelity simulation (Zielke, Evans, Dufour et al. 2009).

We will emulate the definition of the cultural framework outlined above to achieve the living-world construct and create a realistic and believable synthetic training experience. Since the exact situation that must be modeled at any given time is unknown, the simulated environment must be adaptable. For this reason, we have created a living-world architecture that models psychology, culture, emotion and behavior. Part of the ongoing research is to create tools that can be used by subject matter experts who understand the cultural nuances of the mission environments.

Development of the living-world architecture and the cultural models is ongoing. For this reason, the project strategy is to create a living-world architecture that is flexible and can accommodate new inputs, and is robust enough to create a representation of specific cultures now. Figure 3 on the next page illustrates this design strategy.
In FPCT, the NPCs that populate the villages and towns the player visits are not static characters, locked into specific behaviors. Rather, they influence and are influenced by the environment around them, which affects how they interact with other NPCs and the player. Each NPC carries within it a complex emotional model which directly influences its behavior. This emotional model is influenced by many factors, among which are previous events (as outlined in the beginning of each mission through prologues), ethnic background and psychological models, gender, age, and how they have been treated in-game. Figure 4 illustrates this concept.

**DESIGN PHASE ONE: DEVELOP CONCEPTUAL PSYCHOLOGICAL, EMOTIONAL, BEHAVIORAL AND CULTURAL MODELS AND REPRESENTATIONS**

To build the components of culture outlined above, we need to establish a basic psychological model that can serve as a foundation. Our model defines four basic psychological motivators: survival instincts, ego/social relationships, reason and logic, and morality. Each motivator has three major states: strongly negative, strongly positive, and weak. At any moment, each motivator lies within some degree of these three states. For example, a positively influenced morality implies joy while a negatively influenced morality implies sorrow. A weak morality motivator has neither effect.

Table 1 below relates the four motivators and emotions associated with them. These emotions were selected as valuable to our gameplay and relevant to the model.

**Table 1. Psychological Motivators with Relevant Emotions**

<table>
<thead>
<tr>
<th>Psychological Motivator</th>
<th>Strong Negative</th>
<th>Strong Positive</th>
<th>Weak</th>
</tr>
</thead>
<tbody>
<tr>
<td>Survival</td>
<td>Fear</td>
<td>Devotion</td>
<td>Security</td>
</tr>
<tr>
<td>Ego</td>
<td>Anger</td>
<td>Satisfaction</td>
<td>Tranquility</td>
</tr>
<tr>
<td>Reason</td>
<td>Interest</td>
<td>Interest</td>
<td>Apathy</td>
</tr>
<tr>
<td>Morality</td>
<td>Sorrow</td>
<td>Joy</td>
<td>Neutrality</td>
</tr>
</tbody>
</table>

The schema is another component of building psychology. A schema defines mental representations of the world. For our purposes, a schema is the experiential memory of our NPCs that can be revised, refined, or falsified and includes all beliefs about the world. Our schema maps game-world objects to schema nodes, where each schema node contains a list of properties that the character assigns to that object. For instance, if an agent perceives a particular coffee shop, his schema contains a list of everything he believes about that coffee shop. Relational schemata could also be developed to build associations between schema nodes, and the complexity and accuracy of the schema could be controlled by character personality or age as referenced in Figure 5. Piaget (1972) suggests that schemata can acquire greater complexity by assimilation and accommodation through cognitive development.

Figure 3. The on-going five-stage design process for FPCT.

Figure 4. The result of the FPCT design is a dynamic NPC personality that can remember conversations, express moods, and can choose to provide information to the player based on these internal variables.
Cultures act as a seed from which NPCs can be procedurally generated. Some components of our cultural model are based on the research of Martha Maznevski. Maznevski's research has identified at least fourteen variables that define clear cultural parameters for belief and behavior (Maznevski, 2002). These dimensions allow for more specific behavioral definitions than are allowed by our emotional model, "filling in the gaps" of cultural nuance. For instance, the Maznevski dimension of “Can people change?” controls the flexibility of a virtual agent’s opinions toward others and how quickly those opinions solidify.

As figure 6 illustrates, cultures are arranged into a subculture hierarchy. The subculture hierarchy allows cultures to inherit properties from their parent cultures.

Perception, Decision Making and Behavior

In the context of our simulation, perception is defined as all inputs that influence how characters interpret themselves and the world around them. Perception is handled by our event model, which receives broadcasts of internal events (thoughts) and external events (interactions) and passes them into the agent’s mind for processing.

The interaction of cultural expectations, an agent’s schema, and an agent’s mental state makes up our personality system. This system is responsible for driving the decisions and behaviors of our characters via a rule-based contextual analysis of each of these factors.

Determining how to respond to perceptions is the primary form of decision making. Responses include: modification and queries to the schema, emotional change, and prioritization of an errand. Culture drives our list of possible errands. This currently includes, for example, eating, farming, playing, cleaning, and conversing.

Hunger is one example errand. As time passes, characters become increasingly hungry. This constantly re-prioritizes the “eat” errand until it becomes the highest-priority errand. The character’s mental state becomes increasingly anxious as biological needs are not being met proportional to their hunger level until the eat errand has been completed. The character may have several strategies for completing this errand, and his anxiety may be physically manifested in many ways. The personality governs exact manifestations.

Prologues

Prologues are pseudo-cultural entities that represent environmental and psychological influences from recent events in an area. Each environment used by our game has a prologue that changes behavior and personality for that area only. These prologue definitions include information for randomizing population and objects in the world, a situational personality that influences each character, and preset intelligence that will be disseminated by agents during conversation.

Figure 7 on the next page illustrates the full conceptual psychological, emotional, behavioral and cultural model.
DESIGN PHASE TWO: TOOLS FOR EXPERTS TO CREATE CULTURE

Our culture design tool allows subject matter experts to build and maintain subculture hierarchy and define properties. Cultural properties include animation data for gestures, Maznevski cultural variables, and the extent of the culture’s influence on a person. Each culture has a collection of roles that define influences and group expectations. Cultural role properties include acceptable wardrobes, baseline personality, and control variables for defining ages, genders and behaviors. Characters are generated from these cultural roles and deviate depending on the cultural influence on that person. The subculture hierarchy allows cultures to “inherit” properties from parent cultures. Subcultures are by default similar to parent cultures, but may be refined as desired. Parent cultures can be modified, automatically updating each of the child cultures. Figure 8 illustrates the tool’s slider interface.

DESIGN PHASE THREE: DEVELOP VISUAL, AUDITORY, AND BEHAVIORAL REPRESENTATIONS FROM EXPERT CULTURAL DESIGN

We use the constructs described above to make NPCs. Each emotion is associated with a pool of facial and body motion capture data. Blended animations between these extremes produce the final, nuanced animation. The animation associated with a character becomes the weighted average of the corresponding animations.

Virtual Humans

Visually, a given NPC manifests its emotions through

Figure 9. The facial motion capture process. Thirty-nine markers are placed on the faces of live actors to capture culturally specific non-verbal communications.
motion-captured postures, gestures and facial expressions which were performed by highly-specialized Afghani and Iraqi actors. The NPC can use that motion capture data to express a wide variety of emotions including mourning, happiness, and anger. This process is illustrated in Figure 9. When combined with body movement, NPCs are rendered realistically and naturally.

**Fashion Design System & Makeover Studio**

The Fashion Design System is a series of tools that allow us to easily generate large numbers of culturally accurate wardrobes. Our model of human wardrobes is essentially a hierarchical boolean logic system. Each wardrobe is represented as a conjunction or disjunction of wardrobes for nuances like “don’t wear this shirt with this hat.” These wardrobes have an artist-defined system for randomization while allowing us to easily generate large numbers of outfits. To construct an outfit, the program starts at the root of a wardrobe tree and descends recursively. If the wardrobe is conjunctive, then all sub-wardrobes are traversed. Figure 10 illustrates this process.

The most important tool in the Fashion Design System is Makeover Studio. This tool allows artists to see and modify their characters and wardrobes with a real-time view of the results. 

**Environments**

Our environmental design process starts with research of a given geographic area. We use tools such as Google Earth and Google Images, as well as reference photography from subject matter experts, to gather a pool of source materials. We then synthesize our references into a fictitious environment, based on geographic parameters provided from scenario prologues. Our production strategy involves initial development of a basic structural layout using design tools within Unreal Engine 3. We then begin enhancing the environment with the addition of meshes modeled in 3D design programs. Details are added with additional textures and advanced particle emitters that simulate effects such as dirt, dust, and smoke. The environments are finalized by adjusting lighting models to simulate realistic lighting and shadow through the use of radiosity and ambient occlusion, features of the recently updated Unreal Engine 3. The high-fidelity quality that this process produces is illustrated in Figure 11.

**Sound**

Sound in FPCT is mainly designed to ensure the level of complexity and realism necessary for successful immersion in the environment. The sound in the environment will allow for the same type of perceptions as sounds in real life and fulfill four levels of listening (Schaeffer, 1966).

The first level is the indexical level of listening. This means each of the sounds in the environment is related to a visible or plausible physical cause. This level answers the question: “What do I hear?”

The second level pertains to the reception of this sound in regard to the overall acoustical context and to the situation of the listener in this environment. This level answers the questions: “Where am I in this environment?” and “Where does this sound come from?”

The third level is defined as the iconic level of listening because it relates to the purely sonic qualities of the sound. This level is uniquely activated when a sound whose cause has not been defined by the first level is perceived. Qualities of the sound such as gait, variations of harmon-
ic content, variations of intensity, and timbre are used to form an image of the cause of the sound. This level is also activated when semantic information from the fourth level is not available. For example, players can hear a voice speaking in a foreign language and understand the emotions and intentions of the speaker by analyzing qualities of the sound such as variations of intensity and pitch.

The fourth level corresponds to the semantic activity of listening and enables the understanding of the content conveyed by the sound. This level answers the question, “What does this mean?” The answer depends on the knowledge of the listener and the code by which the sound relates to a meaning. The syllables of the phrase, “Sabaah al-khayr,” can mean “good morning” to only some listeners.

Using this approach, we defined four streams of sound used to synthesize any complex environment and to immerse the user into a meaningful auditory context:

- **Global stream**: the background sound of the environment
- **Regional stream**: the background sound of a section of this environment
- **Local stream**: the sounds attached to the visible agents
- **Conversation**: the voices of the agents in face-to-face conversation.

These streams are also defined by their degree and mode of variation or their dependence on other parameters of the program. Figure 12 illustrates, these four streams of sound as they relate to the game environment.

### DESIGN PHASE FOUR: GAMEPLAY

In FPCT, the player is presented with a prologue that outlines the situation (i.e., general perception of the player by the populace, recent violence, unrest, etc.), and then is given a list of goals to meet. The goals are separated into the four stages referenced in the introduction, with each stage having its own discrete set. The player decides when he or she has met the goals of the stage and can advance to the next by selecting the option for his journal.

Once a player has signaled that he wants to advance to the next stage, he is graded on completeness and on his success in interacting with the population. This between-stage grade will factor into the final assessment at the end of the game. By completing the assigned goals, communicating well with the populace, and discovering and documenting important information, the player will advance through the game and ultimately receive a grade for that level. The player can then move on to a different environment level, or replay the current environment with a different prologue shaping the events. Important elements of our game design are discussed below.

#### Time

Time is an important consideration in the living world construct. Each of the four key stages of the game takes time to accomplish. Rushing through any of the stages can make the final stage of gathering intelligence impossible. Time is also driven by cultural norms. This is one of the lessons of the game.

For these reasons, three tiers of time representation overlay and interact dynamically within the game. The first is the narrative representation implied by the prologue. The prologue offers initial definitions of a starting point, a hypothetical future depending on the actions of the player, and an overall game duration to accomplish these actions.

The second time tier, nested in the first, is the ecological representation of time. In this tier, time is represented as it is experienced by the user when immersed into one of the four stages of the game. This time tier is strictly continuous and entirely driven by the dimension of the present in which actions and events unfold continuously for the user as well as for the NPCs. The perception of duration and pace of events and
actions is strictly realistic and similar to the perception of these qualities in the actual world.

The last tier is virtual time, which exists to connect the four stages of the game. This time is only represented by the tangible changes occurring from one stage to another. For example, if the player in stage one agrees with the local authorities on the necessity of digging a new well for the village, then, in stage two, this well will be available. Villagers will organize their errands accordingly due to the new resource. They will also form a different opinion of the player. In general, the gameplay takes advantage of the overall representation of time:

- To offer different levels of complexity for the user as defined by the prologue.
- To avoid a shortcoming of scenario-driven games in which the player, basing his behavior on his sole knowledge of the narrative logic, moves through the game collecting intelligence where it is expected to be.
- To possibly assess the performances of the user at two different levels: environmental interaction and strategic decisions. The first level of achievement can be measured at the end of each stage of the game and the second level can be measured according to the overall success of the mission.

**Written and Spoken Dialogue**

To provide a more immersive cultural experience, NPCs will utilize both written and spoken dialogue to communicate with the player. Nonverbal cues can be ascertained by studying an NPC’s posture and expression. NPC emotional states can also be perceived by the player through spoken dialogue, which was recorded with voice actors of the appropriate ethnicity and background. Dialogue lines were written to correspond to predefined character interaction. The NPC’s spoken dialogue is presented in the appropriate language (Pashto, Arabic, etc.), and a written transcript continually updates the player with translations. Figure 13 illustrates how the dialogue progression influences gameplay and the ultimate outcome.

**Reading Attitudes**

The player reads the attitudes and emotions of NPCs though several feedback mechanisms which are tethered to the NPC’s psychological model described earlier. Throughout the player’s interactions with the various NPCs in a village or town, the ability to interpret and understand a given NPC’s attitude will significantly increase the chance of successful communication.

Taking cues from normal, everyday conversation, in-game conversation patterns cannot simply be memorized to progress to the goal. The player will have to take into consideration the NPC’s emotional state and place this character within a cultural context before he can successfully engage in an important dialogue. The player will come to realize that there can be several paths to the same goal, depending on the situation. An upset character may have to be calmed and reassured before any meaningful dialogue can take place, or the player may have to prove trustworthy before an NPC will talk willingly. Catching an NPC on a “bad day” may mean that a player might have better luck moving on and making contact with this NPC in future stages.

The player can leverage cultural knowledge to successfully interact with the NPCs. For instance, attempting to begin a conversation with an unescorted female in rural Afghanistan may sour NPC perceptions of the player, making future interactions in that area more difficult or even impossible. A perceptive play-

---

**Figure 13.** Dialogue is an important gameplay element – influencing both the attitude of NPCs to the player and overall success in accomplishing the mission.
er will take an NPC’s gender, age, and disposition into account when deciding with whom to interact.

**Journal Entries**

Players will be scored on their ability to collect valuable information and record it in their in-game journal. During a conversation with an NPC, the player is able to use his mouse to drag portions of that conversation into his journal, where that information will be stored and associated with the appropriate NPC. The player may then look at this information at any time, if he needs to recall an important piece of a previous conversation.

At the end of a stage, the player is scored on the quality and quantity of the information stored in his journal. This grade is then factored in with other assessment categories to produce a final grade for the stage. Scoring particularly well or poorly may affect NPC dispositions and goals in future stages.

**DESIGN PHASE FIVE: EVALUATION AND ASSESSMENT**

The Constructivist epistemology proposes that learners are active participants in the learning process, who seek meaning by reflecting on their experiences. (Driscoll, 2005). Experiential learning simulations such as FPCT are clearly supported by these principles. Motivation is one significant factor in a learning environment’s success. The compelling nature of simulations encourages players to continue participating, thereby increasing the time they spend within the learning environment (Medina, 2005). Further, the cognitive complexity and stage objectives of FPCT present the player a difficult, but not unachievable, challenge. Balance between challenge and boredom is an important aspect of motivation in both learning environments and games (Paras and Bizzocchi, 2005; Koster, 2004).

**Evaluation**

We will undertake several forms of evaluation to ensure the continued accuracy of FPCT as a learning environment for cultural awareness training. Formative evaluation is ongoing, as subject matter experts review the simulation’s content and design. A summative evaluation will be performed when the design process is complete, in which the learning outcomes achieved by participants in the FPCT simulation will be analyzed and compared to those who engage in traditional training methods. This evaluation will determine the success of the simulation as an in-

<table>
<thead>
<tr>
<th>Table 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>After-Action Review Tool Design</td>
</tr>
<tr>
<td>Criteria</td>
</tr>
<tr>
<td>Ground Truth vs. Perceived Truth</td>
</tr>
<tr>
<td>Intelligence Collection</td>
</tr>
<tr>
<td>Cause and Effect</td>
</tr>
<tr>
<td>Total Score</td>
</tr>
</tbody>
</table>
structional tool and identify areas for future development.

Assessment

The After Action Review (AAR) tool’s goal is to assess the player’s in-game choices and provide meaningful feedback. The AAR tool will focus on three areas: comparing ground truth to perceived truth, comparing available intelligence to collected intelligence, and analyzing cause and effect. Table 2 on the previous page describes the purpose, components, score and feedback for each of these assessment elements.

FUTURE RESEARCH

FPCT is new development and, as such, the next key step is to field test its usability and training concepts. The project is designed to be modular, so new or improved components can be integrated. We will integrate FPCT into larger, more complex systems. In future iterations, FPCT may become multiplayer and exist within a persistent world. Due to the high-fidelity emphasis of our research, immersive and extreme resolution technology may also be a logical next step for our development and research.

SUMMARY

FPCT is a 3D interactive training simulation that emphasizes learning cultural norms and values in order to successfully accomplish missions such as the identification of IED networks. In particular, the project emphasizes the ability to read culturally endemic non-verbal communication. The project has adopted a challenge to develop extremely high-fidelity representations in order to create an environment that serves as a tool for cultural training before or during actual deployment using the living-world construct. Other important elements of the development are psychological, emotional, behavioral and cultural models, and tools for subject matter experts to create specific cultures. In the future, we expect to integrate our visualization tools with other models and databases to further create representative synthetic cultures.

ACKNOWLEDGEMENTS

This project is possible from a contract from TRADOC G2 Intelligence Support Activity, and we appreciate their ongoing support. We also thank our colleagues at the University of Texas at Dallas for their contributions to the paper, particularly Dr. Frank Dufour, Brent Friedman, Erin Jennings, Lee Brown, Jumanne Donahue, Daniel Hurd, Jason Barnett, William Lemons and Phonesury Ounekeo.

We would also like to thank Dr. Andrew Blanchard, Dr. Monica Evans, Phillip Johnson, Ricardo Flores, Adam Buxkamper, Jay Laughlin, Michael Kaiser, and Michael Jelley for their ongoing support of our living-world research.

REFERENCES