Methods and Tools for the Development of Effective Training Games

Talib Hussain and Wallace Feurzeig
BBN Technologies, Cambridge, MA 02138

Abstract
BBN has developed several successful game-based training systems for the military. From these and other development efforts, we have identified a number of capability gaps in current methods and tools for developing effective game-based training applications. Our observations and conclusions are drawn from research studies and reviews of games and game-based training systems, as well as from our own experience in developing and fielding several such systems. We discuss lessons learned and additional enhancements in current methods and tools needed to advance the instructional utility and power of game-based training technology.

Motivation
There is little doubt that new and emerging technologies will have an impact on how instructional systems are developed and deployed over the next several decades. In this regard, the area of educational games is receiving growing attention due to the perceived motivational power of video and computer games produced for entertainment. Modern computer gaming technology has created highly entertaining games that provide immersive simulation environments enabling rich interactions among distributed players and facilitating the development of strategic thinking and decision-making skills.

Studies of the effectiveness of game-based training have been conducted for a variety of domains from single-player scenarios to those involving large teams. These studies confirm that game-based technology shows great promise for delivering compelling experiential training and for effective assessment of cognitive skills, particularly in stressful situations that call for timely judgment and decision-making. Games have promise for evaluating the impact of psychological traits on leadership, for evaluating the ability of a sailor or marine to adapt to new operating conditions and protocols, and for evaluating the ability of combat units to follow new rules of engagement using new equipment in a hostile environment. However, without the necessary research base to inform instructional design and to guide the development of appropriate training and assessment methods and tools, sizeable investments will likely be made on game-based training systems that are ineffective or marginal.

Systematic research is needed for guiding the design, development, and effective use of games that have serious educational goals and for demonstrating their effectiveness in facilitating learning. We review the current state of the art, identify capability gaps and present some thoughts on how to address those gaps. These represent our initial efforts under an ONR-sponsored project, “Tools for Game-Based Training and Assessment of Human Performance”, dedicated to the development of the learning science, instructional methods, and technological tools to respond to these challenges.

Background
A number of current efforts have explored the use of modern games for training and assessment, and several games have been adapted expressly for military purposes (Bock and Dennen, 2005; O'Neil and Painess, 2005; Weil, Hussain, et al, 2005). These efforts have shown that that gaming technology has many potential benefits, including:

- Games can effectively be used for training individual and team-based military tactical skills, as well as for training cultural and other "non-kinetic" skills.
- Games can represent realistic military missions with a high level of face validity, as demonstrated in America's Army. "In live training, a unit could perform perhaps three evolutions 'on a good day'… On computer, that unit can do up to 40 evolutions, honing skills through repetition and feedback." (Minton, 2005).
- Creating a military training game can be relatively low-cost and require modest development time. DARWARS Ambush! was developed in six months.
- Appropriately designed authoring tools can facilitate rapid development of new content. Ambush! users frequently create new training content in the field.
- Games can produce significant training benefits for teams as well as individuals. Gorman's Gambit, demonstrated the capability of using a COTS multi-player game to support military teamwork training at the platoon level (Hussain & Ferguson, 2005).
- Commercial game engines can be leveraged for rapid development.

Some recent successes in the areas of military training include:

<u>DARWARS Ambush!</u> is a multi-player, game-based training system for convoy operations, developed for DARPA in 2004 and currently in use by our warfighters at numerous bases in the U.S. and abroad. It was designed originally to reinforce the practical skills and TTPs (tactics, techniques and procedures) needed to anticipate and react to convoy ambushes and improvised explosive devices. However, its users have extended its application to address leadership skills, rules of engagement, dismounted urban operations, and many types of missions. Ambush! has been scaled to over 40 simultaneous users. BBN has provided ongoing support for over four years and has helped grow a community of users and trainers. Ambush! enables soldiers and Marines to experience lessons that others have learned and to construct scenarios based on their own experiences. Individual trainees move about in a shared, immersive, first-person-perspective environment to carry out mounted and dismounted operations, operate ground and air vehicles, use small arms and vehicle-mounted weapons, and communicate over multiple radio nets. They learn to anticipate and respond to ambush situations, practice existing tactics techniques and procedures and experiment with new ones.

Ambush! is based on a commercial game engine that has been adapted to suit military training needs. It provides a variety of ways to train in a multi-player environment, with support for synthetic team members and opposing forces. Participants can assume the role of insurgents or observer/ controllers during a mission. Ambush! improves readiness by exposing deploying troops repeatedly to recent lessons learned by in-theater forces. Instructors can author scenarios in the field and conduct After-Action Reviews (AARs) of training sessions using built-in, easy-to-use debriefing support tools.

The Ambush! AAR tool provides enhanced assessment capabilities to support a human trainer in performing an AAR. Ambush! captures such low-level game data as team member locations and orientations, shots fired, communications events, and changes in health and vehicle occupancy. These data can then be used in the AAR tool to analyze, for example, whether gunners were scanning their sectors of fire, or how soon after receiving orders on a comm net participants took action. The Ambush! AAR tool uses BBN speech recognition software, BBNTalk™, a Voice-Over-IP technology that enables users to communicate across an arbitrary number of channels. Each voice communication is automatically transcribed, and the transcription is available in the event list. Instructors can mark moments in time or collections of events as exemplary or problematic, for convenience when conducting the AAR with trainees. Video capture of the game-play can also be displayed alongside the AAR tool's top-down map-based mission playback. Ambush! is currently in active use within the U.S. Army, Air Force and Marines providing team-based training to thousands of warfighters every year. Over 30,000 warfighters have undergone this training, and adoption of the product is increasing dramatically. PEO STRI is actively transitioning it throughout the Army. (Diller et al., 2004; Roberts et al., 2006).

<u>Gorman's Gambit</u>. The BBN Gorman's Gambit project investigated the capability of existing multi-player games to support military training, and identify resources that would need to be developed to increase the training effectiveness of such applications. We conducted an exercise using an existing Commercial Off-the-Shelf (COTS) game. A scenario was designed to support behaviors indicative of effective teamwork. In the "capture the flag" game scenario, two teams of twenty U.S. Army infantry at Ft.Benning, GA played against each other in three thirty minutes sessions while trained observers monitored performance. The attributes of the in-game characters controlled by each soldier were manipulated to encourage coordination among participants. Trained observers monitored performance in three thirty minute sessions, and recorded instances of teamwork using a template based on the existing team performance literature. Results indicate that the multi-player game genre may provide a functional model for emulating distributed multi-user military training with hierarchical team structures. The results highlight several inherent challenges with using COTS gaming systems for assessment-intensive applications such as military teamwork training. (Hussain & Ferguson, 2005).

<u>SABRE Game-Based Testbed for Culture and Personality Research</u>. Under support of the Defense Modeling and Simulation Office and the Air Force Research Laboratory, BBN developed a game-based testbed for evaluating how performance in militarily relevant behaviors, such as teamwork skills, situation awareness, decision-making, and social interaction are affected by individual differences such as personality or cultural traits. The testbed includes personality and cultural assessment instrument and supports complex task scenarios, featuring challenges such as information overload, unexpected failures, and interpersonal conflicts in an immersive, role-play, computer game environment. SABRE employs a number of assessment measures for automated capture within a game environment, e.g., measuring information push or pull among participants,

inferring situational awareness from player activities, and measuring task effectiveness. The project studies the effects of cultural factors and personality differences in team performance. Of particular interest are the sorts of teamwork skills needed for military operations other than war, e.g., when personnel work in a foreign culture and in multi-national coalitions. SABRE allows behavior researchers to customize and extend experimental scenarios themselves, with minimal need for software developers. It is free to researchers and government personnel. The system is currently in use in US and international experiments through the NATO Leader and Team Adaptability in Multinational Coalitions project. (Diller. Ferguson, et al, 2004; Warren et al, 2004).

T-CAST Team Coaching Assistant for Simulation-Based Training In the T-CAST project, BBN developed the Spatial Analysis Engine (SAE) to analyze team performance data extracted from simulations in team-work training for dismounted infantry teams in a game-based training environment. SAE monitors the execution of an on-line, multiplayer, training simulation, recognizes the occurrence of certain situations caused by participant behavior, and reports on those situations (e.g., recognizing that warfighters did not maintain appropriate formations or performed an incomplete search in a house clearing operation). This analysis is designed to help trainers both during the training session itself as well as during an AAR.

Tactical Iraqi. The Tactical Language Training System (Tactical Iraqi) is an initiative funded by DARPA and ONR, aimed at developing a set of active language learning tools, with a concentration on Arabic language and culture. The system, created by the USC Viterbi School of Engineering's Center for Research in Technology for Education (CARTE) and partners, engages students in videogame-based missions in animated virtual environments within which students have to successfully phrase questions and understand answers in Arabic. Lewis Johnson leads the CARTE team. Part of the system, the Mission Skill Builder, resembles an intensive version of the language laboratory programs that have been in use for generations, in which students imitate and practice words and phrases pronounced by native speakers. However, it incorporates some important innovations. Speech recognition technology tailored for language learner speech, is capable of evaluating learner speech and detecting common errors. Pedagogical agent technology provides users with tailored feedback on their performance. A learner model dynamically keeps track of those aspects of the language the user has mastered and those in which the user is deficient.

The program also instructs students in non-linguistic cultural skills that are important for communication, including social skills necessary to build rapport with local people, appropriate degrees of politeness to use in different social situations, and knowledge of how to disagree without offending and how to respond to offers of hospitality. Gesture training includes common Arabic gestures that a Westerner might misinterpret and American gestures (such as thumbs-up) that an Arab might misinterpret. The examination component of the system is designed to provide an unscripted, unpredictable, and challenging test of a student's mastery of these elements. Speaking Arabic, students try to talk their way through encounters with multiple characters and tricky situations in the course of carrying out a military task. (Johnson and Beale, 2005), (Johnson et al, 2007).

ELECT BiLAT is a prototype game-based military simulation for practicing bilateral engagements and negotiations in a cultural context. It provides students with experience in preparing for a meeting including familiarization with the cultural context, gathering intelligence, conducting the meeting, negotiating when possible, and following up on meeting agreements as appropriate. The architecture is based on a commercial game engine integrated with technologies that enable the use of virtual human characters, and scenario customization, as well as coaching, feedback, and tutoring. Although it is too early to report on the training effectiveness of the prototype, a number of aspects of this project are noteworthy. ELECT BiLAT addresses a non-kinetic training domain that is relevant to the contemporary operating environment. It employs customized game-play mechanics to provide an immersive and interactive experience, while making extensive use of story-based scenarios. The system leverages virtual game technologies to support social interaction, and incorporates intelligent tutoring to enhance learning. It is designed to enable rapid scenario development and modification by end-users. (Hill et al, 2006.)

ATL (Adaptive Thinking and Leadership) is a multiplayer simulation game for intercultural communication training launched in December 2004. It was designed to serve as a non-violent virtual sandbox for multiple players to practice communication skills, develop mental agility, and hone cultural awareness acumen. The development team comprised Sandia National Laboratories, the Army Game Project Government Applications Team, and the University of Central Florida. ATL was designed to support classroom training. In-game assessment is performed by peers and instructors who observe play. AAR facilities are available to convey the outcomes to trainees. Trainees often role-play as people from different cultures. They are given a background story and a set of objectives, with the goal of understanding different cultural worldviews. Role-playing is a well-developed technique in the cross-cultural training literature. However, the use of humans as role players can become both cost prohibitive and challenging to control. Massively Multiplayer Online Role Playing Gaming support is among a number of options under consideration for future phases of the project. The game is part of the curriculum for a course in Adaptive Thinking and Leadership in the Army JFK Special Warfare School. The Special Forces school has plans to incorporate the game in Civil Affairs courses concerned with developing interpersonal communication and rapport-building skills. (Raybourn, 2005).

Capability Gaps

However, while these and other initial efforts at creating game-based training systems for the military have shown promise, there are several limitations in current game-based training technology and methods that hamper widespread training game development and application, including:

- Lack of empirical data supporting the learning effectiveness of games.
- Lack of methodology to guide pedagogical design. Commercial games typically lack an instructional design component, requiring significant supplementary effort and materials to use them effectively for instruction.

- Lack of appropriate performance measures. A key limiting factor in the use of gaming technology for training is the lack of analysis tools for enabling rapid collection of relevant high-order behavior events at the right level of detail for tracking and assessment of trainee performance. Most game-based training systems have to rely on the use of human observers, self-report, questionnaires, and/or in-person AARs.
- Lack of a systematic methodology for game-based training development. Current development is ad-hoc, leading to low levels of reusability and poor design decisions. Current authoring facilities severely hamper scenario modification to meet new and changing needs and result in high life cycle costs.

Based on our preliminary analysis of current game-based training methods and military training objectives, we believe that there is a strong need to **produce the capability for repeatable creation of high-quality, effective, low-cost training and assessment products that support cognitive readiness.** To achieve this, we believe that three key advances are needed in our methods and technology:

- A set of readily usable authoring tools to support game scenario design and editing. The tools should guide the developer in applying sound ISD principles, enable the developer to specify and organize events within a game, and aid the developer in embedding automated assessment and feedback of user performance.
- High-level performance capture and assessment mechanisms to support measuring performance effectiveness, diagnosing trainee's underlying problems and instructional needs, delivering feedback, providing remediation and conducting After Action Reviews.
- Support for the learning of cognitive readiness skills exhibited in realistic operational tasks and situations. Cognitive readiness entails a complex of skills including problem-solving, decision-making, situational awareness, adaptability, risk evaluation, uncertainty avoidance, conflict resolution, and team-working. Measurable behaviors associated with these skills can effectively be assessed and trained in the context of immersive experiential games.

In the following sections, we expand upon these needs and offer suggestions on approaches to achieve effective improvements in the state of the art.

Framework for Pedagogically Sound Game-Based Training

To ensure long-term success for the military's efforts in developing game-based training and assessment simulations, they must be built on a pedagogical foundation; performance measures must be based on learning objectives. Developers must apply sound instructional system design (ISD) principles. A specific ISD method will guide the way in which the learning problem is viewed, the way in which the learner is engaged and involved in the process, how the learner's performance is assessed, and what actions are available to the learner and the instructor or instructional system. Different ISD methods are appropriate for different approaches to learning (e.g., behaviorist approach that focuses on specifying training in terms of small, easily measurable tasks, a cognitive approach that specifies training in terms of increasingly complex tasks whose outcome is measurable, or a constructivist approach that provides little overt structure to the training, but encourages the learners to reflect as they perform a complex real-world task).

A principled foundation for developing pedagogically sound, performance-based game-based training should provide processes for developing appropriate instructional, personal, and emotional narratives within the system. These narratives should allow the mapping of pedagogical objectives to specific elements and events in a game, and the linking of specific measures of human performance to specific actions of the learner. For example, the instructional narrative may indicate that a specific event should occur in order to initiate an assessment opportunity (e.g., performance under stress measured in terms of remaining calm, taking in all available information, processing the data, and firing accurately). The emotional narrative may indicate the specific way in which the situation is depicted (e.g., under heavy fire, low visibility). The personal narrative may indicate that the learner's character has limited resources and actions available (e.g., low on ammo, restricted mobility due to injury).

Human Performance Measures

A major limiting factor in the use of games for training is the lack of effective measures and associated methods for capturing performance data during the execution of a game-based training scenario. Instructionally effective training games for domains that address core Navy needs must identify key human performance measures as well as the specific training needs of the end users. The focus should be on measures of high-level cognitive skills (e.g., leadership, maintaining situational awareness) rather than low-level perceptual or motor skills (e.g., hand-eye coordination). For instance, all echelons share a strong need to develop core cognitive readiness skills (see Table 1), as well as job-specific skills (e.g., analysis, projection, flexibility). The focus should also be on measuring changes in performance under different conditions (e.g., high/low stress) and over time (e.g., track progress).

The choice of human performance measures depends on three factors: 1) the importance of the underlying skill in the context of a specific training scenario; 2) the feasibility of measuring behaviors that reflect the underlying skill; 3) the generality of the measured behavior (e.g., relevance to other training scenarios). Some of the behaviors in Table 1 are much easier to capture (e.g., percent of playing time with no player action) than others (e.g., detecting that an individual was monitoring the performance of a subordinate).

Table 1: Skills for In-Game Measurement

Skill	***Measurable Behavior***
Decision-making	Assessment of timeliness of decisions, even under incomplete information
Flexibility/adaptability	Amount of effort devoted to a new approach vs. defending existing one
Tendency to action	Percent of playing time with no player action
Risk taking	Evaluation of different possible responses to risks before taking action
Uncertainty avoidance	Response to unexpected events that may cause deviance from a set plan
Contradiction handling	Actions taken toward resolution of conflicting information
Problem-solving	Assessments of costs/benefits of alternative solutions

Skill	*Measurable Behavior*
Coordination response	Actions taken that are consistent or inconsistent with directions
Directing subordinates	Monitoring individual performance, giving appropriate feedback/assistance
Teaming	Communication of team goals, organization of effective teamwork
Tactical Knowledge	Directions follow established operational practice
Leadership	Subordinates follow orders
Situation Awareness	Directions properly take into account tactical situation
OPFOR Awareness	Directions based on sound knowledge of the Opposing Force, their Tactics, Techniques, and Procedures (TTPs)

A training system should provide a basic capability for measuring instances of specific core skills during game-play. Later, it should broaden its scope to include comprehensive measures of aggregate performance over periods of time within the game. For instance, while a game may initially capture only instances in which the user issued commands, it may eventually be instrumented to capture the user's style of leadership under specific conditions, such as high-stress and high-risk.

Tools for Capture of High-Level Performance Data

The trainee performance data collected during a game needs to capture the "high-level" behaviors that characterize significant aspects of task performance. Capture of a huge stream of low-level details (such as menu options clicked, location over time, recorded use of an item) requires significant post-processing and does not measure the relevant behaviors that characterize performance. A game should be capable of capturing a variety of in-game, high-level user behaviors, such as "issuing a command", "making a decision", or "reviewing alternatives", as well as the context in which those behaviors occur within the game. It should capture context at a high level, such as the level of risk in the user's current position, the level of stress that the user is under, or the number of decision constraints the user is facing.

In order to develop the high level performance measures required for assessment and diagnosis, several types of game events may need to be captured and a varied set of capture tools may need to be implemented, such as the following:

- Trigger objects – objects placed in the scenario and tied to specific actors. Whenever a specific action occurs, a trigger object sends a notification. For example, a proximity detector may record all points in time when the player is within a certain range of an enemy base.
- Event-centric objects – The trainer may define a certain event, statically (such as the appearance of an enemy) or dynamically via a trigger. For this event, the user may define a number of capture mechanisms that are initiated once the event occurs.
- Communication monitors – Capture various types of communication, for example, commands issued, requests for situation reports, delivering of situation reports.
- Response monitors – Record all actions (of specified types) that the player makes within a certain period of time after an event occurs.
- Sequence detectors – Explicitly define a sequence of activities to look for (e.g., via a tree of possible actions), and record whenever that sequence is performed (e.g.,

receive request for information, provide sit-rep). These can be used to identify errors in the user's actions.

- "Over-the-shoulder" recorders – Follow specific players and record all actions they take of specified type and the context for those actions.
- Dialog mechanisms -- Constrain trainee responses to particular events to specific possibilities that are, in turn, linked to specific human performance categories.

In addition to tools for directly capturing performance data, game components for aggregating and analyzing data, and for setting data collection criteria are important. For example, a data context component can capture the current context associated with a given measure, and an event analyzer component can analyze the results of the activities following an event to determine if remedial action (e.g., feedback) is necessary. These components should have immediate access to any messages originating from the players, triggers, or other game actors within the application.

The performance measures captured via such mechanisms may be used in a variety of ways to support training. For instance, they may be used to provide automated feedback, to guide the sequence of events in a training scenario, to support an After Action Review (e.g., via facilities such as logging, playback, and task tracking), to generate performance reports, or to store data for future analysis in order to improve the curriculum.

Authoring Tools for Game-Based Training

An essential requirement for reifying pedagogically sound game-based training is the development of authoring tools that facilitate the creation of games with good instructional design and that enable rapid, iterative development and re-use. This entails advances in three categories of authoring tools:

<u>Tools that guide the game developer in applying sound instructional design principles</u>

As well as establishing a scientific and engineering foundation for guiding the design and development of educational games, work is needed to ensure that game applications can be implemented to benefit training of complex skills such as cognitive readiness. There are two issues related to this challenge. The first involves streamlining the process of developing cost-effective educational games. The second involves embedding the lessons learned from research into formats supported by game development and operation tools that instructional designers find easy to use so that they can produce and run effective educational games.

In responding to the first challenge, several issues must be addressed. To begin with, the processes associated with game development for entertainment purposes are far too costly to be viable in the educational game arena. For example, a typical top-selling computer game can cost upwards of $5M to develop and can take more than a year. If such a game is built on top of a proprietary game engine, run time licenses may be cost prohibitive if distribution to a large number of trainees is desired. Moreover, modern game development practices resemble Hollywood movie making in the sense that there is little, if any, concern for reusing content or resources. Again, this strategy is not viable in a training situation where the goal is to make instructional content sharable across courses. Game design processes are also similar to feature film making in the sense that design and development practices tend to be idiosyncratic, i.e., not subject to a set of common or standard workflow practices. We assert that this must change; what is needed is a strategy to track and rationalize the design process both to reduce costs and to ensure that educational objectives are met.

The second challenge associated with application has to do with providing instructional designers viable help in designing effective educational games. We believe strongly that much of what we know from the learning sciences has not affected the actual design of training systems. While there are many reasons for this, we believe that a major one is that the results of learning science experiments are not often translated into design guidance that is easy to apply in actual system development. Paper-based guidelines such as the "What Works" series (e.g., Montague, 1988; Montague & Knirk, 1993) are clearly a step in the right direction. But we are convinced that taking paper-based guidelines a step further and casting them into an automated design aid is crucial. Our logic relates to the nature of the design task and the realities in which it exists. The task of designing a complex system like an educational game involves making a large series of important decisions throughout the process. In order to be useful, design guidance must have an impact on the decisions that are made as a system develops. We believe that the best way to help in this area is to provide on-line guidance that is available as a continuing, easy-to-access resource during the design process, together with a repository of work that can inform and benefit subsequent educational game development efforts.

<u>Tools that enable the developer to specify and arrange the training events within a game</u>
An important criterion of success for the adoption of educational games in the military will be how easily trainers can apply a pedagogical framework and embed appropriate data capture mechanisms. Developing an effective game-based training and assessment scenario requires the usual specification of low-level game elements (e.g., terrain, characters, equipment, placement of objects). Further, it requires the trainer to instrument a scenario with appropriate data collection activities at appropriate places and times, and to embed appropriate instructional elements (e.g., questionnaires, skill tests, in-action review, real-time remedial feedback, after action review).

We suggest that there is a need for an authoring interface that provided game developers and trainers with capabilities for:

- walking them though high-level instructional design decisions,
- prompting them to add narrative elements that are tied to objectives,
- enabling them to define specific pedagogical events that relate to the narratives,
- helping them setup and link capture mechanisms to appropriate objects and events,
- specifying feedback options that they may tie to specific events or objectives, and
- providing report formats that summarize data on the performance and activities of the player(s) during and after a training experience.

It should also provide capabilities that support and encourage re-use of existing game simulations and simulation components. Examples include enabling developers to save a particular configuration to a library of events, to load saved events, to edit existing scenarios, or to apply a new instructional design template to an existing scenario.

Tools that allow the developer to embed automated assessment and feedback
Good instructional design determines what type of feedback should be given to the user and when it should be provided. Several feedback mechanisms should be supported, including continual in-game feedback (e.g., via a status bar), event-driven feedback (e.g., upon completion of an objective or a test), and an After Action Review (AAR) including summary reports and replay options.

During the AAR, the user performance data should be available to the trainer on an interface that displays high-level user behaviors on an intuitive and scalable timeline. The AAR tool should have capabilities for analyzing and debriefing the trainee's performance via extensive visual displays and speech outputs. The automated system should be capable of generating a real-time replay narrative with facilities for identification of significant events, associated trainee actions, and explanations of performance errors.

Conclusions
Computer gaming technology offers many potential benefits to the area of training and education, and to military training in particular. Many initial efforts have been made and a variety of useful training systems have been developed. However, the field is still in its infancy. Almost all systems developed are “boutique” solutions – using ad-hoc instructional design, minimal re-use and, sadly, minimal sharing across efforts conducted by different organizations. There is a strong need for tools and methods that provide the infrastructure for rapid design of effective training systems, and for consistent practices across the fields. To draw an analogy, the creation of Integrated Development Environments (IDEs) for traditional programming languages provided a significant leap in the capability to rapidly create and modify useful programs. Likewise, there is a need for integrated environments for pedagogical training system development. The challenge is particularly strong since the field of computer gaming itself is highly balkanized, with low re-use and no common development tools. However, even there, there is an increasing realization that such common tools and methods are needed in the gaming field, and there are several on-going efforts to provide such capabilities (e.g., Sun Microsystem’s Darkstar project is developing middleware for multi-player online gaming).

We have offered our thoughts on some initial steps that may help us improve the state of the art in pedagogical game design. BBN Technologies, in collaboration with the University of Central Florida, the University of California Los Angeles (CRESST) and other organizations in the training system development community, has recently started work sponsored under the ONR Capable Manpower program to develop and validate a coherent set of generic, widely applicable authoring and performance analysis tools for developing training scenarios. Our specific objectives are as follows.

- Conduct rigorous empirical studies to determine which game features contribute to learning effectiveness, and how these operate to improve the learning effectiveness of game-based training systems.
- Incorporate cognitive science findings in assessment methodology to improve the accuracy and reduce the time and cost of assessment development and validation,

- Create revolutionary authoring capabilities incorporating well-grounded instructional design methods for rapid, low-cost creation of game-based training systems with compelling narratives that engage the learner and support the learning objectives.
- Develop software for automated assessment of performance in game-based training.
- Develop game-based applications for training cognitive readiness skills in Navy contexts, working closely with the Naval Service Training Command.
- Empirically validate the training benefits.

In summary, we are working toward the development and application of a new science of learning in games founded on a testable pedagogical framework and validated in practice through instructionally effective cognitive readiness training applications. We look forward to presenting our initial results at next year's event.

References

Bonk, C.J. & Dennen, V.P, (2005). Massive Multiplayer Online Gaming: A Research Framework for Military Training and Education. Office of the Undersecretary of Defense For Personnel and Readiness. Technical Report 2005-1. March, 2005.

Diller, D., Ferguson, W., & Leung, A. (2004). A Game-Based Testbed for Cultureand Personality Research. Presentation, NATO Concept Development and Experimentation Project, New Orleans, La., September 2004.

Diller, D., Roberts, B., Blankenship, S. and Nielsen, D. (2004). "DARWARS Ambush! – Authoring lessons learned in a training game," Proc. of the 2004 Interservice/Industry Training, Simulation and Education Conference (I/ITSEC), Orlando, FL, December.

Ferguson, W., Weil, S., Hussain, T., Brunye, T., Sidman J., Spahr L., and Roberts B. (2005).Assessing the potential of massive multi-player games to be tools for military training," Proceedings of the 2005 Interservice/Industry Training, Simulation and Education Conference. Paper 2031. I/ITSEC, Orlando, FL

Hill, Jr., Randall W., Belanich, J., Lane, H. C., Core, M., ; Dixon, M., Forbell, E., Kim, J., and Hart, J. (2006)"Pedagogically Structured Game-Based Training: Development of the Elect BiLAT Simulation", http://handle.dtic.mil/100.2/ADA461575, National Technical Information Service.

Hussain, T.S. and Ferguson, W. (2005) "Efficient development of large-scale military training environments using a multi-player game," 2005 Fall Simulation Interoperability Workshop (Sept 18-23, Orlando, FL), p. 421-431.

Johnson, W. L and Beal, C. 2005. "Iterative Evaluation of a Large-Scale, Intelligent Game for Language Learning." Conference on Artificial Intelligence in Education.

Johnson, W.L., Wang, N., and Wu, S., "Experience with serious games for learning foreign languages and cultures," Proceedings of the SimTecT Conference. Australia. 2007.

Minton, E. (2005). "SoftWARe", Today's Officer", Summer 2005. htpps://dev.moaa.susqtech.com/TodaysOfficer/Magazine/Summer2005/software.asp

Montague, W.E. (Ed.). (1988). *What works: Summary of research findings with implications for Navy instruction and learning* (NAVEDTRA 115-1). Pensacola, FL: Office of Chief of Naval Education and Training.

Montague, W.E., & Knirk, F.G. (1993). What works in adult instruction: The management, design and delivery of instruction (NPRDC TR 93-06). San Diego, CA: Navy Personnel Research and Development Center.

O'Neil, H. & Wainess, R. (2005). "Classification of Learning Outcomes: Evidence From the Computer Games Literature" , The Curriculum Journal (in press).

Raybourn, E.M., Deagle, E,. Mendini, K., (2005) "Adaptive Thinking & Leadership Simulation Game Training for Special Forces Officers", I/IITSEC 2005 Proceedings, Nov. 28 – Dec. 1, Orlando, FL.

Roberts, B., Diller, D., and Schmitt, D. (2006). "Factors affecting the adoption of a training game," Proceedings of the 2006 Interservice/Industry Training, Simulation and Education Conference, Orlando, FL, December 2006.

Warren, R., Sutton, J., Diller, D., Ferguson, W., & Leung, A. (2004). A Game-Based Testbed for Culture & Personality Research. NATO MSG-037 Workshop, October 2004.

Weil, S., Hussain, T., Brunye, T., Sidman, J and Spahr, L. (2005). "The use of massive multi-player gaming technology for military training: A preliminary evaluation" 49th Annual Meeting of the Human Factors and Ergonomics Society. (Sept 26, Orlando, FL)