Guided Conversations about Leadership: Mentoring with Movies and Interactive Characters

Randall Hill, Jr., Jay Douglas, Andrew Gordon, Fred Pighin, Martin van Velsen

USC Institute for Creative Technologies
13274 Fiji Way, Marina del Rey, CA 90292-7008
{hill,jdouglas,gordon,pighin,vvelsen}@ict.usc.edu

Abstract

Think Like a Commander – Excellence in Leadership (TLAC-XL) is an application designed for learning leadership skills both from the experiences of others and through a structured dialogue about issues raised in a vignette. The participant watches a movie, interacts with a synthetic mentor and interviews characters in the story. The goal is to enable leaders to learn the human dimensions of leadership, addressing a gap in the training tools currently available to the U.S. Army. The TLAC-XL application employs a number of Artificial Intelligence technologies, including the use of a coordination architecture, a machine learning approach to natural language processing, and an algorithm for the automated animation of rendered human faces.

Leadership Development

Leadership is difficult to teach, even for people. While there is evidence that some are born with an aptitude for leadership, the traits and skills needed to be an effective leader are often learned only by experience. This holds true across a diverse set of domains, including the corporate world, sports, firefighting and the military, which is the focus of the project described in this paper. Given that the military needs to develop a large number of leaders, it is imperative to find ways to accelerate the development process using whatever means possible.

The U.S. Army defines leadership this way:

> Leadership is influencing people – by providing purpose, direction, and motivation – while operating to accomplish the mission and improving the organization. (FM 22-100, 1999, p 1-4.)

To date, most of the Army's computer-based training systems for leaders use constructive simulations, which create an environment where commanders can practice mission planning and tactics. While these skills are necessary, they focus on the tactical and technical aspects of the job. Learning how to influence people, how to provide purpose, direction and motivation is simply not supported by most constructive simulation environments. While recent research on virtual humans and simulation attempts to address these issues, (e.g., Rickel et al., 2002), there are very few technical applications that support the development of a deeper understanding of interpersonal communication, building a positive command climate, motivating subordinates, and the many other human dimension factors that define an effective leader.

Furthermore, while the current generation of simulations can be used for modeling conventional warfare, today's military leaders face some of the most complex and challenging situations imaginable. To a greater degree than ever before, leaders at the tactical level – captains, lieutenants and non-commissioned officers (NCO's) – are being confronted with situations in the operational environment where their local decisions and actions can have strategic consequences, political and otherwise (McCausland & Martin, 2001). Over the past decade the military has been assigned a new class of missions requiring an expanded set of skills. Whereas the skills needed for war-fighting depend heavily on knowledge of tactics and battle drills, the new missions often have a different set of requirements. Peacekeeping, stability and support operations, humanitarian assistance, and homeland defense requires knowledge of the local culture and politics, as well as skills for dealing with a variety of outside organizations such as non-governmental groups, joint forces (inter-service operations), allied commands, and host nation armed forces.

The challenge for the U.S. armed forces is to develop leaders who have not only mastered the tactical and technical skills necessary to be competent commanders, but to be effective they must also develop intellectual flexibility, self-awareness, adaptability, and be able to deal with ambiguity, all under stressful conditions (Klein, 1999; McCausland & Martin, 2001; Ulmer, 1998; TRADOC, 2003).

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **2003**	2. REPORT TYPE	3. DATES COVERED **00-00-2003 to 00-00-2003**

4. TITLE AND SUBTITLE
Guided Conversations about Leadership: Mentoring with Movies and Interactive Characters

5a. CONTRACT NUMBER
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

5d. PROJECT NUMBER
5e. TASK NUMBER
5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of California,Institute for Creative Technologies,13274 Fiji Way,Marina del Rey,CA,90292

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**8**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

Learning with stories

Knowing how to motivate a subordinate, how to communicate a plan (or intent), and how to create a cohesive team are examples of skills possessed by effective leaders. Sternberg characterizes these skills as *tacit knowledge* (Sternberg et al., 2000), which is a form of procedural knowledge; it is practical by nature and not easily verbalized, and its mastery leads to success in a field or profession. Sternberg and his colleagues have studied tacit knowledge in a wide range of professions, including military leadership. To understand how the members of a profession become successful, stories are collected about problems or issues and the solutions that were either applied or learned by the practitioners. These stories are then used to identify and categorize the tacit knowledge that leads to the successful practice of the trade. In the context of their study of military leaders, Sternberg et al. developed and validated an inventory of tacit knowledge for military leaders that differed by echelon. In addition, they suggested some implications for leader development: (1) use the tacit knowledge categories identified in their inventory as sources to guide the experiences of a leader, and (2) use stories that illustrate a particular point as a launching point for an interaction with a mentor or coach. This is the first guiding principle of our application: use stories that illustrate a situation requiring leadership tacit knowledge to convey an experience to a learner. In fact we took this principle a step further by engaging professional filmmakers to craft the telling of the story.

The choice of Hollywood storytelling as a vehicle for establishing a tactical situation and for exploring key leadership issues was informed by both narrative theory and popular culture. Societal norms have long been transmitted through narrative, in the form of myths, fables, and fairy tales. The ability to form narratives is recognized as one of the important developmental stages in children, and use of narrative is a property of all cultures, not only those with "advanced" communication skills.

From childhood we learn that storytelling is the basis for effective communication. "When I was your age, I had a little red wagon," a parent begins a tale to soothe a child over the loss of a pet goldfish. An alternate approach, a description of nature's life cycle, though technically more accurate, is less emotionally digestible. Once the situation is framed by the narrative, however, factual information can be introduced, information that can affect the listener's behavior beyond the world of the story.[1]

That narrative provides a more engaging process of communication than chronologies (events delivered in chronological order) or other fact-based formats is a matter of anecdotal observation: even a mediocre film or novel lacks the narcotic effects of a textbook or lecture. Narrative theory offers a deeper explanation. As Lev Manovich observes (2001), the reader/spectator actively tests a narrative, making assumptions, accepting or rejecting them, filling in gaps in the narrative text, and creating whole characters out of the sketchiest of traits. Far from passively absorbing a narrative's content, the reader/spectator enjoys an active relationship with it. In turn, this relationship exercises the reader/spectator's belief and knowledge systems:

> ...fictions often have their effect because they call forth from memory real world events and causal possibilities. Even when the import of the original information is canceled out by virtue of its transparent fictionality, the rest of the accessed-belief structure remains intact. (Gerrig, 1993, p. 231)

By leveraging these narrative effects in a learning environment, we hypothesized that the viewer would be engaged on the multiple levels that narrative, and Hollywood, are known for, thereby enhancing the experience.

Learning through discourse

While a story is a powerful medium for communicating another's experience, a mentor can reinforce the salient points to be learned (Sternberg et al., 2000). It has long been recognized that students learn much more effectively when they have a tutor versus what they learn in the classroom. Bloom (1984) showed that tutored students scored on average two standard deviations higher than students who were taught in a traditional classroom setting. Chi et al. (2001) studied what makes learning with human tutoring effective and found that, among other things, tutoring is interactive by nature. Interactivity motivates the student more than passive listening, and it can result in deeper learning by promoting student explanation and reflection. Effective tutors have a knack for scaffolding in a dialogue, which leads to the construction of new knowledge. Graesser et al. (2002) also suggest that getting the student to ask deep questions and make explanations helps them to construct deep knowledge.

The TLAC-XL System

To capitalize on the effectiveness of both storytelling and discourse to achieve leadership development objectives, we developed a software system entitled *Think Like a Commander – Excellence in Leadership* (TLAC-XL). The target population, captains in the U.S. Army, interact with the system in a straightforward manner. First, they are presented with a short movie that depicts a situation where the leadership qualities embodied in the characters influence how the situation unfolds. Second, the users engage in a human-computer dialogue with the system about the leadership issues that are raised.

[1] For a discussion of the effects of narrative on real-world perceptions, see Gerrig (1993).

The dialogue in our system is held between the student and a synthetic mentor, as well as with some of the characters in the story. After viewing the vignette, the student is asked a series of questions by a synthetic mentor, which is embodied as a photo-real animated character. The format of this line of questioning is based on a classroom teaching methodology developed by the Army Research Institute (ARI) at Ft. Leavenworth, Kansas, known as *Think Like a Commander*, or TLAC for short. The purpose of the original TLAC format was to habituate commanders to ask eight critical questions when facing any operational scenario. These questions concerned the mission, the enemy, the terrain, the available assets, timing, the bigger picture, the visualization of the battlefield, and possible contingencies.

The original TLAC discussion format has been used extensively in classroom settings by ARI and the Army to teach commanders critical thinking skills about tactical situations. Our project adapted the original TLAC approach by first engaging the student with a question about the tactical scenario portrayed in the movie, and then raising a leadership issue related to the topic. For instance, the mentor initially asks questions about the mission, beginning with the student's interpretation of the mission and then goes on to ask about how the character in the story appeared to interpret the mission. The mentor then raises a leadership issue related to current TLAC point, where the issue is associated with a character in the vignette. This leads to a dialogue between the student and the vignette character. Here the student can ask the character questions related to the leadership issue, and the character responds in the form of a video clip that is most appropriate for the question.

Figure 1 presents a screenshot of the TLAC-XL user interface. The synthetic mentor appears in the lower right of the screen. A character from the vignette appears in the main upper left window, and responds to questions posed to him by the user.

While we call the interaction between the student and the mentor and the student and the characters a "conversation," it is really a scripted interaction that

Figure 1. A screenshot of Think Like a Commander – Excellence in Leadership (TLAC-XL)

follows the TLAC discourse, while allowing a great deal of flexibility with respect to providing responses to the student based on what the student said. The student interacts by typing questions and responses, but the mentor and the characters all give spoken responses. Thus, the total experience of the student is comprised of watching a movie, interacting with a mentor, and interviewing characters.

Architecture

The TLAC-XL system presents the user with a text-input console, a global navigation menu, a character window and a mentor window. Users can interact with a synthetic mentor and characters from the movie by typing questions into the console. In the TLAC-XL system a number of research efforts needed to come together in one single application. Due to the heterogeneous nature of all the components involved in the resulting application, an architecture was needed that created strong interactive bonds using open-ended software links. Various control and coordination techniques are available to coordinate the input and output of software components within a single or a distributed system, while still allowing them to operate independently of each other. We chose a TSpaces based event heap coordination architecture (Johanson & Fox, 2002) for our system for a number of reasons:

- Both synchronous and asynchronous events can be managed within the same control structure;
- Components attached to the event heap do not need to know about each other. This makes it possible to add a new component without disturbing any existing knowledge sources;
- The event heap facilitates a global interaction standard, instead of custom tailoring each component to each other;

Under most circumstances the system is in control over the navigation between mentor and characters. However a method was needed whereby the underlying software fabric could re-route input and output between components in a natural way. In an event-heap based architecture a number of knowledge sources interact with each other by adding and reading events from the shared data space. This event heap is managed by a control structure that has control over the distribution of events among all the knowledge sources that are subscribed to the event heap. Our control structure is able to seamlessly merge synchronous and asynchronous events, thus allowing partial scripts to be interleaved with spontaneous events. Behind the scenes a number of conversation graphs coordinate the answers of our virtual actors and provide the continuity of the overall dialog. The conversation graphs were originally written in the Java programming language while the main TLAC-XL application was written in C++. Our event heap architecture was designed to include a message based language bridge that can communicate with the event heap directly as a knowledge source.

Leadership Scenario

Students begin their interaction with the TLAC-XL system by watching a video of a fictional military operation where leadership issues arise. In our first TLAC-XL system, we authored a vignette that was based on the real world experiences of U.S. Army captains. We began by interviewing a group of ten captains stationed at the United States Military Academy at West Point. All of these captains had recently completed a tour as company commanders, so they had fresh experiences in that position that was conveyed to us in the form of stories that we solicited to illustrate their points. All ten captains told us about some of their most salient memories as commanders and the leadership issues they faced. Following the interviews we brainstormed ideas for a current operational scenario that could be used as the basis for a vignette. Based on this input we developed a humanitarian assistance vignette that takes place in Afghanistan, entitled *Power Hungry*. Working with subject matter experts from the Center for Army Leadership and the Army Research Institute, we went from a script by a Hollywood writer to a film shot in a mountainous, desert-like area in Southern California.

In the scenario, a company commander, Captain Young, has been given the mission to run a food distribution operation in an area where food is in short supply. The company quickly runs into a number of obstacles, beginning with how to secure the site given the nature of the terrain – soft soil, located in a bowl surrounded by hills and two possible entry points. It is necessary to create lanes with wire to keep control of the crowds that are expected to arrive soon. The company's lieutenants begin rigging the site, but their plan does not satisfy the commander, who directs the executive officer to start over, giving very little guidance other than to stall the food trucks in order to allow time to prepare the site. In the mean time first one then another local warlord appears, offering to "help" with security. Turning away the warlords proves difficult, particularly due to conflicting advice from a brigade command sergeant major (CSM), who happens to be in the company's area site escorting a media crew. The brigade CSM plays a significant but ambiguous role in the vignette. He offers advice that seems to suggest that he has some inside knowledge about the brigade commander's intent. His advice runs counter to the commander's instincts in several instances, and the captain listens. At his suggestion the commander meets with one of the warlords to discuss the situation. Meanwhile the situation worsens as the executive officer is unable to delay the trucks, and after some twists and turns in the story, the warlords hatch their plot to take control of the food. The full duration of the *Power Hungry* vignette is slightly more than thirteen minutes.

This vignette was authored so as to incorporate six specific leadership issues that were raised by the U.S. Army captains that we interviewed. While each of these issues involves the behavior of the fictional captain in our vignette, the vignette was authored in such a way as to associate each issue with a different character. For example, the unexpected presence of a brigade command sergeant major causes some problems for the captain in the vignette related to the influence that is brought to his command decisions. Here the leadership issue is one that concerns the captain, but the issue is associated with the character of the command sergeant major in this vignette. During the interactive portion of the TLAC-XL system, students are given the opportunity to question each of the characters directly about the leadership issue that they are associated with. The six leadership issues in the *Power Hungry* vignette are as follows:

1. Shared vision of intent (LT Perez)
2. Command influence (CSM Pullman)
3. Setting a model of command (LT Wychowski)
4. Clarity of mission (CPT Young)
5. Cultural awareness (Omar the warlord)
6. Respect for experience (SGT Jones)

Classification-based conversations

After watching the video of the vignette, the trainee begins a question-answer dialogue with a virtual mentor. The virtual mentor, visualized as a photo-real animated character, poses questions to the student, who responds by entering natural language text using the keyboard. Within the course of this interaction, the virtual mentor introduces characters from the vignette, and allows the student to compose questions to them directly. Responses from the vignette characters are presented as video recordings.

In each dialogue mode, either answering questions from the mentor or asking questions of vignette characters, appropriate responses must be presented to the trainee to achieve a sense of coherence in the dialogue and as well as pedagogical goals. To accomplish this, we follow statistical, machine learning approach for processing the natural language input of the user. At any point in the interaction in either dialogue mode, there are a fixed number of pre-authored media items that are possible to present to the trainee, each of which would move the conversation forward one turn. The task, therefore, is to select the most appropriate member of the set of possibilities given the trainee's textual input. By using a statistical, machine learning approach, where the trainee's input is classified based on the available supervised training data from previous users, acceptable levels of performance can be obtained in a manner that is robust to slight variations in language use.

Classification algorithm

To perform a correct classification of the textual input of a trainee using a machine learning approach, we employ a Naïve Bayesian classification algorithm (George & Langley, 1995) implemented in the WEKA open source toolkit (Witten & Eibe, 1999). To construct feature vector instances for training and test data, we treat user text inputs as a set of features consisting of individual words (unigrams) and adjacent pairs of words (bigrams). Feature vectors are constructed for instances without using stop-lists filters, without truncating the features space, by ignoring punctuation and variation in case, and using feature counts for feature values, although feature counts are very rarely greater than one for a given instance.

In order to aid in the development of an operational prototype, the training data used for classification of trainee textual input was seeded with training examples fabricated by our development team to serve as a placeholder in the absence of real data from our user population. As more legitimate data was being collected, it became evident that the seed examples were indistinguishable from the real data in form and content, and were retained in the complete training data set. Examples of the seed data for a single class are as follows:

Class: *Mission-intent*
What was your understanding of the mission?
What was your mission?
What do you think the purpose of this operation was?
What were you trying to accomplish here today?
What is the goal of this food distribution operation?
Did you understand the purpose of this mission?

Classification performance

To evaluate the performance of this approach to trainee input classification, a cross-validation analysis (10-fold) was performed using 6 sets of supervised training data, one for each of the classifiers that is used to select the most appropriate response to a trainee's question during character interviews. Although both the mentor interaction and the character interviews employ the same classification approach, the mentor interaction was structured in a way where there were at most two possible mentor responses for an answer typed in by a trainee (corresponding to agreement or disagreement). In contrast, the character interviews are much more demanding on the classification algorithm, where there are an average of 13 possible character responses available.

Figure 2 presents the results of the cross-validation analysis for each of the six character interview classifiers used in our system. Accuracy is presented as the likelihood that a novel input will be correctly classified, and performance levels for the initial seeded training data are presented along with that obtained through the addition of

Character classifier	Classes	Seed instances	Seed accuracy	Total instances	Total accuracy
Jones	8	48	58.3%	128	62.5%
Omar	11	66	72.7%	187	68.4%
Perez	15	90	72.6%	175	73.1%
Pullman	13	78	62.3%	221	65.2%
Wychowski	10	60	58.3%	142	61.3%
Young	19	114	66.7%	309	63.8%
Average	*12.67*	*76*	*65.15%*	*193.67*	*65.72%*

Figure 2. Character Interview Classifier Performance (10-fold cross validation)

legitimately collected instances. Interestingly, the admittedly modest amount of legitimate training data that we have been able to collect thus far has not significantly improved the level of performance beyond what was obtained using the initial seed data. The Naïve Bayesian learning algorithm outperformed several other approaches that we evaluated for this classification task, with C4.5 rule induction performing almost as well. However, contemporary kernel methods and support vector machines were not evaluated, and we expect that greater performance could be obtained by capitalizing on recent advances in these methods.

Conversation graphs

In order to design effective interactions between trainees and the system, we encoded the set of possible trainee/system dialogues as a directed finite-state graph. Each node in the graph represented a dialogue turn where the system said something (using media), and each arc in the graph represented a classification of the trainee's typed input. Every node in this graph that has more than one arc transitioning away from the node requires a separate classification of the trainee input. The section of this graph representing the mentor interactions include 12 separate classifiers for this purpose, mainly to determine whether or not the mentor should agree or disagree with a trainee's response to a mentor's preceding question. However, each of the six character conversations is driven by a single classifier, which selects the most appropriate answer from the character. Graphical representations of the mentor graph and a character interview graph are presented in figures 3 and 4.

As seen in Figure 3, the mentor interaction can be viewed as an eight-tiered interaction, where each tier corresponds to a line of questioning that concerns one of the eight Think Like a Commander (TLAC) points used in the previous work of the Army Research Institute. Within each tier, the mentor begins by asking a few preliminary questions about the topic (e.g. "What was your understanding of the mission?") that lead to one of the six critical leadership issues that were brought up in the vignette. To explore these leadership issues (if necessary, based on the user's response to a poignant question), the mentor will allow the character to conduct an interview with a relevant character from the vignette. Each node labeled with a letter in the mentor graph indicates a point where the mentor introduces a character, invoking an embedded subgraph corresponding to a character interview. At the end of an embedded character interview, the mentor asks a follow-up question aimed at determining the trainees understanding of how the leadership issue relates to the given Think Like a Commander point, then moves on to the next point.

Figure 4 illustrates the general shape of an embedded subgraph for supporting a trainee-led character interview. A single classifier is used to route a trainee's question to

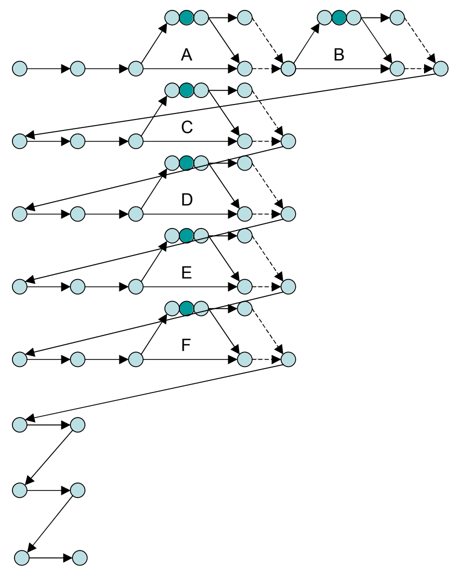

Figure 3. The mentor graph

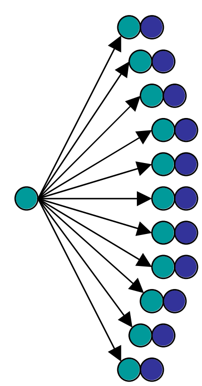

Figure 4. A character interview graph

one of a set of possible character responses. The embedded subgraph is used repetitively to allow the trainee to ask multiple questions, until they indicate to the system that the interview is over by means of a user interface button. When a trainee's textual input is classified to the same category over multiple repetitions (the system believes they are asking the same question twice or more), a secondary media item is presented to the user, typically where character states that they've already answered that question, and they have nothing more to say on the matter.

Animated Mentor

To support the conversational interactions with the mentor, we developed an animated character (Figure 5). One of the requirements for our character is that he should look lifelike and engaging to the trainees. We leveraged computer graphics technology to bring this character to life and build a digital talking head that can be animated for an arbitrary input sentence. Our approach falls within the realm of visual speech synthesis: the facial animation system takes as input a speech signal and output the corresponding animation.

Realistic animation of a synthetic human is a difficult task due to the complexity of the human body, one that traditionally involves many digital artists in the special effects industry. We took advantage of motion capture technology to bring realism into the synthetic mentor at an affordable cost. Motion capture allows the accurate recording of live actors' motions. We used this technology to record a large database of speech related motions from a live actor. We then analyzed this data to build a generative statistical model of these actor's facial motions. This model used the database of motions indexed with speech. We organized this database according to the phonemes of the recorded speech: each phoneme is associated with a large number of motion fragments.

Figure 5. Synthetic Mentor

To generate animations from our model, given an input speech, we first segment it into phonemes. This string of phonemes is then used as a guideline to extract from the motion database a corresponding sequence of motion fragments. The motion fragments are optimally chosen to maximize the fidelity of the synthesized motion. We stitch the sequence together to produce a facial motion that both matches the input speech and is visually realistic.

Results

At the time this paper was written, two sets of evaluations have been conducted by the Army Research Institute to study the effectiveness of the TLAC-XL system. The first consisted of an initial series of formative evaluations at Ft. Lewis, WA, aimed at developing the evaluation method itself. As TLAC-XL involved a non-traditional interaction with students and subtle training objectives, it was necessary to investigate appropriate techniques for obtaining pre-test and post-test data from subjects. This first evaluation provided us with one specific and unexpected result. In most military training scenarios the final outcome of the operation is overwhelmingly positive. However, our story ends in a failure of the mission. As a consequence, our test subjects were highly disgruntled by what they saw, in most cases. At first, the evaluation team viewed this negative response as an apparent failing of the system. However, the agitation expressed in our subjects appeared to support the interaction that occurred after watching the story. Most test subjects used the interactive portion of the session to vent their frustrations concerning the mission to the virtual mentor and virtual characters.

A second set of evaluations was performed at Ft. Drum, NY. Here, more evidence was gathered to suggest that the frustration evoked by watching the vignette can provide a strong force for learning, leading our subjects (U.S. Army captains) into heated discussions. In this set of evaluations, subjects would spend 1 1/2 hours to 2 hours with the system on average, and engage in additional discussions with evaluators concerning various possible outcomes and solutions. To evaluate the relative value of guided conversations with interactive characters versus traditional classroom methods, a comparison was conducted between TLAC-XL and a slideshow version of the scenario. Early results of this comparison suggest that the slideshow variation was effective at presenting the scenario in a way that enabled students to remember facts about the mission. However, subjects using the TLAC-XL application had an additional understanding of the interpersonal dynamics that contributed to the failure of the mission that went beyond the factual details of the scenario.

Through these and other evaluations, we have learned a number of lessons about the guided conversations. When students ask questions within the scope supported by the conversation graph, the answers can appear to be highly realistic and engaging. When students ask questions of virtual characters that are outside the expected scope, the irrelevant answers that are given in response can be frustrating to the student, but can also give the appearance that the character is simply avoiding the question. Also, it appears that failures in classifying students' questions can be mitigated somewhat by responding with engaging content. That is, the students may be less frustrated with a character response that is not relevant to their question as long as it is interesting in its own right and relevant to the larger topic of conversation.

The TLAC-XL system has been demonstrated to a broad range of U.S. Army officers ranging in rank from lieutenant to general. The universal reaction to the vignette has been that it is very engaging and stirring. Besides good storytelling, one of the reasons we believe that the vignette has been so well received is that it hits several areas that the Army currently needs to cover in leader development, but does not have any technological support. The scenario encompasses a contemporary operational environment, a food distribution operation in Afghanistan, which is in the Army's new spectrum of operations. Furthermore, it raises cross-cultural issues, interpersonal communication, command climate, and a number of the other human dimensions of leadership.

Future Work

There is a lot of work we would still like to do on this project. To more fully support deep learning we plan to take seriously the need for student modeling, analysis of the input, and providing customized feedback. In addition we plan to incorporate tutoring strategies based on the kinds of questions asked by participants. It has been observed by our ARI colleagues that less experienced leaders may not have the ability to ask the right questions. A skilled tutor knows how to ask telling questions in these instances, to prompt the generation of a more focused question that may not have been considered otherwise. In addition, we plan to expand the capabilities of the animated tutor to incorporate text-to-speech technology, enabling an even greater degree of customization. At the prompting of our colleagues in the Army, we plan to provide multiple identities for the mentor to represent other races and genders.

Acknowledgements

This paper was developed with funds from the U.S. Army Research Institute for the Behavioral and Social Sciences under ARO contract number DAAD 19-99-D-0046. Any opinions, findings and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the Department of the Army. We wish to thank the members of the ARI and the Center for Army Leadership team—Stanley Halpin, Robert Solick, Rex Michel, Larry Laffitte, Jason Ward, Michelle Zbylut, Zita Simutis, COL Mike Flowers and Clark Delavan—for their support on this project.

References

Bloom, B., 1984. The 2 Sigma Problem: The Search for Methods of Group Instruction as Effective as One-to-One Tutoring. *Educational Researcher*, 13,6, June/July 1984: 4-16.

Chi, M., Siler, S., Jeong, H., Yamauchi, T., Hausman, R., 2001. Learning from Human Tutoring, *Cognitive Science* 25, 471-533.

Department of the Army, 1999. *FM 22-100, Army Leadership: Be, Know, Do*, Headquarters, Department of the Army.

TRADOC, 2003. *Pamphlet 525-66, Military Operations: Force Operating Capabilities*, U.S. Army Training and Doctrine Command.

George, J. & Langley, P., 1995. Estimating Continuous Distributions in Bayesian Classifiers. *Proceedings of the Eleventh Conference on Uncertainty in Artificial Intelligence.* 338-345, San Mateo, CA: Morgan Kaufmann.

Gerrig, R, 1993. *Experiencing Narrative Worlds: On the Psychological Activities of Reading*. New Haven: Yale UP.

Graesser, A., Person, N., Hu, X., 2002. Improving Comprehension Through Discourse Processing, *New Directions for Teaching and Learning*, no. 89, Spring, San Francisco, CA.: Jossey-Bass.

Johanson, B. & Fox, A. (2002) The Event Heap: A coordination infrastructure for interactive workspaces, Proceedings of 4th IEEE Worshop on Moble Computer Systems and Applications, Callicoon, NY.

Klein, G., 1999. *Sources of Power: How People Make Decisions.* Cambridge, MA: The MIT Press.

Manovich, L., 2001. *The Language of New Media*. Cambridge, Massachusetts: MIT Press.

McCausland, J. & Martin, G., 2001. Transforming Strategic Leader Education for the 21st-Century Army, *Parameters*, Autumn 2001, U.S. Army War College, Carlisle, PA.

Rickel, J. Gratch, J., Hill, R., Marsella, S., Traum, D., & Swartout., B. 2002. Toward a New Generation of Virtual Humans for Interactive Experiences, IEEE Intelligent Systems Journal.

Sternberg, R., Forsythe, G., Hedlund, J., Horvath, J., Wagner, R., Williams, W., Snook, S., and Grigorenko, E., 2000. *Practical Intelligence in Everyday Life*, New York: Cambridge University Press.

Ulmer, W., 1998. Military Leadership into the 21st Century: Another "Bridge Too Far?," *Parameters*, Spring 1998, pp. 4-25, U.S. Army War College, Carlisle, PA.

Witten, I. and Frank, E., 1999. *Data Mining: Practical Machine Learning Tools and Techniques with Java Implementations.* San Francisco: Morgan Kaufmann.