Critical Thinking Training: Proven New Technologies for Engaging DoD Personnel

Dan M. Davis
HPC-Education/USC
Long Beach, California
dmdavis@acm.org

Nicholas J. Kaimakis
Univ. of Southern California
Los Angeles, California
kaimakis@usc.edu

Howard Spaulding
Training Consultant
Arlington, Texas
pirate612@cablelynx.com

ABSTRACT

Critical thinking is widely accepted as a requisite skill for successful military leaders. Today's defense leaders must thwart sudden assaults with preemptive parries and innovative defensive tactics. This requires a level of critical thinking and meta-cognition that is not always found in educational processes today. Effectively enhancing that skill set via a range of training techniques has a well-documented history, but the current application of those techniques remains problematical for a number of reasons. The authors' past and current research is in the use of conversational on-line agents in a number of contexts and in the use of distributed computing and simulation. They assert these can be used to address problems in critical thinking training. These capabilities tender a solution to global application and continuous availability of this training for DoD personnel. The need for critical thinking is taken as a given. A brief survey of effective critical thinking pedagogies is followed by an analysis of which would be most effectively adapted to the distributed simulation/agent environment. Then this paper discusses both the distributed computing simulation capabilities that have been demonstrated and the emerging technologies enabling the delivery of engaging conversational agents. These agents are capable of carrying on a life-like verbal exchange with the users. Examples of where and how both the distributed computing and agent interfaces have proven effective are given, with analysis about the impact these capabilities would have on critical thinking training. A discussion is presented concerning the relative benefits of using animated agents vice using video clips of *live* personnel as the response vehicle for the conversation. The efficacy of using an active voice recognition-enabled dialogue instead of a text delivered didactic approach is discussed. The paper closes with a description of how such a system could effectively serve DoD personnel.

ABOUT THE AUTHORS

Dan M. Davis is a consultant for the University of Southern California (USC), focusing on large-scale distributed DoD simulations and conversational computer agents. Pre-retirement, for a decade he was the director of USC's JESPP project for JFCOM. As the assistant director of the Center for Advanced Computing Research at Caltech, he managed Synthetic Forces Express, bringing HPC to DoD simulations. Prior experience includes serving as a director at the Maui High Performance Computing Center and as a software engineer. He has served as the chairman of the Coalition of Academic Supercomputing Centers and has taught at the undergraduate and graduate levels.. He saw duty in Vietnam as a cryptologist and retired as a commander, USN. He received B.A. and J.D. degrees from the University of Colorado in Boulder.

Nicholas J. Kaimakis is a Computer Science major in the University of Southern California's Viterbi School of Engineering. He has been active in research at USC's Institute for Creative Technologies since 2016. His current research thrusts are in the use of computer-generated avatars or video clips, animated and directed by natural language optimized artificial intelligence programs that present a life-like dialogue capability to interact with remote users. The project aims to improve knowledge of technical fields across varied demographics through virtual interaction.

Howard Spaulding is a retired cryptologic technician (interpretive) senior chief petty officer and recently retired from Schneider National where his duties focused on training. His military service spanned three decades including 28 months in Viet Nam as a marine cryptologic linguist. His naval service included tours in the Philippines, Thailand and Japan. He also served as the assistant officer in charge for a highly technical project under the supervision of CNSG and SPAWAR and was instrumental in the interior design of the ES-3A aircraft. He made two deployments aboard an allied naval vessel as the lone area/target expert. He received a BA from California State University, Long Beach in 1983 while still on active duty**.**

Critical Thinking Training: Proven New Technologies for Engaging DoD Personnel

Dan M. Davis
HPC-Education/USC
Long Beach, California
dmdavis@acm.org

Nicholas J. Kaimakis
ICT/Univ. of Southern California
Los Angeles, California
kaimakis@usc.edu

Howard Spaulding
Training Consultant
Arlington, Texas
pirate612@cablelynx.com

INTRODUCTION

This paper will assumes that critical thinking is vital to a strong defense posture and will advance two theses: 1) Critical thinking can best be inculcated with skill training and 2) critical thinking can optimally be implemented with globally accessible computer-aided instruction. The support presented will be both from the authors' personal research and insights arising out of their own military service and from the community at large.

The paper will begin with the definition of some terms that will be used.. The authors take the position that improving these skills would optimally be a training function for today's military forces. The barriers to the adoption of these concepts in what the authors believe to be an already over-tasked and under-funded defense force will be discussed. The advantages of computer-aided training will be laid out and several existing and emerging technologies will be highlighted. There are sections on the vision for implementation of such a system and on the envisioned users' experience.

Terminology

As used in this paper, the term critical thinking will refer to an intellectual process that is characterized by being purposeful, reasoned and goal directed, all the while not being disrupted by emotion, prejudice, or mysticism. This process is invoked to solve problems, formulate inferences, calculate likelihoods, and make decisions (Halpern, 1997). Critical thinking encompasses Col. John Boyd's OODA Loop concept (Coram, 2002), which is mentioned.

It is the authors' position that fostering and enhancing critical thinking is most amenable to training. That may be at odds with the notion that is suggested by the focus on the differing types of education that are thought to achieve these goals, *e.g*. the use of Socratic methods or constructivist approaches to education. The literature seems to support that a directed approach to training a set of steps for critical thinking will result in a more reliable inculcation of these skills in the target group (Lehman & Nisbett, 1990). Naturally, some researchers use different terms for this formal reasoning process.

Meta-cognition is a high-order intellectual process in which one examines one's own internal thinking processes and uses that knowledge to plan future mental activity, monitor the efficacy of that approach, and make corrections as appropriate. Whether driven by intuitive insight or guided by external training, it allows the practitioner to recognize and improve rational contemplation of future actions (Flavell, 1971). An early approach to using computational technology to enhance meta-cognitive skills was advanced in an earlier paper (Davis, Curiel & Davis, 2010).

The use of computer generated on-line representations of human-like interfaces can be called avatars (computer controlled) or agents (user controlled). Conversational on-screen interfaces seem to the authors to range back and forth between these two definitions. Both terms are used below. In the work discussed below, they are not even purely virtual, but are video recordings of live humans. While they act on their own, their actions are triggered by user input. Test users have asked, "Is this person really live and just answering my questions from elsewhere?"

CRITICAL THINKING SKILLS TRAINING

Positing the manifest and patent benefits of critical thinking, the question arises as to the possibility of enhancing the skill level in this domain. It must be acknowledged that there are many factors contributing to critical thinking: genetic inclination, childhood experience, formal training, motivation, and natural selection based on trial and error.
Many feel that critical thinking can be taught, but some are skeptical (Willingham, 2007). From the literature, the authors have extracted three major ways in which the improvement of critical thinking is approached: 1) training a

rubric to be followed, 2) forming the process by the Socratic method, and 3) using constructivist methods emphasizing self-discovery. Many authors report using combinations of these three approaches. This paper does not have the space to resolve the varying benefits of the differing approaches (Pither & Soden, 2000), but will proceed on the assumption that the rubric training approach may be the most applicable to the defense environment, all the while leaving open the use of the other methods as may be appropriate.

Many studies have found that rubric training is effective. One found that the students demonstrated an effective use of the techniques they had learned and effectively displayed it in an encounter during which the researcher did not reveal that they were assessing the subjects' critical thinking skills (Lehman & Nisbet, 1990). One Belgian review of approaches found the multitude of variables made assessing the efficacy of instruction very difficult to adequately quantify in order to allow comparative analysis of the various techniques (Tiruneh, 2014). A similar study in a professional school setting, nursing, came to virtually the same conclusion (Carter, et al., 2016). This study is especially germane when looking at the stress, time pressures, and criticality of sound decisions that are common to both combat and the emergency room. These all stress the need to rigorously establish and scrupulously monitor skill parameters in any initiative to improve critical thinking (Halpern, 2002).

There may be a preliminary consensus on the steps of the process from which a working model can be extracted (Lai, 2011).. These may be something on the order of:

Collecting data	Reviewing initial conclusions
Categorizing and analyzing	Combining ideas and expanding uses
Using the insights gained	Internalizing the high order concepts

All this activity must be accomplished in the face of the ever-present human tendency to reject logic for a more comforting environment, where wishes and mythology reign (Ariely, 2008).

Teaching the critical thinking processes to students has been conducted via a wide range of educational techniques:

Didactic lectures	Text book exposition
Small group exercises	Model analyses
Text book exposition	Socratic dialogues

BACKGROUND AND INTRODUCTION TO GERMANE TECHNOLOGIES

This section will present a quick survey of approaches and developments at the University of Southern California, primarily at the two home institutes of the authors: the Information Sciences Institute (ISI) and the Institute for Creative Technologies (ICT). The technologies will be discussed in greater detail below.

ISI is a four-decades old research institute, almost entirely funded by various entities of the DoD. An early proponent and developer of the Internet, it has focused its activities on high performance computing and communications in all areas, especially artificial intelligence and battlefield simulations. ISI was a leader in the use of high performance computing in battlefield simulations (Davis & Lucas, 2006) and implementing graphic processing units as accelerators (Wagenbreth, Lucas, & Davis, 2010). For four-years, their high performance-computing group has operated one of the four operating quantum computers in the U.S. This technology promises to bring almost unimaginable capabilities to new training approaches (Schuld, Sinayskiy, & Petruccione, 2015).

ICT, grew out of ISI about a decade and a half ago. It is an US Army funded University Affiliated Research Center (UARC). Its *raison d'être* is the melding of USC academic research and the Los Angeles area's world-renowned entertainment industry skills. The major thrust is in training and immersive virtual reality. They have developed a number of recognized breakthroughs in training and information presentation. Three of the best known, and most highly relevant to this papers focus, are the immersive training system for preparing US Army personnel facing IEDs (CHAOS), a program providing a mentor for PTSD patients (SimCoach), and a program to archive and deliver conversational memories of holocaust survivors (New Dimensions in Testimony; ICT, 2018).

Deep learning enhances the abilities to analyze huge data sets by invoking hierarchical neural nets, GPU acceleration is a well-proven technique for increasing the power of huge cluster computers with inexpensive hardware, and emerging capabilities in quantum computing which could increase computational power in computer education by orders of magnitude. Deep learning (LeCun, Bengio & Hinton, 2015) is currently used effectively in image recognition and natural language processing. The use of GPU's is now recognized as an enabling technology for these techniques (Oh & Jung, 2004). Assuming deep learning extensibility to behavior recognition and trend isolation, attention to hardware support should be considered for GPU use and the instruction for their implementation (Wagenbreth, Davis & Lucas, 2010). The combination of this technology (Yao, Kaimakis, Liu and Davis, 2018) with emerging quantum computing capabilities (Lucas, Davis & Burns, 2015) is also a potential anew area of research.

A COMPUTER-AIDED TRAINING APPROACH

New advances in high performance computing, globally distributed computer communications, animated or video-captured agents, natural language processing, and human-computer interaction technologies can effectively interface with the user, both in terms of geographical location and personal proclivities. They will be discussed in the order originally mentioned, with no implication as to the relative seriousness of the issue or the efficacy of the recommended computer system's response to that issue.

At ICT, three types of display are used to best communicate with different target audiences, see Figures 1, 2, and 3. If an animated avatar or agent is used, the appearance of the avatar can easily be altered to present different ages,

Figure 1 - Animated SimCoach CGI with alternative characters

Figure 2 - Videotaped holocaust victim in 3D Holographic display

Figure 3 - Videotaped interview displayed on 2D monitor

genders, ethnic appearance and social/military status. Once this capability is prepared, the agent can be available to engage in a "one-on-one" conversation with as many users as the servers can support, and seen anywhere in the world there is network capability. Further, it can be provided at any time and is not seriously disrupted by unanticipated breaks in instructional time, being ready to resume where it left off immediately upon being cued that the user is on-line again. This all-day availability and immunization against disruptions due to intruding operational requirements should addresses many of the above noted operational constraints.

A large portion of communication becomes more compelling, more cogent and more comprehended if presented by a human (EduNote, 2016), but the provision of humans is expensive and sometimes dangerous, *e.g*. sending instructional personnel into combat zones. Technologies conceived, developed, and validated at ICT have fielded very engaging and lifelike conversation computer interfaces (Reger, Rizzo & Gahm, 2015). These make use of a wide range of computer capabilities to input language via both keyboard typed and spoken words from the user, analyze that content, select the appropriate response to the input, then cue up and display the response on-screen, so quickly that it accurately mirrors the pace of a human-to-human conversation (ICT, 2015a).

Spawning new avatars or agents requires very little marginal costs, mostly additional computer processing power and storage space. Carefully and easily modifiable design characteristics of the agents also can provide a range of styles and *looks* to satisfy the majority of users' needs for a mentor fitting their image of a senior advisor. In fact, they can switch with whom they interface with no penalty in terms of costs or continuity. While websites like the Khan Academy (See: Khan, 2016) offer excellent instruction via written and visual materials, those kinds of interfaces lack the compelling quality of a face-to-face conversation with a responsive agent. Especially in the area of

Socratic dialogues, the responsive avatar has been shown to be more arresting and retain interest of the user longer and elicit more conversational input from the user (ICT, 2015b).

So, what would such a system look like and what would be the benefits of such an approach? This kind of system should be cross-platform compatible, which is easily accomplished with current Web technologies, including detecting the client platform and adjusting for bandwidth constraints. For the program to attract and retain a large user community, it must be readily accessible from a wide range of platforms (*e.g.* desktops, laptops, tablets, and smart phones) and be globally accessible any time. The authors have supported such efforts in the past and recognize both their potential and their hurdles *e.g*. speed of light induced latencies when server and client are separated on the order of five thousand-miles. (Gottschalk, Yao, Wagenbreth, & Davis, 2010) While the system discussed in that paper was designed to advance analysis and evaluation for the joint experimentation directorate of JFCOM, its training function was mentioned by many of the active duty participants who provided the red and blue forces for the exercises. The authors observed the total immersion of the personnel in the combat on the screens and the enthusiasm with which the participants engaged their duties (Davis, Lucas, Amburn & Gottschalk, 2005). One post-exercise celebratory meal was nearly disrupted by a fist fight between the red force team and the blue force team until intercession by the academics present restored civility to the proceedings, which the authors took to be indicative of the users' emotional involvement in the immersive experience.

The trans-continental capabilities of such exercises stand as solid evidence that a global network could reliably provide computer-aided training to advance critical thinking skills. Figure 4 shows a notional diagram of the network for one of these exercises and it exhibited round trip latencies well under 500-milliseconds, which should be adequate for conversational flow. Operators in Virginia could access computer capabilities on Maui with no apparent loss of realism. There were some issues with use of pointing devices such as mice, as these are susceptible to operator disorientation from latencies even as short as half a second. The cursor would move on the local monitor, but the resultant impact on the image, controlled by the remote process, was a quarter second later. These issues were met by prepositioning interface data on local computers and caching of instructions. Conversational delays of this length are not uncommon in everyday speech, so should not raise the same concerns. Bandwidths are not constraining during normal operations and training would be suspended during times of national emergency.

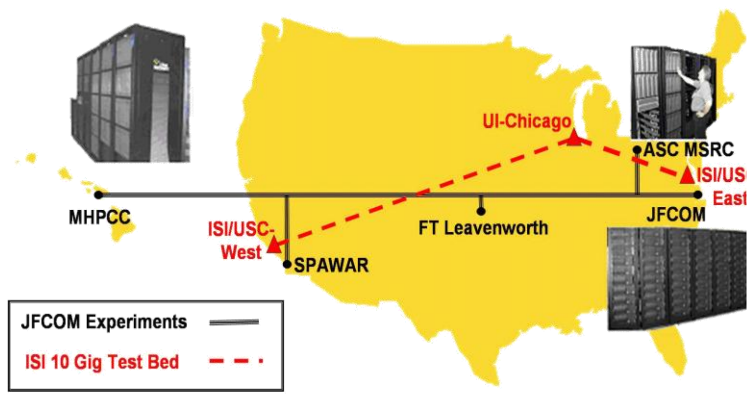

Figure 1 - USC-enabled Trans-continental Network Testbed.
Maui to Virginia round trip latencies: 204 msec

Having this infrastructure advantage, the system could present a range of didactic techniques, ranging from pure lecture to interactive tutorial approaches. It could cover historical examples, long a favorite of military educators, or futuristic *what if* simulations. In both cases, the students would be driven by direction from the tutor/mentor to consider their critical thinking strategies and critique their approaches in a meta-cognitive analysis.

Advanced artificial intelligence technologies, especially deep learning (Deng & Yu, 2014), would allow the system to evaluate new trends in student issues and recommend to programmers the changes needed to make the system increasingly topical and compelling. Simulation techniques would allow application of decisions made by the students to be rendered in realistic settings, be they historical, current or prospective. These could be invoked to assess the validity of their decisions. The outcomes would, of course, be caveated with the warnings that they represent a stochastic representation of the world and outcomes will vary, as they do in real world situations (Wong, et al., 2007). The major goal is to get the students to become engaged in the habit of continuously analyzing their thought processes and making corrections as the situations mandate (Tetlock & Gardner, 2016).

For those who are more engaged with tactical duties, the Boyd OODA Loop analyses may be more appropriate. There are a number of approaches in use that purport to improve OODA Loop performance. Many of these (von Lubitz et al., 2004) would be readily implementable into the proposed global critical thinking network. Again, high-resolution battlefield scenarios would be an excellent vehicle for both training and evaluation. Experience with popularity of America's Army, a product designed with an attractive interface and interactive structure, demonstrated its mettle and measure performance training levels (Jean, 2006).

In a different setting, i.e. academia, skills such as identifying limitations and opportunities were increased four times compared to the skills of subjects who had no training. Trained subjects were twelve times more likely to propose changes spontaneously (Holmes, Wieman & Bond, 2015).

Implementation Vision

While there are a wide range of implementation approaches that would satisfy the primary goal of this initiative, the following example of an approach may give a little more clarity to a future implementation of a new globally accessible critical thinking training enhancement system. The individual technologies are available; implementation requires only system-engineering and careful execution by an experienced military, industry, and academic research team. A necessary first step would be identifying and parameterizing the goal. Much previous research on this matter will be of use especially when it focuses on computer implementations addressing critical thinking, like a study that quantified critical thinking and argued that a computer tool to impact this, demonstrated improvements (Garrison, 2001). The primary goal would be the establishment of a consensus view of critical thinking and meta-cognition, leading to a set of accepted measures of the skill sets impinging on those processes (Leedom, et al., 2007).

The suggested use of computer-generated simulated situations as a vehicle for improvement requires that some thought be given to the situations that should be implemented. One solution would be to create a multi-dimensional matrix of these scenarios. The first dimension would be the scenario's position in world history. In order to engage and maintain the interest of each participant, having this range of settings could allow each of the users to pick the period of history that is the most enticing. The impact of America's Army on training was generated by the original entertainment value of the program, which had been developed as a recruiting tool, not a training program. Having a set of problems from different time periods could go a long way to retain users. The periods offered could be: classical, medieval, Napoleonic, U.S. Civil War, the World Wars, Vietnam, and today. These would offer attractive alternatives to personnel with differing personal enthusiasms including a present day or future setting.

Selection of the particular pedagogical approach to enhancing critical thinking skills would also require careful consideration. Depending on time and funding constraints, several approaches could be analyzed and the best ones implemented for dissemination on a global basis. Even though these varying approaches are all not available at all times in a single-teacher class room, the pedagogical styles offered to the learner in this system can span the entire spectrum with the use of a single teacher or agent.

The next dimension would be the command level of the personnel. In the officer ranks, this would be something on the order of company grade, field grade and general officer for the Army and the Marine Corps and division officer, commanding officer, and flag officer for the Navy and Coast Guard. Each level would require a careful examination of the issues likely to be faced by officers of that rank. In the enlisted ranks, modules could be developed for three appropriate ranges, *e.g*. the Navy might want to trichotomize their training into seamen, petty officers and CPOs.

A research issue worthy of interest would be ascertaining the final parametric for the ability to synthesize all of the skills needed to effectively optimize critical thinking.

A major feature of this initiative would be the provision of on-going mentorship by avatars or agents. Whether considered mentors or Socratic tutors, these activities would require a significant amount of research on their responses, video recording of answers/questions, and editing, in order to provide the seamless interface needed for conversational fluidity. Previous research has indicated that a response repository on the order of five hundred to fifteen hundred responses is hardly ever stumped by a user, but mentorship would require more than just appropriate information, e.g. mentorship is often based more on critical questions than it is on providing informational statements.

As a model for what the projected *conversational avatar* implementation critical thinking interaction may look like, the authors recount their experiences with the staff at Next Generation Leaders. It is an educational counseling and scholastic augmentation institute that focuses on inculcating skills and providing exposure to ideas that are not present either in formal school settings or for home-schooled students. One of their foci is on critical thinking (Snyder, & Snyder., 2008). They follow the method that poses a series of question to the students: “What do you think about that?”, “How confident are you in that?”, “What other possibilities are there?”, “Can you make a good argument for your theory?”, “Can you make a good argument for the opposing ideas?”, … Like Rogerian client-centered therapy (Rogers, 1966), the benefit is not in the specificity of the questions, it is in the relationship of the mentor/therapist and the student/client and the ability of the relationship to foster self examination by the subject. The SimCoach project mentioned above found evidence of the establishment of a good relationship (ICT, 2015a).

While the initial effort would be directed to producing a small prototype of the system for test and evaluation purposes, if the system is selected for global distribution, the issue of speed of light/network latencies again becomes critical (Davis & Lucas, 2006). Past experience has shown that network latencies were on the order of 250-milliseconds for a 5,000-mile round trip circuit, see Figure 4 above. Human interaction can support something on the order of 500-milliseconds latency without being discomforted, which would suggest a global system would require on the order of half a dozen strategically placed computer clusters. Vessels at sea and other stations served only by satellite communications might require a pre-session downloading of the entire module onto local computers or servers, as satellite latencies are too disruptive and will not support the illusion of a *live* avatar.

Given the recent development of interactive virtual humans, there are many causes for concern in regards to the effectiveness of such a system. Those who are accustomed to personal interaction might have aversions towards technology, while digital natives are more likely to think that conversations with avatars can create the same trust and interest that human conversation do. Once the user grows accustomed to the new paradigm, the comfort level rises. The modern commodity of technology is a part of many people’s daily routine, and interaction with chatbots, *e.g.* Siri or Alexa, is commonplace. With this observation in place, the concept of interaction with virtual avatars is not so farfetched. In fact, several projects done at the Institute for Creative Technologies have found positive relationships between virtual avatar and user (ICT, 2015a).

The New Dimensions in Testimony (NDT) project developed an interactive storytelling avatar/agent after interviewing several Holocaust survivors. Students used the technology and evaluated their experience in surveys. The post-surveys showed that the system gave students a connection to the survivor, kept their attention, and had a positive impact, even to the point that users were observed apologizing to the agent on screen (Traum et al., 2015). The data gathered from the NDT experiments illustrate that interacting with a virtual avatar is as effective, if not more effective as interacting with a live survivor as shown in Table 1 below. Students reported a greater connection with the survivor and predicted that the experience would have an impact on them in the future. Though the live survivor more effectively retained the attention of the users, the system performed very effectively without a major decrease in attention. The interesting story given by the holocaust survivor did not suffer because it was displayed within a new medium, and the interaction proved to be relatable, intriguing, and significant.

Table 1 Percent of users rating the statements as “Strongly agree” or “Agree”. (Traum et al., 2015.)

	Live survivor (N = 28)	System (N = 25)
I felt that I could connect with the story of the survivor.	57	72
I felt that the activity kept my attention.	100	80
I think that my experience in this activity will have a positive impact on me.	86	92

Beyond NDT, the SimCoach project, also developed at ICT, hypothesized that users would be more willing to trust and interact on a personal level with a virtual avatar than with another human. The reason for this is that an interaction with an avatar seems to contain a lack of judgmental humans, whereas an interpersonal human conversation would suffer from the fear of judgment. After testing one interaction in which subjects were told that they were being observed and one in which they were not, the study demonstrated, “[Virtual humans] are able to have this impact because they allow patients to feel as though their responses are not currently being judged. Because the only difference between frames was the belief that another human was observing responses during the interview session, we can establish that the power of VH-interviewers to elicit more honest responding comes from the sense that no one is observing or judging" (Lucas, Gratch, King & Morency, 2014.). To this effect, it is likely that interactions with virtual avatars actually elicit a greater degree of trust than human interaction. Using the frame of a virtual avatar has been shown to be effective in increasing trust, all the while retaining an unaffected interest from the user. In the SimCoach project, the avatar retained interest more than an on-line live coach (ICT, 2015a).

Another concern regarding the effectiveness of virtual humans is that the lack of faultless automated speech recognition (ASR) renders successful virtual human interaction ineffective. The merit for this concern is that if a system does not understand a question, how can it provide an effective response? The answer to this question comes from the experiences at ICT. First, even without an entirely precise transcription of user utterances, a response classifier can still generate an adequately accurate reply. The classifier chooses the best matching response based on the input it receives compared to the bank of answers that it contains. As such, even an incomplete recognition of user utterances can provide enough context to determine a feasible answer. Second, even if the response isn't an exact match, as long as it is not completely off base, the answer will likely satisfy the user. It is unlikely that the voice recognition entirely misinterprets user input, and with even limited context, the classifier can provide appropriate feedback.

Because such classifiers determine the probabilities that each response matches the given input, choosing the response with the highest probability, the system will still determine a reasonable match based on the information it has. As such, even limited recognition results in somewhat suitable responses, which illustrate to the user that the system did, even if to a limited degree, identify the theme of the user's prompt. Even such small evidence gives the user the impression of comprehension by the virtual avatar. Finally, on top of the fact that effective classifiers can provide accurate responses despite limited or incorrect recognition, current ASR technologies are rapidly approaching an acceptably accurate response rate. A blog comment reported that Google's ASR was the most accurate, "Google achieved 73.3% of exact recognized phrases with a 15.8% [word error rate]" (Kudryavtsev, 2016). When combined with natural language processing technologies, all these processes continue to improve and are already at a point where they are quite useful. Considering all of these factors, natural language processing and virtual interaction are reaching levels wherein virtual agents can, to a degree sufficient to support conversation, comprehend and interact productively with live humans, who also have a word error rate.

A more closely analogous system would be the MentorPal project (Nye, et al., 2017) on which the authors are currently working. It is designed to provide mentoring to high school students considering science, technology, engineering and math (STEM) careers. Some early data supports the human quality of the virtual mentor presented, which supports the authors' contention that a conversational avatar, presented via video clips of live interviews, could sustain the *personal* relationship necessary for a critical thinking Socratic exchange. One of the on-going issues in the MentorPal project is the issue of creating a question bank for the mentor, as well as the existing answer bank. This capacity would be necessary for the critical thinking tutor as well.

OPERATIONAL VISION

The authors' vision of the operation of such a system is certain to evolve significantly as subject matter experts provide input and as development progresses. . Upon entry into the program, the user would be given a short briefing as to the goals and processes of the system, then invited to enter a significant range of personal data to help optimize the presentation of material to be covered. They would then take a short diagnostic test to assess their current levels of critical thinking and meta-cognition. These tests would be designed to establish skill levels in all of the germane areas of the two concepts. All of this could be accomplished by either text entry followed by a screen display of instructions and data or, it could be accomplished by voice recognition input and synthesized speech responses, with or without animated or video-taped avatars. The system would then present feedback and a proposed schedule of future instruction and evaluation, which would be available to the user, system administrators, and cognizant command personnel, as appropriate.

Thereafter, whenever the users log in, they would be presented with the continuation of the lesson, exercise or tutorial conversation, continuing from where it was left off. The system would present a brief recap of the previous actions and location in the pedagogical process, designing that presentation based on the time delay since last log-on. In the case of didactic material, it would be presented by displayed text or images, or by oral speech, or by oral speech with animated or video-taped avatars, as the users' personal preferences or operational constraints determined.

At the end of a fairly substantial introductory exposition of critical thinking concepts, the users would receive the first of a series of modules designed to highlight and improve individual skills needed for critical thinking enhancement. These modules would typically be comprised of academic materials, either written or oral, followed by an interactive situational simulation that was founded on and based in the matrix for scenario creation described above. After that was completed, the users would be evaluated and the next step would involve analysis and questioning by an avatar/agent, which would have the characteristics defined by the users or by higher authority. Should this process suggest the need at this time for additional work on this issue, a new set of instructional materials would be presented, along with a new scenario and evaluation.

This process would be repeated until all of the skills areas in both critical thinking and meta-cognition had been mastered to a level acceptable to the system monitors. The levels would be adequately identified so users could request additional or brush-up work in any area they or their command thought appropriate. When mastery levels indicated an over-all competency, the modules would begin to focus on a synthesis of all the skills in an operational setting, albeit a simulated one.

At each level, feedback to the users and to others in need of the data would be compared to other users with similar career positions. External evaluation marks from the user's service record could be input and compared with machine generated results from participation in this program and any discrepancy alarmed so command and system personnel could rationalize the dichotomy and make corrections or take action as required.

When nearing the end of a career phase, (US Navy: division officer and commanding officer or US Army/Marines: company and field grade), the users could be instructed to begin training at the next highest level. After the system gains some consensus of validity within the services, proficiency levels in these simulated tasks could become important input parameters for advancement. Using the aforementioned deep learning algorithms, the system would continuously seek out emerging nexuses that may suggest making improvements in simulated scenarios, analytic inferences, feedback methodologies, creative threat generation, and tutor/mentor interactions.

Given the granting of the required service-wide access permissions, which are often limited by privacy concerns, the system could further collect information on graduates and non-graduates and compare their performance, promotion, and retention. Such data is often considered to be valuable, but unless automatically collected and analyzed, the authors' experience is that it is rarely reviewed and heeded to support future decisions.

CONCLUSIONS

If one agrees that critical thinking is a valuable asset for the defense of the United States, then the DoD could easily infer that this area merits more attention. The authors maintain that many of the successful approaches to inculcating critical thinking in adults have been shown to do well if implemented by carefully trained and highly compelling trainers. The paucity of these highly qualified personnel assets and the ostensible disruptive nature of military operations schedules militate in favor of a computer-enable approach that would leverage and extend the reach of the most enthralling mentors and be more ubiquitously available. The authors' previous research demonstrates that virtual training works for this population, and should be effective for critical thinking training..

REFERENCES

Ariely, D. (2008). *Predictably Irrational* (p. 20). New York: HarperCollins.

Carter, A. G., Creedy, D. K., & Sidebotham, M. (2016). Efficacy of Teaching Methods used to Develop critical Thinking in Nursing and Midwifery Undergraduate Students: A Systematic Review of the Literature. *Nurse Education Today*, 40, 209-218.

Coram, R. (2002). *Boyd: The fighter pilot who changed the art of war*. New York, NY: Little, Brown.

Davis, D. & Lucas, R. (2006). Joint Experimentation, Data Management and Analysis Enabled by Trans-Continentally Distributed Linux Clusters, at *HPCMP Users' Group Conference*.

Davis, D., Lucas, R., Amburn, P. & Gottschalk, T., (2005), Joint Experimentation on Scalable Parallel Processors, *The ITEA Journal of Test and Evaluation.*

Davis, L. K., Curiel, J., & Davis, D. M., (2010), HITL and Metacognition: Self Analysis and Leadership Enhancement During Simulations, in the Proceedings of the *SISO Fall 2010Simulation Interoperability Workshop*,

Deng, L., & Yu, D. (2014). Deep Learning. *Signal Processing,* 7, 3-4.

Dunn, Rita, D., & Dunn, K. (1993). Learning styles/teaching styles: Should they.... can they... be matched. *Educational leadership.*

EduNote, (2016), *Under What Circumstances Oral Communication is More Effective*, retrieved from the internet on 28 November, 2016, from https://iedunote.com/effective-oral-communication

Flavell, J. H. (1971). First discussant's comments: What is memory development the development of?. *Human development*, 14(4), 272-278

Garrison, D. R., Anderson, T., & Archer, W. (2001). Critical thinking, Cognitive presence, and Computer Conferencing in Distance Education. *American Journal of Distance Education*, 15(1), 7-23.

Gottschalk, T. D., Yao, K-T., Wagenbreth, G. & Davis, D. M., (2010), Distributed and Interactive Simulations Operating at Large Scale for Transcontinental Experimentation, in the Proceedings of the IEEE/ACM *Distributed Simulations and Real Time Applications 2010 Conference*.

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring. *American Psychologist*, 53(4), 449.

Halpern, D. F. (2002). *Thought and Knowledge: An Introduction to critical thinking*. London,, U.K., Routledge.

Holmes, N. G., Wieman, C. E., & Bonn, D. A. (2015). Teaching Critical Thinking. Proceedings of the *National Academy of Sciences*, 112(36), 11199-11204.

ICT, (2015a), *Video of SimCoach in action*. USC Institute for Creative Technologies, Retrieved on 27 June 2015 from: https://www.youtube.com/watch?v=2bsMESwBeyg&index=14&list=PLBF277FAE78E8CB39

ICT, (2015b). *Selected Slides from presentation to the Army Research Laboratory*. USC Institute for Creative Technologies. Retrieved on 29Nov16 from: http://www.hpc-educ.org/AFIT-Init/Materials/Slides/VH-or-Voice-or-StaticImage.pdf

ICT, (2018). *Project One-Sheets*. USC Institute for Creative Technologies. Retrieved on 01 March 2018 from : http://ict.usc.edu/about/pdf-overviews/

Jean, G., (2006), "Game Branches Out Into Real Combat Training", *National Defense Magazine*, and Retrieved 02Dec16: http://www.nationaldefensemagazine.org/archive/2006/February/Pages/games_brance3042.aspx?

Khan Academy, (2016), *The Khan Academy*. Retrieved on 08 November 2016 from: https://www.khanacademy.org/

Kudryavtsev, A. (2016). *Automatic Speech Recognition Services Comparison*. Grid Designer's Blog. 11 Jan. 2016. Retrieved on 29 April 2017 from: http://blog-archive.griddynamics.com/2016/01/automatic-speech-recognition-services.html>.

Lai, E. R. (2011). Critical thinking: A literature review. *Pearson's Research Reports*, 6, 40-41.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436.

Leedom, D.K., McElroy, W., Shadrick, S.B., Lickteig C., Pokorny R. A. Haynes J. A. and Bell, J. (2007), *Cognitive Task Analysis of the Battalion Level Visualization Process*. Technical Report 1213, United States Army Research Institute for the Behavioral and Social Sciences. Fort Belvoir, Virginia.

Lehman, D. R., & Nisbett, R. E. (1990). A Longitudinal Study of the Effects of Undergraduate Training on Reasoning. *Developmental Psychology*, 26, 431-442.

Lucas, R.F., Davis, D.M., & Burns, D.P. (2015). System of Systems Complexity Addressed by Practical Adiabatic Quantum Computing. *The ITEA Journal of Test and Evaluation*, 36(4), 311-321.

Lucas, G., Gratch, J., King, A. and Morency, L.-P. (2014). It's Only a Computer: Virtual Humans Increase Willingness to Disclose. *Computers in Human Behavior*: *37* 94-100.

Nye, B., Swartout, W., Campbell, J., Krishnamachari, M., Kaimakis, N. and Davis, D. (2017). MentorPal: Interactive Virtual Mentors Based on Real -Life STEM Professionals . In the Proceedings of the *Interservice/Industry Simulation, Training and Education Conference*.

Oh, K. S., & Jung, K. (2004). GPU Implementation of Neural Networks. *Pattern Recognition*, 37(6), 1311-1314.

Pithers, R. T., & Soden, R. (2000). Critical Thinking in Education: A review. Educational Research, 42(3), 237

Rogers, C. R. (1966). Client-Centered Therapy (p. xi). Washington: American Psychological Association.

Reger, G. M., Rizzo, A. A., & Gahm, G. A. (2015). Initial Development and Dissemination of Virtual Reality Exposure Therapy for Combat-Related PTSD. In *Future Directions in Post-Traumatic Stress Disorder* (pp. 289-302). New York, NY: Springer.

Schulz, M., Sinayskiy, I., & Petruccione, F. (2015). An introduction to quantum machine learning. *Contemporary Physics*, 56(2), 172-185.

Snyder, L. G., & Snyder, M. J. (2008). Teaching Critical Thinking and Problem Solving Skills. *The Journal of Research in Business Education*, 50(2), 90.

Tetlock, P. E., & Gardner, D. (2016). *Superforecasting: The art and science of prediction*. New York: Random House.

Tiruneh, D. T., Verburgh, A., & Elen, J. (2014). Effectiveness of critical thinking instruction in higher education: A systematic review of intervention studies. *Higher Education Studies*, 4(1), 1.

Traum, D., Jones, A., Hays, K., Maio, H., Alexander, O., … Swartout, W. (2015). *New Dimensions in Testimony: Digitally Preserving a Holocaust Survivor's Interactive Storytelling*. 269-81. Springer International Publishing.

Von Lubitz, D. K. J. E., Carrasco, B., Levine, H., & Richir, S. (2004). Medical readiness in the context of operations other than war: Development of first responder readiness using OODA-Loop thinking and advanced distributed interactive simulation technology. In Proceedings EMISPHERE 2004 symposium on European-Mediterranean Virtual Hospital, Istanbul (Turkey).

Wagenbreth, G., Davis, D.M. & Lucas, R.F. (2010). GPGPU Programming Courses: Getting the Word Out to the Test and Evaluation Community. *ITEA Annual Technology Review*. Charleston, South Carolina

Willingham, D. T. (2007). critical thinking. *American Educator*, 31(3), 8-19.

Wong, W. L., Shen, C., Nocera, L., Carriazo, E., Tang, F., Bugga, S., ... & Ritterfeld, U. (2007, June). Serious video game effectiveness. In Proceedings of the international conference on *Advances in Computer Entertainment Technology* (pp. 49-55). ACM.

Yao, K-T., Davis, D. M., Liu, J. J., & Kaimakis, N. J.. (2018, Pending). New Technologies to Enhance Computer Generated Interactive Virtual Humans. In the Proceedings of the *SISO Fall Simulation Innovation Workshop*. (Note: Accepted in 2017, SIW conference was cancelled due to hurricane Irma: paper available on-line at - http://www.hpc-educ.org/Papers/2017/SIW/QuantumVH-SIW-Paper-Accepted.pdf)