Efficient Development of Large-Scale Military Training Environments using a Multi-Player Game

Talib S. Hussain
William Ferguson
BBN Technologies,
10 Moulton St.
Cambridge, MA, 02138
617-873-6861, 617-873-2208

Keywords:
Military Training, Multi-Player Games

ABSTRACT: *In a recent study, we demonstrated that it is feasible to perform large-scale military training using a commercial off-the-shelf game with low development time and high re-use of training content. Our approach was to focus on using the game environment as a means for enabling interactions between teams of human players, and to elicit complex interactions through the instructions and objectives given to the players rather than through the complexity of the game scenario. Because multi-player games are designed to provide experiences that will entertain and engage a non-captive audience, they latently meet critical requirements for the creation and execution of team training activities. We exploited authoring tools provided with the game to design a relatively simple, "capture-the-flag"-style scenario with minimal game objects and non-player characters. However, we defined several different types of player characters with different tactical strengths and weaknesses, and created two different sets of mission objectives. We conducted a training event in which 40 soldiers played against each other in two teams. Multiple trials were held, and in each trial the mission objectives were provided as verbal instructions to the teams. Pre-mission planning and post mission de-briefing were performed in person. Observation of the event and comments from the soldiers showed a rich set of interactions, a high level of interest, positive training potential for the scenario, and the ability to effectively reuse the same environment for two different sets of mission objectives. Placing the onus of complexity on the participants and using the game as it was intended with no modifications other than simple content design proved effective and economically efficient.*

1 Introduction

Recently, the success of commercial gaming technology for engaging multiple players, from several to thousands, in a shared virtual experience has led to the examination of that technology for military training purposes. The use of computer simulations for military training has a long established history. However, effective simulations for training on a particular type of task generally require a significant development time and have limited flexibility in their use for different types of training tasks once developed. In addition to providing robust gaming engines for enabling interactions among many players, many modern multi-player games (MPGs) also offer powerful content authoring tools that enable rapid development of new gaming scenarios. Together, these capabilities may provide a rich basis for the creation of robust team-based military training simulations for a variety of tasks with low development overhead.

The Gorman's Gambit project [1,2] examined the issues involved in using an MPG to develop and deliver military training for large teams of soldiers. The project was part of the larger DARWARS program, a Defense Advanced Research Project Agency (DARPA) funded initiative that aims to establish warfighter training superiority through a new generation of training systems. DARWARS is being developed as an integrated training platform that brings together a unified pedagogical infrastructure with advanced personal computer (PC) based training and simulation systems. Within DARWARS – and in military training in general – MPG technologies promise to advance simulation-based training by enabling distributed training at a level of participation, intensity, and fidelity previously unrealized.

We describe the development efforts undertaken in Gorman's Gambit to create a game-based training scenario, present the results of an exercise conducted at Fort Benning, GA with 40 soldiers using the scenario, and summarize the lessons learned from the project.

2 Gorman's Gambit

The Gorman's Gambit project aimed to answer the following questions:

- To what extent can existing MPGs support military training of teamwork skills?

- What are the key development issues that need to be addressed for effective use of existing MPGs for military training?

The basis of the project was a thesis put forward by General Paul Gorman (U.S. Army, Ret.) that teamwork skills can be taught effectively using modern commercial off-the-shelf (COTS) MPGs, as the software supports the fundamental behaviors, and that, furthermore, there is no need for the game to be realistic with regard to modern military operations [3]. For instance, although the level of technology and the obstacles faced are very different, the elements of teamwork may be similar for attacking or defending a medieval castle and for some modern military operations. Based on this thesis, many available games have the potential to provide an experience that has pedagogical value, even though they may not have been initially designed with this in mind. The thesis has a firm ground in cognitive theory, in that learning by analogy is an effective means of obtaining skills that will be applied in unrelated contexts [4].

We concentrate on teamwork skills because of their central importance for distributed military operations in the information age. We must understand the capabilities that current COTS gaming technologies offer, as well as the critical gaps that remain. Based on this understanding, we can make progress toward requirements, so that new technologies being developed build on real capabilities and fill in real capability gaps.

3 Design

The Gorman's Gambit project had three main developmental phases over a period of several months (September to December, 2004), culminating in a final exercise held at Fort Benning, GA (in mid-December). The developers and testers were from BBN Technologies and Aptima, Inc.

3.1 Game Selection

We explored the merits of various games for use in the project and selected one to use as the basis for the training environment – Neverwinter Nights (NWN; Trademark of Wizards of the Coast, Inc. http://nwn.bioware.com/). NWN is a role-playing game set in a fantasy setting. The game allows a variety of terrains, both indoor and outdoor, and a variety of Non-Player Characters (NPCs) that are controlled via automated scripts. Several key factors determined our selection of the game:

- NWN enables the study of a large number of users (up to 64) working in collaboration.

- The game is several years old and its technology is stable. In particular, the process for setting up and running a multi-player gaming session is straightforward and robust.

- Each human player controls an avatar (the character in the game) with capabilities determined by avatar-specific skills and inventory. This enables different players to have different roles within a game.

- NWN provides a powerful authoring tool called the Aurora Toolset (Trademark of Bioware Corp, copyright 1997-2005) for developing new game scenarios.

- A special avatar called the Dungeon Master may be used invisibly within the game to observe the other avatars and non-player characters (as well as control NPCs and cause specific events to occur). This provided a mechanism for the facilitators to observe during the final exercise.

- NWN provides a tutorial game scenario to help new users learn the mechanics of playing the game and learn to exploit the main capabilities of their avatars.

3.2 Initial Phase

During the initial phase, a key factor influencing our design of the training exercise was the degree of scenario "scripting". The Aurora toolset is very powerful and provides the capability to adapt terrain, NPC behaviors, equipment properties and a myriad of other aspects of gameplay. On the one hand, a strongly-scripted approach could be followed in which the players are carefully guided through a specific scenario using in-game cues, events and NPC actions. On the other hand, a weakly-scripted approach could be followed in which the terrain and NPCs are made relatively simple, but the gameplay options open to the human players are plentiful. Since our focus was on eliciting and observing teamwork within the game, we adopted a weakly-scripted approach to require the players to consciously behave as a team rather than being constrained through scripted activities.

We adopted a "Platoon v.s. Platoon" design in which the game scenario contained a basic world with a separate "camp" for each platoon, a "long but safe" way between camps that entailed travel through a variety of terrains, and a "short but dangerous" way between camps that entailed travel through areas containing hazardous traps

(see Figure 1). Each platoon was given a simple mission – namely to capture the enemy camp and defend their own.

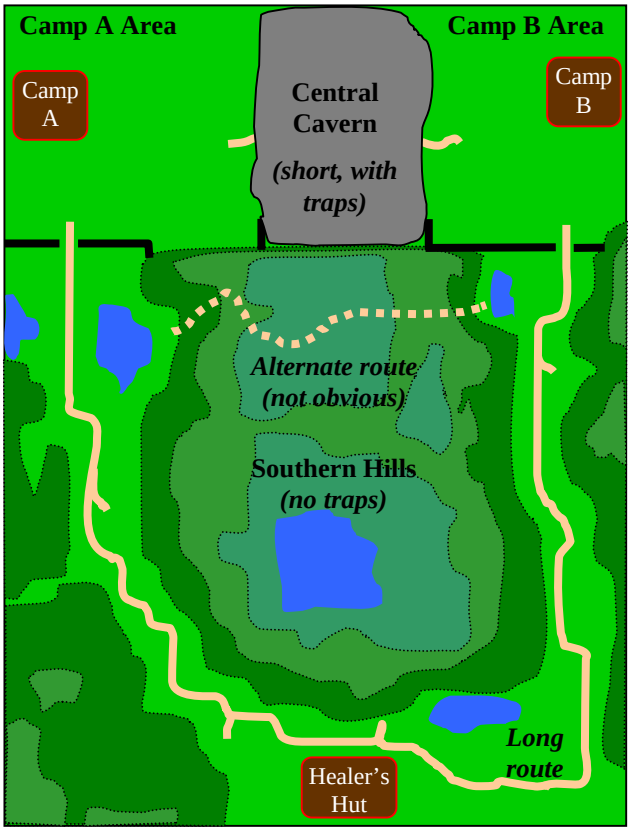

Figure 1: Terrain for Teamwork Training Scenario

A key principle we adopted was to encourage teamwork by using multiple types of avatars, each with different strengths and weaknesses. This allowed skills critical for mission effectiveness to be distributed. Avatars were distinguished primarily by the items carried (e.g., weapons, healing items, magic items) and by their intrinsic skills (sword-fighting, trap detection, spell casting and speed of travel).

To further encourage teamwork, each "platoon" was separated into multiple "squads", and leaders for each platoon and all squads were identified. Communications among players was achieved using the in-game chat mechanism. In particular, a hierarchical communication framework was used in which members in one squad could not talk directly to members of another squad, unless they were a squad leader, and where the platoon leader could only talk to the squad leaders.

In keeping with a typical game property, we enabled avatars to come back to life (or "re-spawn") once they were killed. One minute after death, the avatar would re-enter gameplay at the "Healer's Hut", located at a distance from both platoon camps (see Figure 1).

We conducted a pilot run on October 29th with 12 BBN personnel. (Two platoons of size 6, with two squads each of size 2 or 3, plus a platoon leader). All players except one had never played NWN before. All players first participated in the built-in NWN tutorial for 45 minutes and then conducted the main training exercise. The main exercise was run several times with the mission goal to occupy the enemy's camp for 2 minutes. Occupation was defined as having two of the team's players alive in the enemy camp for the entire 2 minute period.

The pilot validated a number of our design decisions and identified a number of improvements that were required. In particular, the hierarchical chat communications interfered significantly with the ability of the players to play. It produced a "type or fight" effect in that the amount of time required to type a message to a team-mate or leader was too long – if the enemy was nearby, most avatars would be killed before the user could finish typing their message. In addition, many players had difficulties learning quickly how to play NWN effectively using the built-in tutorial.

The mission goal that was used led to a very fast-paced game and did not encourage the players to exploit the varying capabilities of the different squad members. Further, since both teams had the identical, single goal, the result was purely a race condition that required little strategizing.

3.3 Second Phase

We implemented several key changes based on the results of the first pilot.

We modified the capabilities of several avatar types in terms of endurance, strength, and other parameters. The avatars developed were named scout, medic, archer, artillery, and tank, as well as a platoon leader and squad leader. As before, the avatars varied in their individual capabilities in order to promote teamwork and collaboration between players. However, it was also important to develop a good balance of skills between avatars to avoid having one avatar always having the advantage. To achieve this balance, avatars were tested against each other in pairs to determine that one was better in some situations but worse in others. For example, the tank was extremely powerful at close-up battle and heavily armored, but extremely slow, while the artillery was capable of inflicting significant damage from a distance and moved at normal speed, but was very lightly

armored. Thus, at a distance artillery had the advantage, but up close, the tank did. Figure 2 shows the platoon composition for a platoon of 20 players, with three squads of 6 or 7 (mostly) distinct avatars. The heterogeneous composition of the squads was adopted to make squadmates dependent on one another to accomplish certain tasks.

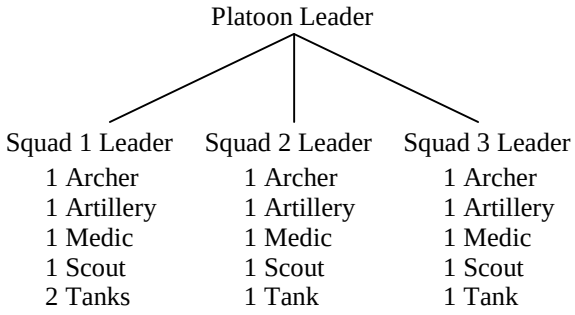

Figure 2. Platoon and Communication Organization

We incorporated a BBN-developed Voice-over-IP (VoIP) system with a hierarchical organization for voice communications similar to that of the original chat communications (see Figure 2). The VoIP software has been used successfully in the DARWARS Ambush! system [5] and is capable of readily handling 20 to 40 players. The VoIP software was setup external to the game and players wore a headset and microphone. A player could communicate with the others (according to the hierarchy) by pushing a specific key on the keyboard.

The most significant change was to the tutorial process. To address the limitations expressed during the pilot regarding the built-in tutorial, supporting written materials were developed to compress the tutorial time required by guiding the players through the tutorial. We also created two additional scripted tutorial scenarios. In the "Tutorial Arena", the unique skills of each avatar were explicitly explained and practice exercises assigned. NPCs were used to impart avatar-specific information to each player. In the "Tutorial Village", a single squad was given a mission that required the skills of multiple players to be used together and at specific points in order to achieve success. The training village contained NPCs that would provide clues to guide the players, as well as enemy NPCs that would fight the squad. This provided an opportunity for the players to learn how to play together (each squad could perform the scenario independently).

We conducted a second pilot on Nov 22 with 7 Aptima and BBN personnel. The new training process was validated and some remaining scenario limitations were identified.

3.4 Final Phase

The scenario was changed based on the results of the second pilot. The scenario and mission objectives were adapted to encourage strategizing and differentiated roles for the squads. In particular, a third "camp" was added to the game scenario. This camp was placed at a random undisclosed location and protected by several NPCs. The mission goals were adapted to encourage exploration, capture and protection of the hidden camp as well as the main two platoon camps. It was hoped that this would produce richer group coordination. To simplify the determination of when a camp was "captured", a flag was placed in each camp that indicated possession or ownership of that territory. To change possession of that territory, participants had to pull a lever placed adjacent to the flag. Once a lever was pulled, the territory would remain in the possession of the puller's platoon until a member of the opposing platoon gained control of the lever, thereby claiming it as their own.

The platoons were now provided with three goals:

- Goal 1: Defend your flag
- Goal 2: Capture the enemy's flag
- Goal 3: Capture and hold a third flag hidden in an unknown location

To further encourage different strategies, we defined two different "winning conditions" for a mission. In the first condition, the winner was the team that held the most flags at the end of 30 minutes. In the second condition, the winner was the team that held flags for the longest total amount of time by the end of 30 minutes.

A final pilot was conducted on Dec 8 with 32 BBN personnel. In addition to testing the new additions to the scenario, a focus of the pilot was to stress the server and clients by placing large numbers of avatars in the same area. The game server was a Pentium M 1.7 GHz with 1 GB RAM computer. A variety of client machines were used, ranging from Pentium III or 4 with 512 MB RAM and 32 MB graphics cards to Pentium M 2.0 GHz with 2 GB RAM and 128 MB graphics cards. The pilot showed that the server was able to handle the 32 participants and hierarchical voice and chat communications seamlessly. No clients exhibited problems when 20 avatars were in the same area. Low-end clients experienced moderate slow-down in play with 25 avatars and significant slow-down with 32 avatars. Mid-range clients experienced slight slow-down in play with 32 avatars. High-end clients experienced no problems.

3.5 Teamwork Measures

A key focus of the project was the development of tools for measuring the teamwork skills that could be exhibited within the training scenario by the players. Research has shown that skills that support team effectiveness do exist, and that these skills can be defined, trained, and assessed [6-12]. Table 1 summarizes the skills that we considered in the development of our measurement tools, such as monitoring and back-up, information exchange, and leadership. In addition to testing the development of the training scenario, the pilot studies above also enabled us to refine and validate our data collection approaches.

In developing our measurement tools, our goal was to obtain a sample of behaviors to provide evidence that existing MPGs can support military training of teamwork skills. Our goal was not to capture every instance of teamwork. As such, we emphasized the careful observation of teamwork skills rather than the automated capture of instances of those skills during gameplay.

Several data collection approaches were used. Our primary measurement tool was an *In-Session Observation Form*. Observers watched the gameplay either in an "over-the-shoulder" manner or via the Dungeon Master avatar. Using a paper and pencil based form, observers documented events that involved teamwork. Event details included the time, the participants involved, a detailed description of the event, the presence or absence of teamwork skills (as described in Table 1), ratings of those skills, and comments. Observers trained on their role during the pilots. In the exercise, some observation was also done using a video camera, to allow for later coding.

To encourage the use of teamwork skills, and to reflect the realities of military operations, we incorporated a *Team Planning* period prior to each game session and a *Team Debriefing* period immediately following each game session. During the October 29 pilot, planning was conducted using either in-game chat or verbal communication. The former was shown to be difficult to conduct and to observe. In the final exercise, planning and debriefing were conducted verbally by each team and observations were captured using free form notes and by video. The planning and debriefing periods provided excellent opportunities to capture instances of adaptability, leadership, team orientation, as well as overall gaming experiences.

Our third measurement approach was a *Post-Exercise Questionnaire*. The questionnaire contained 21 questions, nineteen of which required responses on a Likert scale of low (unfavorable) to high (favorable) ratings. Five questions allowed for open-ended answers to further detail their experiences and opinions (in addition to a rating). Questionnaires were completed anonymously and were labeled only by their team and avatar type (e.g., Team A, Medic), and participant number. Questions probed several areas, including game experience, avatar knowledge, team interactions, self and team assessment, situational awareness of self and team, and practical applications of the game as a training tool.

Table 1. Teamwork Skills Supporting Team Effectiveness

Leadership: The ability to direct and coordinate the activities of other team members, assess team performance, assign tasks, develop team skills and abilities, motivate team members, plan and organize and establish a positive atmosphere.
Monitoring: The ability to develop common understandings of the team environment and apply appropriate task strategies and processes in order to accurately monitor teammate performance.
Back-Up Behavior: The ability to anticipate other team members' needs through accurate knowledge about their responsibilities. Includes the ability to shift workload among members to achieve balance during high periods of workload pressure.
Adaptability: The ability to adjust strategies based on information gathered from the environment through the use of compensatory behavior and reallocation of intra-team resources; altering a course of action or team repertoire in response to changing conditions (internal or external).
Team Orientation: Propensity to take others' behavior into account during group interaction and the belief in the importance of the team's goals over individual member's goals.
Closed Loop Communication: The practice of confirming receipt and understanding of others' communications. This practice builds trust in the communication skills, knowledge, and intent of others and ensures that information is accurately conveyed.
Team Mental Models: The ability to accurately represent the capabilities of others, their responsibilities, and their perception of the state of the world.
Coordination: The practice of planning, preparing, organizing people and/or tasking to accomplish a goal.
Communication Push: The practice of sharing or sending information with/to others.
Communication Pull: The practice of seeking information from others or other data sources; asking questions; attempts to gather intelligence.

Once all game sessions were completed, an *After-Action Review (AAR)* was conducted. The AAR was semi-structured with a set of core questions to field to the participants. Facilitators conducted the final AAR and captured the event on videotape, with participants' permission. The questions addressed participant likes and dislikes of the game, team challenges, teamwork examples, strategies and adaptations, relevance to Army tactics and training, and comparison to other MPGs.

4 Exercise

The Gorman's Gambit exercise was conducted on December 15 and 16, 2004 at Fort Benning, GA.

4.1 Participants

Forty members of a U.S. Army Infantry Platoon participated in the exercise. The soldiers were between 19 and 33 years of age (mean of 23.6 years) with between 1.5 and 174 months of military experience (mean of 51.2 months). Participants ranged from E-2 (Private) to O-1 (Second Lieutenant), and twenty-five of the soldiers had been deployed within the last year. The soldiers averaged 3.7 hours per week of computer use in the last year, with an additional 4 hours of game-playing (on personal computers or video game console systems) experience per week in the last year.

Participants were separated into two competing platoons of equal size (i.e., 20 participants each). Each platoon was composed of 3 squads, each with a similar mix of player types. Each participant was assigned a specific role and avatar (platoon leader, squad leader, scout, medic, archer, artillery, tank), as illustrated in Figure 2.

On the first day, the soldiers trained on how to play the game and use their assigned avatar with the three tutorials (built-in tutorial with written aids, tutorial arena and tutorial village). Six hours were required to complete all the tutorials and resolve any technical difficulties the soldiers encountered.

4.2 Gorman's Gambit Exercise

On the second day, the main Gorman's Gambit exercise was held over a period of four hours. The soldiers conducted three successive game sessions using the main training scenario. Each session started the game scenario with the same initial conditions and with each platoon in its own camp. However, the sessions used different mission winning conditions and different communication modes, as summarized in Table 2. The winning condition manipulation was conducted to encourage different strategies between sessions. The communication manipulation was conducted to gauge the effects of different communication mediums on teamwork behaviors and to emulate two use case possibilities: high and low bandwidth network conditions.

Table 2. Winning Condition, Communication by Session

Session	Winning Condition	Communication Mode
1	**Possession at end of 30 minutes**	**Voice-over IP**
2	**Possession for longest duration**	**Text chat**
3	**Possession at end of 30 minutes**	**Voice-over IP**

At the beginning of the day, the soldiers were given written instructions that described the events of the day, and were provided with the mission objectives and a map of the terrain. Soldiers were also instructed to maintain contact with their leaders, to assist their teammates, and to engage the opposing team in combat when necessary.

Each session lasted 30 minutes and was preceded by a 15 minute planning period and followed by a 15 minute debriefing period. At the end of the final session and debriefing, the post-exercise questionnaire was completed by all participants. A 30 minute AAR was then conducted. A camera was used to videotape selected activities during each planning period, game session and debriefing, as well as to record the entire AAR.

During planning, the platoons were placed in separate rooms for privacy. During this time, platoon leaders relayed their strategies for the current mission. Both team had access to a white board.

During each session, participants sat at computer stations which contained a monitor, keyboard, mouse, and headset. Team A was located in one large room, where they were seated as pairs at computer cubicles. Due to facility constraints, Team B was distributed among four rooms, with squads largely seated together.

During post-exercise debriefing, the two platoons were again divided into separate rooms. The debriefings were conducted by the platoon and squad leaders.

The AAR was conducted jointly with both teams, based on questions asked by the facilitators.

5 Results

The Gorman's Gambit exercise provided many insights into the possibility of using existing game technology for training purposes. Broadly, we discovered that existing games (with custom scenarios) provide a powerful tool for teaching teamwork skills in a context that is divorced from the biases of any particular real situation. This may have great value in teaching warriors how to cope with situations for which they were not trained and to use resources that they did not know they would possess or need. On the negative side, we learned that it is hard for students and instructors to accept the potential validity of simulation-based training if that training resembles a game. Thus we learned of a new source of requirements for appropriate fidelity in training simulations: the users (trainers and learners) must be able to take it seriously. This requirement will probably vary greatly according to the institutional culture of each group of potential users.

5.1 Exercise summary

The tutorials on the first day went smoothly and very few soldiers experienced significant difficulties in learning to play the game at a basic level and to understand the basic capabilities of their avatars. However, many soldiers failed to learn the full skill set of their avatar.

The planning sessions were effective, though the two teams showed distinct styles. One platoon leader used the written map as a basis for describing and discussing the plan with his squad leaders; the squad leaders in turn relayed the plan to their squad members. The other platoon leader used the white board for describing and discussing his plan with his entire platoon.

Team B won the first and third sessions, and Team A won the second session. There were two technical issues that arose. In the second session, a technical problem with the healer's hut occurred and was resolved after a few minutes. Also, some soldiers (on both teams) discovered an unanticipated way to re-spawn in the same location where the avatar had died. This was eventually observed and the soldiers were instructed not to do that in the second session. They persisted in the third session and a way was found to disable that ability.

The debriefing periods were used effectively as opportunities for each team to reflect on specific events during the session, such as successes, failures, and surprises. Both teams identified lessons learned and applied those to future sessions.

5.2 Lessons Learned

Nine key lessons were learned regarding the critical technological, logistical and pedagogical issues in the Gorman's Gambit exercise. These results reflect our goal of informing the development of a MPG-based training platform, and shed light on the requirements gaps and capabilities necessary for moving from gaming to training.

1. Teamwork skills can be observed in MPGs.

Using multiple data sources including gameplay video, notes, and observer in-game evaluation forms, more than 115 examples of specific teamwork skills were extracted and reviewed. Most events involved multiple teamwork skills, as teamwork skills often work in conjunction with each other. For example, the *leadership* and *monitoring* skills have overlapping definitions that would make it probable, although not without exception, that where you have *leadership* you also have *monitoring* and vice versa. The teamwork skills demonstrated in this exercise suggest that MPGs can support large-scale exercises in which many individuals with differing but complementary skills work together towards a single effort.

2. Any game-based training system must be robust.

After Action Review responses indicated high frustration levels when opponent team members were able to take advantage of software vulnerabilities. For example, several soldiers found ways to bypass intentional game manipulations; by exiting and restarting the NWN application upon death, players were able to immediately re-spawn on location. In addition, by inviting opponent team members into chat groups, several players were able to view enemy communications during the text chat gaming session. These examples of soldier resourcefulness led to frustration and reduced effectiveness of the opposing team, and undermined the facilitator-imposed scenario structure. For these reasons, while the training product must be robust enough to allow for flexibility, it must also guard against user or environmental factors that could mitigate training value.

3. Targeted avatar design and rapid modification are critical for training effectiveness.

Training effectiveness is contingent upon modification and development of avatar characteristics in a manner that is quick and effective enough to correspond with the changing demands of players and trainers. For example, on the first day of the exercise, the platoon leaders indicated that they were not satisfied with their avatar's slow speed and low combat abilities, and they believed these deficiencies detracted from their ability to lead their

platoons effectively and gain their soldiers' respect. Targeted attribute changes were made quickly and easily (i.e., in a matter of minutes), allowing content developers to increase the speed and combat effectiveness of these avatars. The platoon leaders reported increased levels of satisfaction and effectiveness as a result of this rapid manipulation. The incorporation of authoring tools tailored for NWN allowed for easy initial authoring as well as rapid avatar modifications. These tools require moderate amounts of technical expertise, but do have a graphical user interface (GUI) which allows rapid familiarization with functionality.

4. Targeted scenario design is critical for training effectiveness and operational acceptance.

The ability to easily alter existing scenarios or create new targeted scenarios is fundamental to the development of game-based training simulations. Existing gaming technology may provide a venue for occasionally practicing specific skills within the virtual gaming environment. However, highly effective training demands tailored scenarios that are *designed* to exercise specific skills, in a repeatable and consistent manner. Our development of infantry-specific gaming scenarios considered several aspects of basic infantry training in an effort to correspond with the participant pool. For instance, the NWN software is largely limited to ground-based interaction, congruent with the operational requirements of the infantry. There were several powerful analogies. NWN allows the artillery avatar to summon a small flying goblin. The soldier's point of view could be changed to take on the goblin's perspective and fly quickly through dangerous areas, avoiding traps and enemy engagement, thereby allowing for safe reconnaissance. This was quickly recognized as being analogous to the real-world function of an unmanned air vehicle (UAV). In addition, several soldiers commented that the resource and communication manipulations corresponded to actual military conditions. Gorman's Gambit required the manipulation of gaming aspects so that they would correspond with infantry operations. However, a variety of military users (e.g., Air Force, Marines, Navy, etc) will require the ability to cater game aspects to specific trainee populations.

5. Ease of conducting and maintaining a training session are critical for successful participation.

When managing the interaction of 40+ players, simple connectivity that is mostly transparent to the user is a necessity. Within NWN, setting up the client and getting started playing the game occurred in a matter of seconds. The simple networking capability of NWN allowed easy deployment on a large number of machines. If a restart of the server was needed (e.g., to start a new mission), the users could easily join the new game (within seconds). Without this capability, the exercise would have been constantly delayed. If the process of starting a session or correcting technical problems involves a significant amount of time or complications, the level of user participation and "immersion" will be highly affected.

6. Using gaming technology for training requires additional assessment tools.

One limitation of the present research was the inability to incorporate built-in assessment measures for use during training. COTS gaming systems do not generally have the assessment components that are critical for effective training. For instance, while it is possible to capture the locations of avatars at any time within NWN, this data is insufficient to determine that one of the key teamwork skills has been demonstrated. Training systems require the ability to observe from a global perspective, collect data from measures of performance fully integrated into the simulator (e.g., text logs to allow for later communications analysis), and incorporate tools for AAR. Thus, the use of the NWN COTS software did not provide the opportunity to pursue such assessment methods.

7. Training superiority can be improved by carefully controlling level of fidelity.

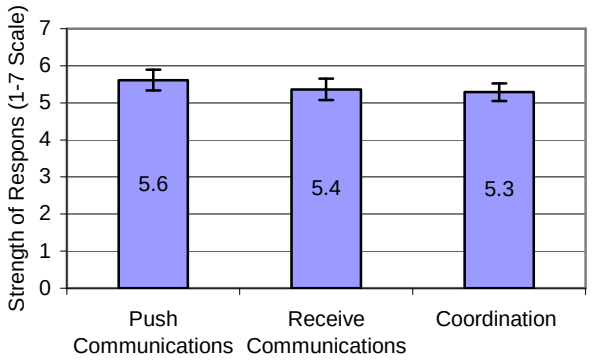

Figure 3. Average Ratings for Three Teamwork Behaviors

Laptop- and desktop-based simulators do not have the same physical or visual fidelity as large simulators, but can reflect real-world characteristics to a high degree. Varying the degree of fidelity for certain dimensions based on training goals will provide training value to a low cost system. For instance, important components of teamwork skills are communication and coordination. Thus, in the work reported here we attempted to provide an adequate and realistic communication capability. Our efforts were rewarded, for the soldiers commented that the communication system was a good analogue to their

typical hierarchical system. The questionnaire results show (Figure 3) that soldiers thought they were able to push information at a high level to others, receive a high level of communications, and coordinate at a high level.

8. Training on the interface and avatar roles and skills is essential.

In this exercise, the tutorial process took several iterations, with more than one phase – first basic gameplay skills, then individual avatar skills and finally teamwork skills – to facilitate effective interaction and coordination. In the first tutorial phase, participants learned characteristics and capabilities of their avatar, and in the free-for-all format of the next training phase, participants practiced exercising the capabilities of their avatar and learned more about the dynamics of the game. These two aspects of the training paid off well. In the post-exercise questionnaire, the mean response was 6.2 out of 7 to a question asking how well they understood what their avatar was supposed to do. Clearly, the soldiers felt competent in exercising the capabilities of the avatar they controlled in the game. During the final training phase, all the members of a team engaged a specially designed practice scenario to help them develop teamwork and foster coordination. Not only did this phased approach to training succeed, but the post-exercise questionnaire results indicated that the better that soldiers understood their characters (i.e., the better trained that they were to play the game), the greater the capacity for training and user acceptance. Moreover, soldiers rated the tutorial process as moderately to highly effective, with an average rating of 4.9 (out of 7).

9. Low operational realism has both benefits and drawbacks.

The fantasy setting afforded the exhibition of the teamwork behaviors predicted by General Gorman. Observers noted that participants were easily able to identify parallels between the components of the fantasy realm (small flying goblins) and equivalent military objects (UAVs). However, the departure from their operational environment also makes it more difficult for buy-in from military participants, as evidenced by their feedback. Some soldiers reported having difficulty taking on a military mind-set while playing the game. It is plausible that this bias could inhibit skill transfer. The most frequent negative comment voiced during the AAR by soldiers was poor realism (43%). This indicates that a sizable minority of the soldiers was conflicted by the game environment and believed that more operational realism is required for adequate transfer of training to occur. However, the non-operational setting could easily be used to train critical thinking skills and to challenge participants to engage appropriately in unfamiliar environments.

6 Conclusions

The military is interested in supporting effective, large-scale, distributed, simulation-based training that will enhance and expedite soldier instruction. The results of the Gorman's Gambit project demonstrate that commercial multi-player game technology may serve an important role in the development of future training systems.

Observation of the event and comments from the soldiers showed a rich set of interactions, a high level of interest and positive training potential for the scenario. Our "weakly-scripted" approach proved to be effective for eliciting teamwork with low development overhead. In a brief development period of approximately three person-months, a number of scenario variations were developed and new changes were rapidly incorporated based on feedback from pilot studies. Even during the main exercise itself, content changes were easy and quick to make and test. Placing the onus of complexity on the participants and using the game as it was intended with no modifications other than simple content design proved effective and economically efficient. Further, the approach enabled reuse of the same environment for different sets of mission objectives.

A variety of evidence demonstrated that teamwork does occur during MPG game-play, in the form of coordination, leadership, monitoring, team orientation, back-up, adaptability, communications pushing and pulling, and closed-loop communication. The MPG environment, therefore, is sufficient to produce instances of teamwork, and may provide an appropriate medium for instruction of military teamwork skills.

However, a number of capabilities gaps exist. For instance, while it is possible to make a variety of content changes using COTS authoring tools, existing MPG gaming technologies do not have the components critical for teamwork training assessment, such as embedded tools for observation, in-game performance metrics, and tools for supporting AAR. If these capabilities gaps are met successfully, then existing commercial gaming technology will provide a powerful tool for the development of effective, large scale, military training environments.

7 Acknowledgements

This work was supported by DARPA, under the direction of Ralph Chatham, via Office of Naval Research (ONR) contract N00014-03-C-0279. The views and conclusions contained in this document are those of the authors and should not be interpreted as presenting the official policies or position, either expressed or implied, of the Department of Defense or the U.S. Government unless so designated by other authorized documents. Shawn Weil, Tad Brunye, Jason Sidman, Frederick Diedrich and Lisa Spahr were the key Gorman's Gambit team members from Aptima, Inc., and provided critical contributions to the research presented here and to the identification of the lessons learned. We thank General Paul Gorman (U.S. Army, retired) for the initial inspiration; Harold F. O'Neil and Richard Wainess of Advance Design Information, Inc. for their contributions to the project; Len Gittleman for his excellent videography; and Bruce Roberts and Elliot Entin for their ideas and help. Particular acknowledgements are given to the soldiers who participated in the final exercise at Ft. Benning, GA.

8 References

[1] Hussain, T.S., Weil, S.A., Brunye, T.T., Diedrich, F., Entin, E.E., Ferguson, W., Sidman, J., Spahr, L., Macmillan, J. and Roberts, B.: Gorman's Gambit: Assessing the Potential of Massive Multi-Player Games as Tools for Military Training, Final Project Report, BBNT Solutions, LLC and Aptima, Inc., 2005.

[2] Weil, S.A., Hussain, T.S., Diedrich, F.J., Ferguson, W., and MacMillan, J.: "Assessing distributed team performance in DARWARS training: Challenges and methods." Proceedings of the Interservice/ Industry Training, Simulation, and Education Conference, Orlando, FL., 2004.

[3] Gorman, P.: Comments at DARWARS program meeting. Washington, D.C., July 2003.

[4] Gentner, D.: "Mechanisms of analogical learning." In S. Vosniadou and A. Ortony, (Eds.), Similarity and analogical learning. London: Cambridge University Press, pp. 199-241, 1989.

[5] Diller, D.E., Roberts, B., Blankenship, S., and Nielsen, D.: "Darwars Ambush! – Authoring lessons learned in a training game." Proceedings of the Interservice/Industry Training, Simulation, and Education Conference, Orlando, FL, 2004.

[6] Salas, E., and Cannon-Bowers, J. A.: "The science of training: A decade of progress." Annual Review of Psychology, 52, pp. 471-499, 2001.

[7] Cannon-Bowers, J.A., Tannenbaum, S.I., Salas, E. and Volpe, C.E.: "Defining team competencies and establishing team training requirements." In R. Guzzo and E. Salas (Eds.), Team effectiveness and decision making in organizations. San Francisco: Jossey-Bass, pp. 330-380, 1995.

[8] Freeman, J., Diedrich, F. J., Haimson, C., Diller, D. E., and Roberts, B.: "Behavioral representations for training tactical communication skills." Proceedings of the 12th Conference on Behavior Representation in Modeling and Simulation, Scottsdale, AZ, 2003.

[9] Serfaty, D., Entin, E.E., and Johnston, J.H.: "Team coordination training." In J.A. Cannon-Bowers and E. Salas (Eds.), Making decisions under stress: Implications for individual and team training. Washington, DC: American Psychological Association, pp. 221-245, 1998.

[10] Sims, D.E., Salas, E., and Burke, C.S.: "Is there a 'Big Five' in teamwork?" 19th Annual Conference of the Society for Industrial and Organizational Psychology, Chicago, IL, 2004.

[11] Smith-Jentsch, K.A., Johnston, J.H., and Payne, S.C.: "Measuring team-related expertise in complex environments." In J.A. Cannon-Bowers and E. Salas (Eds.), Making decisions under stress: Implications for individual and team training. Washington, DC: American Psychological Association, pp. 271-295, 1998.

[12] Smith-Jentsch, K.A., Zeisig, R.L., Acton, B., and McPherson, J.A.: "Team dimensional training: A strategy for guided team self-correction." In J.A. Cannon-Bowers and E. Salas (Eds.), Making decisions under stress: Implications for individual and team training (pp. 271-295). Washington, DC: American Psychological Association, 1998.

Author Biographies

TALIB S. HUSSAIN is a Senior Scientist at BBN. He holds a Ph.D. in Computer Science from Queen's University at Kingston. He is an expert on genetic algorithms, genetic programming and neural networks, with a broader interest in learning and training for both machines and humans. He was the Principal Investigator (PI) on the recently completed, DARPA-sponsored Gorman's Gambit project, which studied the design issues involved in training human soldiers teamwork skills using modern multi-player game technology. He was also the PI

on the recently completed, Army Research Lab (ARL) sponsored Advocates and Critics for Tactical Behaviors (ACTB) project, which applied domain-dependent genetic algorithms to evolve the paths and actions of a simulated unmanned ground vehicle. In the past, he was a key contributor on the Future Combat Systems (FCS) Unit of Action Supportability project, which applied genetic algorithms to the problem of scheduling fuel distribution for moving combat vehicles. He was also the lead investigator for the DARPA-sponsored Colonies of Neural Networks (CoNN) effort, which applied genetic algorithms, neural networks and attribute grammars for the representation and evolution of complex neural network solutions to a battlefield course-of-action prediction problem.

WILLIAM FERGUSON is a Division Scientist at BBN. His background is in artificial intelligence, simulation, computer-based training, and commercial game technology. He is currently Co-PI of the integration effort under DARWARS, a DARPA program to develop large, deployable simulation-based instructional systems using technology and ideas from the commercial gaming world. He also serves as Co-PI of the Cultural Modeling Testbed, a Defense Modeling and Simulation Agency (DMSO) and Air Force Research Lab (AFRL) jointly sponsored project to use commercial games to study culture and personality. He was the PI on a recently-completed, ONR-sponsored project on behavior authoring techniques used in commercial games. He worked for many years as the technical lead for the Analysis of Mobility Platform (AMP) project, USTRANSCOM's transportation programmatics modeling system. In the past, he served as technical lead for CSMART/ABC, a component-based system for exploring the behavior of large, heterogeneous software systems; as well as system architect and technical lead for the OASYS (Operability Assessment SYStem) project. OASYS was a simulation-based design system that provided tools for non-programmers to assemble simulations of proposed new equipment and software. OASYS supported human-in-the-loop experiments to predict human-system performance and assess concepts of operations.