

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 DoD 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 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 avatars 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 continuing availability of this training for DoD personnel. A short review is offered of the need for critical thinking. A brief survey of effective critical thinking pedagogies is followed by an analysis of which would be most effectively adapted to the distributed computing avatar 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 avatars. These avatars are capable of carrying on a life-like verbal exchange with the users. Examples of where and how both the distributed computing and avatar 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 avatars 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 given. The paper closes with a description of how such a system could effectively serve the globally-dispersed and time-constrained DoD personnel.

ABOUT THE AUTHORS

Dan M. Davis is a consultant for the University of Southern California, focusing on large-scale distributed DoD simulations. 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 at the Jet Propulsion Laboratory and Martin Marietta. He has served as the Chairman of the Coalition of Academic Supercomputing Centers and has taught at the undergraduate and graduate levels. As early as 1971, Dan was writing programs in FORTRAN on one of Seymour Cray's CDC 6500's. He saw duty in Vietnam as a USMC Cryptologist and retired as a Commander, Cryptologic Specialty, U.S.N.R. He received B.A. and J.D. degrees from the University of Colorado in Boulder.

Nicholas Kaimakis has been active in research at the Institute for Creative Technologies of the University of Southern California. His current research thrusts are in the use of computer generated avatars or video clips, animated and directed by natural language optimized Artificial Intelligence (A/I) programs that present a life-like dialogue capability to interact with remote users via the internet. He has demonstrable success designing efficient project structures, teaching programming, and managing multi-faceted teams. His current project is funded by the Navy and is designed to help improve knowledge of STEM fields across varied demographics with the development of an interactive interface that makes STEM information more accessible on-line. He is an undergraduate student studying Computer Science (CS) within the Viterbi School of Engineering at the University of Southern California.

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 during the development of a highly technical pro-

ject 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 has been at the forefront of career path decisions of other personnel in both the military and civilian sectors. 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 advance three theses: 1) Critical thinking is vital to a strong defense posture; 2) Critical thinking can best be inculcated with skill training; 3) Critical thinking can optimally be implemented with globally accessible computer-aided instruction, especially in the light of current staffing and funding constraints. While evidence will be adduced to support all three theses, it is acknowledged that a short paper is not the proper vehicle for an irrefutable proof of any of these theses. 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. This will be followed by examples from history emphasizing the necessity for critical thinking of leaders in the past and present. These issues have currency in today's daunting defense issues that mandate redress by leaders equipped with and fortified by critical thinking skills. The authors take the position that this would optimally be a training function. The barrier to the adoption of these concepts in 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 high-lighted. There are sections on the vision for implementation of such a system and on the envisioned users' experience. The conclusion will feature the path forward and the effort required for implementation.

Terminology

As used in this paper, the term critical thinking will refer to that 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 analyzed.

While the dichotomy between the fields of training and education is subject to discussion, 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, *et al.*, 2010)

BACKGROUND OF CRITICAL THINKING IN DEFENSE

Many assert that concern for critical thinking today is central to military effectiveness (Guillot, 2004). While it took the Greeks generations to develop the Phalanx (Hanson, 2009) and even longer to develop the ships used at Salamis (Williams, 1968), today's DoD officer must recognize the need for new defensive tactics, then conceive, develop and field effective measures within a short time window of months or years. While the cited Grecian technologies were effective weapons for millennia, today's weapons may succumb to obsolescent in a matter of weeks due to counter-measures or societal constraints.

Nearly two thousand years after the Grecian tactics and technologies were developed, great masses of Phalanx-like troop formations, now armed with muskets, fought each other at Waterloo using the Napoleonic Infantry Square (Chandler, 2009). The carnage increased fifty years later, due to improved technologies used in the American Civil war. Late in that war, one Major with innate critical thinking skills pointed out that skirmishers, fighting in dispersed formations, gained objectives with far fewer casualties than those suffered by massed formations during frontal assaults. This insight was taken up and championed by Brigadier John Watts de Peyster (Randolph, 1905) However, that lesson to disperse was not effectively implemented until after the incredible slaughter in the First World War. But even this dramatic lesson did not cause a general examination of the evaluative process. The British, among others, had gone into that War relying on tactics they had found useful in their colonial wars. (Scales, 1976) A similar analysis would obtain a scant two decades later as manifest in the reluctance to accept change by the command structures of both of the great navies in the 1940s' Pacific war. The Japanese clung to the focus on large capital surface ship tactics despite the dominance of carrier task force capabilities until surface ship losses and carrier operations successes forced their acceptance of the new weapons systems. (Durham, 1998)

While the U.S. did develop more economical approaches in the use of their personnel, other nations persisted in the use of tactics that led to unimaginable slaughter of their own people, often without the desired result. The losses of the Russians at the battles of Stalingrad and Leningrad (Beever, 2006), as well as the Japanese sacrifices on Iwo Jima (Bradley, 2000), stand as stark monuments to such horrendous expenditures of life.

After that war, the U.S. sought to offset the numerical superiority of the USSR with the destructive power of nuclear weapons and their intercontinental delivery using bombers and then missiles. A second offset was effective by the U.S. use of advanced technology to offset the reduced size of its defense forces in the post-Vietnam era. A third offset is now sought to defray the dangers posed by a new set of foes that are often asymmetric in their threat posture and by nation states that have their own technology capabilities. Like previous changes, this may drive the need for Critical Thinking deeper down into the command structure, making it necessary for even junior personnel.

However, one lesson has been learned, over and over; a nation is at peril if it ignores the invaluable contribution of the first line of leadership, the ones who face the enemy on a personal basis. So fluid is the current situation that rote training, doctrinal imperatives, and reliance on blind obedience do not well serve either the individual or the nation. But, the past is past; the situation today requires a new vision.

More than one military leader has probably asked, "How do we develop thinking skills in tomorrow's leaders?" The answer has been stated in many ways. (Guillot, 2004) One senior enlisted warfighter defined it as: Quite simply stated, it is to encourage questions regarding specific instances and then use a bevy of "What if?" scenarios to clarify the issue under contemplation." Starting this process early in a military career is essential. The authors feel that there are two main areas of critical thinking; context free and context dependent. Context dependent concerns only a specific circumstances and is usually the one with which those who are actively engaged will concern themselves.

Context free covers a wide range of possibilities, and it concerns itself with the overall picture. As service personnel advance in pay grade, responsibility, and knowledge, they will, and should, transform to context free critical thinking which will involve many different non specific subjects. In previous eras, Officers were mostly college educated while the majority of Enlisted Personnel were seldom educated beyond the high school level.

Beginning somewhere in the 1980's, this education differential was no longer the case. Educational levels increased virtually across the entire military sector. Jobs that once required mid level (O-4/5) Officer supervision were now being handled by college educated Senior Enlisted personnel, which meant that their former tasks were now being performed by the next lower rank junior personnel. In some of the technical fields, detachments previously led by junior officers were now being led by senior sergeants and chief petty officers. Their billets of Watch Section Chief and similar position were being filled by E-6 level non-commissioned and petty officers. At the same time, the pressure on mid-grade enlisted personnel to earn associates and bachelor's degrees was increasing. Military mentors are known to be advising senior enlisted personnel to seek at least an associate degree to document educational achievements in order to be competitive for advancement to E-7, 8 or 9. In many fields it is not unusual to find NCO's and CPO's with bachelors, and even post-graduate degrees are increasingly appearing in these ranks. A DoD study from 2015 found that nearly seven percent of enlisted personnel had bachelor's degrees and one percent had post-graduate degrees. In a trend analysis, it was found that the number of bachelor's degrees amongst the enlisted ranks had nearly doubled in a little more than a decade (DoD, 2015).

One of the benefits of increasing the process of gradually becoming proficient in the critical thinking process is that those individuals may often have intimate knowledge about a certain task than those senior to them. Anecdotally, the authors are familiar with one Chief Petty Officer who was assigned as the Assistant Officer In Charge to develop a new Battle Group Communications system which was loosely based on an existing Air Force program. During sea trials there were numerous problems retaining communications continuity with the airborne platform. After all of the computer programmers were unable to resolve the issue, the CPO and his operators were asked for input. The Navy personnel inquired of the developers if they had reprogrammed the positioning software to take into account the mobility of a Navy ship. The contractors had not realized the criticality of that. It was pointed out that an Air Force base is a fixed point on the globe; U.S. Navy Ships are not stationary. Upon investigation, it was found that the system had been programmed with the coordinates for Pier 12 in Norfolk, VA as the default, when actually the ship was off the coast of South America. Such critical thinking on the part of that CPO and his team were not only expected of them, it was delivered when needed. This Navy team was made of intelligence ratings, so were trained in and practiced critical thinking as part of their profession.

In considering the need for critical thinking today, the above observations should be coupled with the fact that many senior general and flag officers are now briefed by senior enlisted personnel. These professionals need both the communications skills provided by higher education and they need to possess the critical thinking skills to be able to respond to fundamental and probing questions with their own germane and rational answers. It is the authors' opinion that critical thinking is now vital, and will be increasingly so in the future. It can and must be learned.

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 are literally volumes that have been written on the plethora of approaches to defining, enumerating, and teaching Critical Thinking. (Lai, 2011) Most of these analyses are not set in a military or defense environment, so, while illuminating, they should not be considered controlling. Given that *caveat*, there may be a preliminary consensus on the steps of the process from which a working model can be extracted. These may be something on the order of:

- | | |
|-------------------------------|--|
| 1. Collecting data | 4. Reviewing initial conclusions |
| 2. Categorizing and analyzing | 5. Combining ideas and expanding uses |
| 3. Using the insights gained | 6. 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. The behavioral economists have done much to expose the human frailties along this line. (Kahneman, 2011 and Ariely, 2008)

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

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

Note that most of these imply or require the presence of a tutor or lecturer and some of the earlier cited studies on program efficacy note the skill and personality of the instructor play a major role in the success of the educational process. Transferring that process into the active-duty or reserve environment for defense personnel suffers from two major debilities: lack of sufficient qualified and engaging instructors as well as inaccessibility of dispersed personnel, all of whom are likely to have immutable ops-schedule constraints. In these days of restricted funding, the idea of dramatically increasing the number of instructors or swelling the ranks to ease scheduling limitations are putatively infeasible.

Further, it has been observed that much of the enthusiasm students have for a particular topic, hence its impact on them, is based on their relationship to the instructor. While there is some commonality as to preferences for instructors, there is also some variability; even the most popular professors have some students with whom they are not well matched. At least in major universities where there are classes with several sections, students have the option of finding a professor they like, but the DoD rarely affords this choice to its personnel.

A COMPUTER-AIDED TRAINING APPROACH

Many of the inhibitions mentioned above can be successfully overcome by Computer Aided Education. New advances in high performance computing, globally distributed computer communications, animated or video-captured avatars, natural language processing, and Human-Computer Interaction technologies can effectively address all of these issues. 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.

A large portion of human 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 the Institute of Creative Technologies have fielded very engaging and lifelike conversation computer interfaces. (Reger, *et al.*, 2015) These make use of a wide range of computer capabilities to: input 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). At ICT, three types of display are used to best communicate with different target audiences, see Figures 1, 2, and 3.



Figure 3 - Fully Animated SimCoach CGI with alternative characters



Figure 2 - Video-taped holocaust victim in Holographic classroom display



Figure 1 - Video-taped Interview Head and Shoulders on monitor

If an animated avatar is used, the appearance of the avatar can easily be altered to present different ages, genders, ethnic appearance and social/military status. Once this capability is prepared, the avatar can be available to engage in a “one-on-one” conversation with as many users as the servers can support, 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 24 by 7 availability and immunization against disruption due to intruding operational requirements addresses the above noted operational constraints.

Spawning new avatars requires very little marginal costs, mostly additional computer processing power and storage space. Careful and easily modifiable design characteristics of the avatar 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 web sites like the Khan Academy (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 avatar. 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 and be globally accessible any time. The authors have supported such efforts in the past and recognize both its potential and its hurdles *e.g.* speed of light induced latencies when server and client are separated on the order of five thousand miles. (Gottschalk *et al.*, 2010) While this system was designed to advance analysis and evaluation for the Joint Experimentation, 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, 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.

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 below 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.

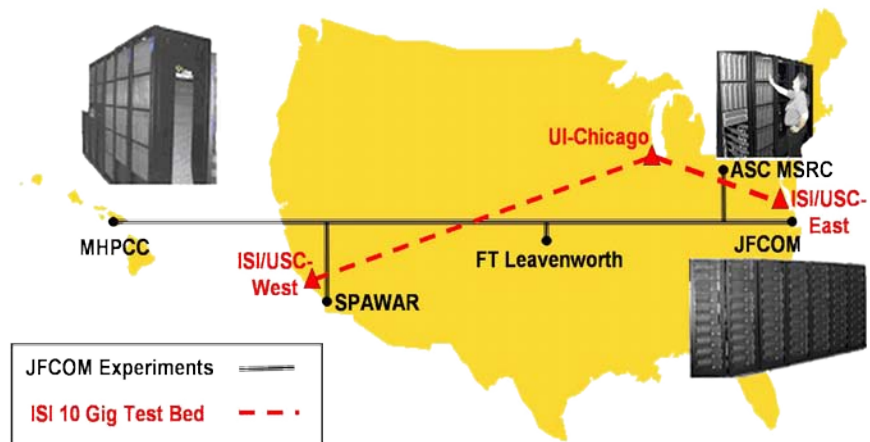


Figure 4 - USC-enabled Trans-continental Network Testbed.
Maui to Virginia round trip latencies: 204 msec

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 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 situation mandate.

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 America's Army is an example of how a product designed with an attractive interface and interactive structure could rapidly show its mettle via measurable performance improvement in training levels. (Jean, 2006)

The thinking skills taught would make a difference in the way personnel analyze data and take action to improve operations. In a different setting, 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, 2015)

Implementation Vision

While there are a wide range of implementation approaches that would satisfy the primary goal of this initiative, the following approach will give some clarity to a future implementation of a new globally accessible critical thinking 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 (Garrison, 2001), but the nature of the defense issues requires that these early efforts be assessed by a panel of senior DoD personnel. 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)

Selection of the particular pedagogical approach to enhancing critical thinking skills would also require careful consultation with experienced and proven DoD leaders. Depending on time and funding constraints, several approaches could be analyzed and the best ones implemented for dissemination on a global basis. Different learners show a range of amenability to a similarly large range of teaching styles (Dunn & Dunn, 1993). Even though these 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, even, as the Dunns suggest, with the use of a single teacher or avatar.

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 alter-



Figure 5 -A Present Day Setting in Indonesia using JSAF

natives to personnel with differing personal enthusiasms including a “Present Day” or future settings.

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.

The last dimension is the most open to discussion. As the program focuses on skills required for critical thinking, these may be parsed out in ways that would suggest or require individual modules to isolate issues and enable quantification of improvement. This is a research issue worthy of interest. In any case, the final parametric would surely be the ability to synthesize all of the skills effectively to optimize critical thinking. Perhaps a good initial start would be a dimension with loci of emphasis such as: data collection, analysis, decision, action and synthesis

Following the development of the training and the scenarios, a major feature of this initiative would be the provision of on-going mentorship by avatars. Whether considered mentors or Socratic tutors, these avatars 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. Considering the eight by nine by five by two (historical period, command level, skill issues, and ground or sea) matrix there are 720 potential programs. Fortunately, many of the appropriate questions will be the same across most of the matrix. Early research and proof of concept efforts would likely find it propitious to restrict work on just a few historical periods and command levels. Filling out the rest of the matrix would be relatively *pro forma*, with the exception of the creative work on the historical period scenarios (a good project of the Service Academy student bodies?).

Both Socratic teaching and meaningful mentorship involve more question-asking than direction-giving, so the questions must have a set of cogent follow-on comments prepared for a range of responses. This is where the capabilities in natural language processing and deep learning algorithms come into play. Not only must the questions and comments come quickly to sustain the illusion of “humanness” of the avatar, the insights offered must be useful. Naturally, there must be a series of good “break out” responses, if frustration, boredom or antagonism is detected in the user. The user would then be guided to an appropriate live “Master Tutor” for further action.

The system would be well able to carefully monitor the initial skill-level of the user, assess their progress, correlate advancement with program content, modify approach or pacing, and report quantified assessments of capabilities for consideration later. A serious ethical issue arises as to whether measures of capability should remain private, be available to current command personnel, or be used as a reporting item on evaluations and fitness reports. The system could easily be constructed to support whatever decision is made in this realm.

As a model for what the projected implementation of what a “conversational avatar” critical thinking interaction may look like, the authors recount their experiences with the staff at Next Generation Leaders, 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, *et al.*, 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?”, *etc.* 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 Sim-Coach project mentioned above found evidence of the establishment of a good relationship. (ICT, 2015)

While the initial effort would be directed to producing a small portion of the system for test and evaluation purposes, should the system be selected for global distribution, the issue of speed of light/network latencies again becomes critical (Davis, 2006). Past experience has shown that network latencies were on the order of 250 milliseconds for a five thousand 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 avatars, there are many causes for concern in regards to the effectiveness of such a system. Within a society accustomed to personal interaction, despite its aversions towards technology, particularly in the younger generations, it is old fashioned to hypothesize that conversation with avatars cannot possibly create the same trust and interest that human conversation contains. This claim does hold some merit, and it is often the case during the initial interaction of new technologies that the user feels out of place and thrown off by the experience, like the first time using a cell phone, or typing on a computer. However, once the user grows accustomed to the addition, 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 Google Voice, 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.

The New Dimensions in Testimony (NDT) project developed an interactive storytelling avatar 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” (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. Some users wept and some offered condolences and apologies to the avatar.

Table 1 Percent of students 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 contains no human judgment, 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, G. *et al.*). 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 a live coach. (ICT, 2015)

Another concern regarding the effectiveness of virtual humans comes in the misconception 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 in three parts. 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, 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 study done by Andrew Kudryavtsev showed 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, 2017) on which the authors are currently working. It is designed to provide mentoring to high school students considering STEM (Science, Technology, Engineering and Math) careers. Some early data supports the human quality of the 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. The Likert chart shows that, even at this small a sample size, the overwhelming impression on the students was one of openness to the two avatars presented and a subjective assessment that their information was valuable and reliable.

Table 2 – Likert Chart of Student Responses

Results from Users, N=16					
<i>Please Circle the number that best describes how you feel.</i>	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1. I think or worry about my future life every day.	1	2	2	11	
2. Having easy access to better advice would be useful.			2	10	4
3. I liked this on-line mentor and learned from him.		1	3	6	6
4. The answers were on-target and useful.		1	3	6	6
5. The responses seemed "conversational" and real.		1		7	8

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. Nevertheless, it may be useful to outline the current vision as it is perceived in the first quarter of 2017. 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, 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, (Navy DivOff and CO or Army/Marines Company and Field Grade), the users could be instructed to begin training at the next highest level. After the system gains some sense 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, feed-back methodologies, creative threat generation, and tutor/mentor interactions.

Deep learning (LeCun, 2015), an advanced method of machine learning of big data, is currently used effectively in image recognition and natural language processing. The use of Graphics Processing Units (GPUs) 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 *et al.*, 2010) The combination of this technology with emerging quantum computing capabilities (Lucas *et al.*, 2015) is also an intriguing new area of research.

Given the required service-wide access permissions, the system could further collect information on graduates and non-graduates and compare their performance, promotion, and retention. Such data is always needed, but unless automatically collected and analyzed, it is rarely reviewed and heeded to support future decisions.

CONCLUSIONS

In order to meet the dynamic threats posed to the U.S. in the 21st Century and still honor the constraints of societal norms and financial limits, the DoD must make optimal use of the strengths of the nation that are not common in other cultures. Among these are a highly educated and motivated intellectual elite that has an unparalleled set of infrastructure tools available to teach improved analytic and decision approaches to military personnel at every level and in every environment. The authors maintain that critical thinking is vital, can be inculcated by training and is most practicably implemented by distributed computer-aided instruction and skill maintenance. Naturally, the choice between the varying approaches would have to be assessed using standard System Engineering methodologies.

REFERENCES

- Ariely, D. (2008). *Predictably Irrational* (p. 20). New York: HarperCollins.
- Beevor, A. (2012). *The Second World War*. Hachette UK. p203.
- Bradley, J., & Powers, R. (2000). *Flags of Our Fathers*. London: Pimlico.

- 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.
- Chandler, D. G. (2009). *The Campaigns of Napoleon*. Simon and Schuster.
- Coram, R. (2002). Boyd: *The fighter pilot who changed the art of war*. Little, Brown.
- Davis, D.M. & Lucas, R. F. (2006). Joint Experimentation, Data Management and Analysis Enabled by Trans-Continentially Distributed Linux Clusters, at *HPCMP Users' Group Conference*, Denver, Colorado
- 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 2010 Simulation Interoperability Workshop*,
- Davis, D., Lucas, R., Amburn, P. & Gottschalk, T., (2005), "Joint Experimentation on Scalable Parallel Processors," *The ITEA Journal of Test and Evaluation*, Fairfax, Virginia, 2005
- Deng, L., & Yu, D. (2014). Deep Learning. *Signal Processing*, 7, 3-4.
- DoD, (2015). *2014 Demographics: Profile of the Military Community*. Office of the Deputy Assistant Secretary of Defense for Military Community and Family Policy (ODASD), Washington DC. 39-40
- Dunn, Rita, D., & Dunn, K. (1993). Learning styles/teaching styles: Should they.... can they... be matched. *Educational leadership*.
- Durham, R. W. (1998). *Operational Art in the Conduct of Naval Operations*. Monograph submitted to School of Advanced Military Studies United States Army Command and General Staff College, Fort Leavenworth,
- 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>
- Hanson, V. D. (2009). *The Western Way of War: Infantry Battle in Classical Greece*. Univ of California Press.
- Halpern, D. F. (2002). *Thought and Knowledge: An Introduction to Critical Thinking*. 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.
- 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*, Fairfax, Virginia
- Guillot, C. W. M. (2004). Critical Thinking for the Military Professional. *Air & Space Power Journal—Chronicles Online Journal*, 17.
- Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring. *American Psychologist*, 53(4), 449.
- ICT, (2015). *Selected Slides from presentation to the Army Research Laboratory*. <http://www.hpc-educ.org/AFIT-Init/Materials/Slides/VH-or-Voice-or-StaticImage.pdf>
- 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?
- ICT, (2015a), *Video of SimCoach in action*. Retrieved from the internet 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*. Retrieved on 29Nov16 from the internet at: <http://www.hpc-educ.org/AFIT-Init/Materials/Slides/VH-or-Voice-or-StaticImage.pdf>
- Kahneman, D. (2011). *Thinking, Fast and Slow*. Macmillan.
- Khan Academy, (2016), *The Khan Academy*. Retrieved from the web on 08 November 2016 at: <https://www.khanacademy.org/>

- Kudryavtsev, Andrey. (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.
- 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, Gale M., 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*, Orlando, Florida, 2017.
- 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
- Randolph, H. L. (1905). *Biographical Sketches of Distinguished Officers of the Army and DoD*. Page 87. Retrieved on 02Dec16 from: <https://books.google.com/books?id=4iIZx9QohXIC&printsec>
- 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). Springer US.
- Scales, R.H. Jr., (1976) *Artillery in Small Wars: The Evolution of British Artillery Doctrine, 1860-1914*. *Duke University PhD dissertation*, 1976.
- Snyder, L. G., & Snyder, M. J. (2008). Teaching Critical Thinking and Problem Solving Skills. *The Journal of Research in Business Education*, 50(2), 90.
- 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, David, Andrew Jones, Kia Hays, Heather Maio, *et al.* (2015). *New Dimensions in Testimony: Digitally Preserving a Holocaust Survivor's Interactive Storytelling*. 269-81. Springer International Publishing, 2015.
- von Lubitz, D. K. J. E., *et al.* "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." 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
- Williams, R. T., & Morrison, J. S. (1968). *Greek Oared Ships: 900-322 BC*, University Press.
- 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.