Proceedings of the 2005 Winter Simulation Conference
M. E. Kuhl, N. M. Steiger, F. B. Armstrong, and J. A. Joines, eds.

SIMULATING SCENARIOS FOR RESEARCH ON CULTURE & COGNITION USING A COMMERCIAL ROLE-PLAY GAME

Rik Warren

Air Force Research Laboratory
AFRL/HECS
Wright-Patterson AFB, OH 45433–7604, U.S.A.

David E. Diller
Alice Leung
William Ferguson

BBN Technologies
Cambridge, MA 02138–1119, U.S.A.

Janet L. Sutton

Army Research Laboratory
Fort Sill, OK 73503, U.S.A.

ABSTRACT

Most research on culture and cognition uses self-report tasks such as paper and pencil questionnaires. Such tasks are inexpensive, quick, and easy to score, but they are vulnerable to response bias and manipulation effects. Action-based or performance tasks can be more absorbing and permit more of someone's natural behavior to emerge but are rarer due to increased costs, lower experimenter control, and difficult logistics. Computer games can potentially regain the benefits of real performance and immersive play while retaining experimenter control and keeping costs low. Properly constructed, computer games can simulate action-demanding scenarios which embed opportunities for personality and culturally-conditioned behaviors to manifest themselves. This is especially true when computer-simulated non-player characters are included which exhibit carefully modeled behaviors. However, such simulations are not themselves panaceas. This paper examines some of the concepts we have tried, the challenges we have faced, and the lessons we have learned.

1 INTRODUCTION

Most research on culture and cognition uses self-report tasks such as paper and pencil questionnaires (Brouwers et al. 2004). Such tasks are inexpensive, quick, and easy to score. But they are vulnerable to response bias and manipulation effects. Response bias can be due to outright attempts to misrepresent oneself for private reasons and portray a different person than one truly is. However, response bias can be more subtle, insidious, and even unknown to the participants themselves. There are cultural differences in the way people use rating scales (Fischer 2004). For example, Arabs tend to use the extreme ends of typical rating scales whereas East Asians tend to use the middle of the scales when they are asked if they agree or disagree with a statement.

Action-based or performance tasks can be more absorbing and permit more of someone's uncensored behavior to emerge but are rarer due to increased costs, lower experimenter control, and difficult logistics. Computer games can potentially regain the benefits of real performance and immersive play while retaining experimenter control, keeping costs low, and reducing the logistics of an experiment. Properly constructed, computer games can simulate action-demanding scenarios which embed opportunities for personality and culturally-conditioned behavior to manifest itself. This is especially true when computer-simulated non-player characters are included which exhibit carefully modeled behaviors (Diller et al. 2004).

Moreover, computer-based games can be used to study cultural effects that emerge only during team play. Although paper and pencil self-report questionnaires are instruments for studying individual people, games can be used to study the play of arbitrary aggregates of people. That is, teams can be composed of people from a homogeneous culture, or they can be deliberately composed of people from mixed culture and nationality groups. Hence, we can study not only differences in cognition and decision making across cultures, but also the performance effects of mixed-culture teams. In today's world of multinational corporations, multinational coalitions, and cross-cultural conflicts, research on cross-

cultural and mixed-cultural effects is critical and game-based techniques can become invaluable enablers.

However, such game-based simulations are not themselves panaceas. This paper examines some of the concepts we have tried, the challenges we have faced, and the lessons we have learned

2 PROJECT GOALS & STATUS

Our goal is to develop a testbed, based on a low-cost commercial game, to enable research on the effects of culture on cognition and decision-making in individuals and teams. Moreover, the game is to be immersive to permit natural behaviors to emerge. Finally, the testbed is to be agnostic with respect to any theory of culture or personality. That is, the testbed should allow easy creation and manipulation of scenarios and situations which allow for the testing of any cultural or personality factor the experimenter wishes to test.

The game-based testbed is in a beta-level of development in July 2005 and should be ready for distribution in fall 2005. Preliminary data has already been collected as part of the testing and validation of the testbed.

The testbed software will be freely available to researchers, but the commercial game, *Neverwinter Nights* – Platinum Edition (Bioware 2004), on which it is based, must be purchased separately.

3 TESTBED GAME

The game-based testbed uses a "re-skinned" version of the commercial off-the-shelf medieval fantasy role-playing game *Neverwinter Nights* (Bioware Corp. 2004; Warren et al. 2004). The re-skinned version features a simulated modern cityscape with people wearing modern clothes and engaged in non-magical activities (See Figure 1 for a screen shot). This game was selected because it can be used to simulate cooperative team tasks and it facilitates scenario authoring and customization. The built-in game editor tools are designed to allow users to customize the size and contents of the game world, including synthetic character behavior and the creation of customized items.

Neverwinter Nights is a role-playing game. This means that the participant controls the actions of a single character during the game. Typical game modules are designed to last for dozens of hours, over multiple sessions of play, allowing the player to guide character development. During game play, events unfold continuously, with a real-time feel. There are no obvious "turns" like there are in games like chess or bridge.

Characters can move around the world, pick up and use items, go into buildings, read maps and signs, and interact with other characters. Other characters can be player characters (PCs) controlled by other participants or researchers,

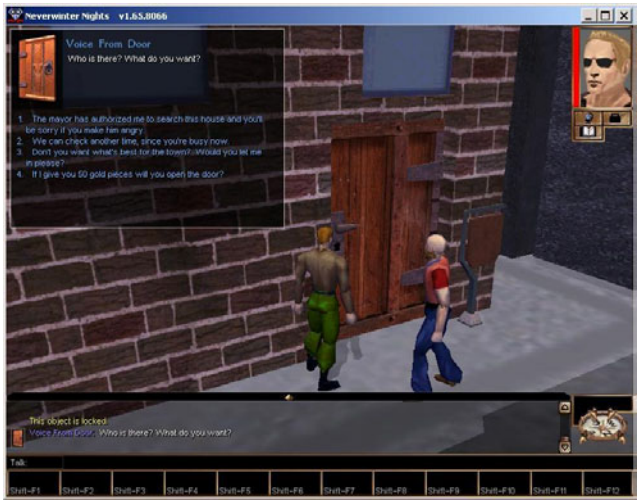

Figure 1: Screen Shot From The Testbed

or non-player characters (NPCs) controlled through computer programs or scripts. Although typical game modules involve combat as part of the adventure, the game-platform is not focused on fighting. This game is not a "first-person shooter" and does not emphasize quick reflexes or violence.

The *Neverwinter Nights* game interface is relatively easy for someone who has never played computer games to learn. It includes a journal for tracking information and a dynamic map.

The following in-game behaviors are part of *Neverwinter Nights.*

- Communications
  - Free form typed text between players
  - Selection from dialog menu choices
  - Emotes or animated expressions
  - Journal
  - Map
- Actions
  - Gain/lose items
  - Use items
  - Open doors or containers
  - Combat
- Movement
  - Entering areas
  - Changing location within an area

4 PILOT EXPERIMENT SCENARIO

In order to verify the testbed as a mechanism for experimentation and validate its utility as a means for conducting team research into culture and personality, we conducted a pilot

experiment examining the impact of culture and personality on team decision-making, coordination, and performance. We examined the behavior of homogeneous teams of Foreign and American males.

Because the evaluation study was designed to illustrate the types of task situations which can be implemented using the game-based testbed, we incorporated a range of elements allowing for the examination of culture on (1) team organization, including designated leader and leaderless conditions, (2) preferences for negotiation styles, (3) willingness to engage in tasks not related to the primary mission and the impact of the requesting individual's status, and (4) response to insults as moderated by insulter social status.

The study comprised 8 four-person teams of Foreign males and 8 four-person American male teams for a total of 64 individuals. Participants were active college students or college graduates, ages 18 to 35. American participants were required to have spent no more than six months between the ages of 2 and 18 residing outside the U.S. Foreign participants had to have spent no more than six months between the ages of 2 and 18 residing in the U.S, and have English language aptitude scores (TOEFL scores) greater than 550 (paper test) or 213 (electronic test), or have spent at least 24 months in the U.S. as a full time student or working in an English speaking job.

As part of the study, participants were asked to complete a background questionnaire, the NEO FFI Personality Inventory (McCrea and Costa 1987) and the Cultural Styles Questionnaire (Matsumoto 2000, Matsumoto and LeRoux 2003). The questionnaires were followed by a game tutorial and training session consisting of both task-work and teamwork skills critical to execution of the principle team task within the game. The training session was self-paced and lasted between one and two hours. Following training, teams were briefed within the game on the details of the upcoming task and then allowed to plan as a team how they would accomplish the task. The mission briefing and planning phases of the experiment lasted approximately 30 minutes. The execution phase of the task lasted 60 minutes, followed by a debrief questionnaire. Details of the pilot experiment scenario are described below.

In the basic experiment, team members are assigned roles (e.g., patrol leader, weapons specialist) depending on the experimental condition and the team is given the high-level task of locating and acquiring caches of weapons hidden within a town. The team is provided with equipment to help with the task (sensors of varying capabilities designed to help locate weapons caches and tools for opening doors and crates) and must decide how to allocate those resources. Additionally, team members have collaboration tools, allowing information to be shared between individuals and locations flagged or marked within the virtual environment.

Performance is team-based, with participants able to increase the team score by completing tasks which have rewards while managing costs and penalties. The primary task is recovering weapons from hidden caches – outdoor caches are worth 100 and indoor caches are worth 300. Subtasks have variable rewards. Penalties are associated with a number of activities, including entering private residences to search for caches (50 points penalty), regardless of whether permission has been obtained to enter the residence; opening containers that do not contain weapons (100 points penalty); and opening booby trapped crates (250 points penalty).

In the planning phase, the scenario begins with a mission briefing followed by several group planning tasks. Participants are first asked to distribute two high fidelity sensors, two low fidelity sensors, and one lock pick among the team. Then they are each required to select individual two-step negotiation strategies for entering private buildings. At each step, they can select to try to persuade, intimidate, or bribe the resident, or they can accept the resident's refusal and not gain entry. Next, participants are each given two unique intelligence reports about suspected weapons caches. Participants are free to share and organize these intelligence reports as they see fit. Finally, the team is shown a map of the town, and is asked to formulate an overall strategy, with one participant summarizing the group plan.

During the task-execution phase, the team is transported to a small town roughly 4 square city blocks in size (See Figure 2). The town is populated with townspeople (NPCs), some of whom can provide information useful to the team's task. While some of these NPCs will provide information only if asked, others will seek out team members and provide information. We push some information to team members in order to ensure the team receives a minimum number of intelligence reports. The information provided by NPCs can include locations where caches are suspected, locations where caches are not suspected, or townspeople who are likely to have additional information. A small number of these tips are false and some intelligence is time-sensitive (i.e., the player is informed that for the next five minutes, a piece of information is true).

All intelligence reports, or tips, received and some other types of information are recorded as entries in the participants' journals. Participants may use their journal-management tools to mark entries as completed. Participants may also use this tool to share a tip (sending a copy to a teammate) or assign a tip (sending a copy to a teammate and marking their own tip as completed).

Participants communicate with each other through text chat. Chat messages can be sent to specific teammates, simulating a point-to-point radio; or broadcast to all nearby individuals, simulating conversational speech.

Over the course of the experiment, team members are approached by NPCs requesting assistance with tasks unrelated to the main mission of locating hidden caches. These

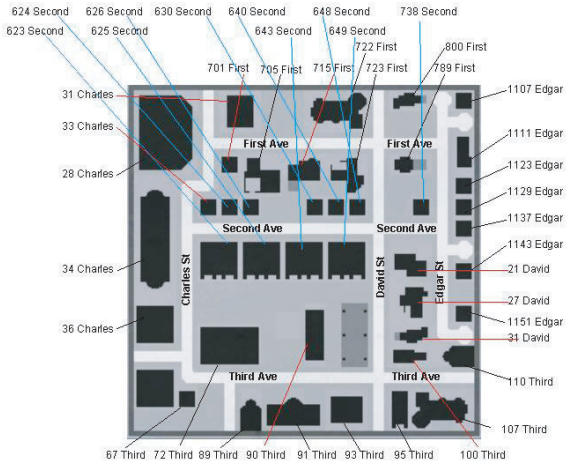

Figure 2: Pilot Experiment Scenario Map.

distractor tasks include requests to (1) plant listening devices in buildings (2) assist police in dealing with criminals (3) recover a stolen necklace (4) find a missing child. Participants may choose whether to accept these requests, whether to act on the requests, and whether to involve and inform their teammates.

Participants also engage in a number of other interactions with NPCs. For example, the doors to private buildings are locked and many have NPC residents with whom the participants may negotiate for entry. While there are a number of different negotiation strategies available to a participant, all strategies have a 50% chance of success except bribery, which is 100% successful, but costs 50 points, and backing off, which never results in opening a door. Participants have two chances to negotiate entry into a building. If unsuccessful, the door can only be opened using the lock pick. Additionally, as the participants move about the town, some NPCs of varying social status will approach and insult them. Participants can choose to escalate/de-escalate hostilities by responding respectfully, ignoring the insulter, or returning the insult.

5 USAGE NOTES & ADVANTAGES OF GAME-BASED EXPERIMENTATION

Development activities and preliminary testing have yielded numerous lessons learned and have strong implications for further development and use of such games for research. We review design and task considerations here and some lessons learned in the next section.

5.1 Eliciting & Examining Behaviors in Virtuality

Simulated tasks need only elicit the general behaviors of interest under the conditions of interest, and the simulation need not realistically reflect an actual situation (Warren and Riccio 1985). Our testbed allows us to examine behaviors such as reaction to insults, negotiation, group decision making, communications, etc. under conditions such as time pressure, limited information, and high workload, even though the actual simulated situation is more "abstract" than "realistic." The use of a game-based rather than realistic task set allows the experimenter more freedom to adjust the test conditions (so that the task is not too easy or too difficult) and to focus on particular situations or behaviors not easily posed or experienced in real life.

Another point is that a game-based testbed elicits behaviors, but these behaviors should be interpreted relatively rather than literally. For example, many people in real life might hesitate to accept a mission to explore the ocean in a leaky submarine. In a game, many of these same people would agree to the dangerous submarine mission, so their in-game decision cannot be interpreted literally.

However, if the population of participants was analyzed for how much information or motivation the game had to give the person before they agreed to accept the mission, relative differences would emerge between individuals and between cultures that would be expected to correlate to real-world differences in decision making.

5.2 Experimental Control in Free-Play Virtual Environments

In this section, we discuss some implications of doing experiments within a free-play environment such as a computer game. Overall, the experiment designer must keep in mind the inherent advantages and disadvantages of a free-play environment when designing the experimental scenario.

First, the experimenters should consider that there exist competing desires for constraining behaviors in order to produce situations amenable to quantitative analysis and hypothesis testing, while still providing more 'realistic' situations for participants to play out behaviors. Allowing free-play and creative solutions inevitably leads to behaviors and actions unanticipated by the experimenter. It is important that these behaviors, at a minimum do not negatively impact the experiment, and ideally are captured and evaluated using post-hoc analysis. For example, in our pilot scenario, a team of four participants is asked to search an area. It is desirable to determine whether the players walk around the area all together, individually, or in smaller subgroups. The participants are asked to state their plan, including a grouping decision. However, since during free-play the players might dynamically form different sub-groups than they stated in the plan, it is necessary to capture each player's location as a function of time, and to calculate their time-averaged proximity to the other players.

A second consideration is that game-based behavior experimentation is likely to require altering the scoring/win condition aspects of a commercial entertainment computer game if the experiment wants to examine the natural behav-

iors of participants. This is because an entertainment game's scoring system will often strongly/consistently reward certain types of behaviors, and rely on the players' knowledge of this in-game reward structure. However, an experiment may want to observe someone's choice of behavior under conditions where the choices have unknown/uncertain likelihood of rewards. Often these sorts of uncertain conditions are not "fun" and would not be included in an entertainment game scenario.

In our pilot experiment, we use the amount of gold as a proxy for the team's performance score. Team members are informed that they will be rewarded with gold whenever they find the objects they are searching for, and they will be penalized gold for making certain mistakes during the search process. Players are told that items inside buildings are worth more gold, but are riskier to pursue because they may be trapped. In a commercial game, it would be likely that pursuing the riskier option would result in a higher score, and that the instructions would encourage players to take risks. In our game, we have set the rewards and risk probabilities so that pursuing either the riskier or safer course of action will, on average, result in the same score and the instructions are designed to be neutral towards both options so that observation of what the player decides to do will reflect an actual behavioral tendency rather than a behavior intended to maximize game performance.

The random factors that may exist in entertainment games in order to increase the fun aspects of the game may need to be reduced, controlled, or eliminated within an experimental paradigm. For a game-based experiment, reproducibility and consistency between participant sessions may be critical. Game-based platforms may not provide access to critical factors such as seeds for random number generation making random factors even more difficult to manage.

In our pilot scenario, one decision we are examining is the players' choice of negotiation strategy when trying to convince residents to permit them to search houses. Each negotiation encounter presents the player with four possible choices of phrases. There is one phrase each that uses intimidation, persuasion, bribery, or acquiescence. The exact phrase is randomly selected from a pool of phrases representing each strategy option, and the order that the choices are presented is randomized. This is an example of a situation which would typically not be random in a game, but needs to be random in an experiment.

There are also cases where a situation in a game would be randomized, but needs to be controlled for an experiment. For example, our pilot experiment features several trapped items that cause the players to lose points. In a commercial game, each time a player picked up an item, the presence of a trap would be determined probabilistically. However, to prevent some teams from becoming discouraged by encountering traps early in the experiment, our items are set up with deterministic traps, so that the third item picked up by the group will have a trap no matter which item that happens to be.

5.3 Transforming & Repurposing a COTS Game

Repurposing a commercial game and transforming it into a testbed for experimentation presents a number of challenges not typically encountered when developing software from scratch. In most situations the developer may be both underconstrained and overconstrained by the game's capabilities. Not having access to the underlying source code often makes it difficult to add new features to the game. For example, control over interface elements can be difficult, if not impossible. This lack of control can lead to the opposite problem of underconstraining user capabilities, making it difficult to keep participants from performing certain actions irrelevant, or even detrimental to the experimental condition. In some cases clever tricks and "hacks" can be exploited to provide solutions to the aforementioned problems. For example, in *Neverwinter Nights*, the right mouse button brings up a control menu which provides more advanced capabilities and user control. By using a mouse driver which provides control over the mouse button settings, we were able to eliminate access to these controls. Other tricks have involved replacing screen graphic images with those more appropriate for the experiment, or even removing screen icons which are not experimentally appropriate.

The architecture of the game and the degree of authoring capabilities built into the game influences the degree to which you can customize the game. In most situations, there will always be capabilities that would greatly improve the platform but are simply not possible without some cooperation with the developer. In our testbed, the inclusion of timestamps as part of event logs was a critical new feature added to *Neverwinter Nights* that greatly improved our capabilities and reduced our development effort.

6 LESSONS LEARNED

In addition to the opportunities and advantages of research using a virtual world, there are a number of pitfalls. There are lessons to be learned not only in experimentation and data collection but also in training and even in legal issues.

6.1 Training to Use the Environment

Over the course of usability testing, we had 16 teams run through our scenario. We used their experiences to refine both the experimental scenario and the in-game tutorial for familiarizing the participants with the game interface. The design of the familiarization and training session was critical to ensure that every study participant had adequate ability to play the game before going into the experimental scenario.

We found a steeper learning curve than expected for the average individual to learn to play the game, in spite of the fact that *Neverwinter Nights* is a relatively easy game interface to learn compared to many first person shooter, strategy, or building games. We also simplified this interface for use in the experiment. Thus, any researchers planning to use a commercial game platform should anticipate a significant need for participant training. This is partially due to the fact that the test population of interest is not restricted to gamers and hence we did not require that participants be gamers. For any game experiment, basic computer literacy is strongly suggested; we believe that participants who are not familiar with basic mouse and keyboard usage require considerable training in order to learn to play the game.

Along with the definite need for training, we observed a large variability within the population in terms of time required to learn the interface. Gamers take naturally to the environment, whereas non-gamers can have some difficulty mastering game controls. This means that for a self-pace tutorial session, some participants finished within 1 hour while others took 2 hours or more. This can create a situation where some participant must wait for the others to finish training before the group experiment. Thus, we added some activities to the end of the single-participant training session which would entertain the faster players without giving them significant performance advantage in the coming experiment.

Without some form of selection filtering for participants, some players will have considerable difficulties with the task. When performing team-based experiments, having a single outlier in the group often will significantly skew team behaviors and performance. Each target population will differ in its game experience, so different strategies might be necessary to prevent the inclusion of a participant who is unable to capably play the game. One method we used was to have an extra participant begin game training, and to select the top performers to continue with the team experiment.

Especially while players are learning a game, there can be many unexpected situations that detract from or add noise to the experimental manipulations of interest. Usability testing is a must so that these situations can be handled. For example, we found that some participants might neglect to pick up equipment that they needed later. Thus, it was necessary to check whether the items had been picked up, and to direct the player to go back for items if necessary.

Overall, the training module required more development effort than anticipated. Our goal was to produce a training module that did not require individual instruction or significant intervention by a human experimenter. We found that building this level of automated training into the experiment required careful design and required extensive testing and redesign.

6.2 Data Collection

While providing a rich environment for experimentation, the free-play aspects of *Neverwinter Nights* and the testbed increases the complexity of data collection and logging within the testbed. As testbed developers, we recognize that experimental participants are highly likely to behave in diverse and unexpected ways and that it is critical to capture these behaviors for evaluation. For example, despite there being no utility for changing an avatar's attire, we have experienced situations in which players have changed or removed clothing from their avatar. Players have experimented with constructing their own equipment, engaged in unexpected combat with NPCs (e. g., attacking instructors), etc. In order to capture the kinds of unexpected behaviors highlighted here, we now log behaviors at a general level in order to capture all possible activities.

While logging low-level events is one method for ensuring that all important events are captured, it typically increases the complexity of the post-processing required in order produce the desired analysis. For example, collecting data at the level of keystrokes and mouse movements, while complete, is extremely difficult for a researcher to utilize. Therefore, we log information at multiple levels of granularity, capturing typical actions and behaviors in a format easily analyzed by a researcher, but also capturing a lower-level running log of game events, actions, and in-game communications. This allows a researcher to analyze the lower level events, depending on the needs of the experiment, while hopefully minimizing the need for additional effort extracting data from lower level events.

Despite the advantages of free-play, there are many situations in which an experimenter wants to experimentally control and manipulate a situation in order to ensure the appropriate collection of data. For example, to assure that participants receive certain information during the course of the experiment, we "push" the information out to a participant using an NPC which will seek out a player and force a conversation. Furthermore, it is possible to collect multiple datapoints on a factor of interest by either building in multiple situations which may be encountered by a participant, or as previously described, forcing situations to occur over the course of the experiment.

6.3 Experimentation

Using a rich game environment as the basis for an experiment provides both opportunities and additional challenges for the experimenter. Rich virtual environments and open-ended interactions allow participants to become immersed, and perhaps make choices that are relevant to the experimental

hypothesis, in a scenario where these decisions form part of a naturally conceived activity. The challenges of using a game include the fact that many of the participant's actions may not be relevant to the experimental hypothesis, and worse, that much of the data may have to be aggregated, scored, or otherwise processed in order to uncover its relevance. Also, the scenario and the virtual world may be difficult to design in such a way that the critical actions of the participant are well defined enough to relate to the hypothesis. For example, if one is studying leadership, a questionnaire can refer specifically to aspects of leadership abstractly (for better and for worse, as far as self report information goes); in a game based experiment, the experimenter must judge leadership from the virtual actions that the participant takes.

In order to achieve the benefits and to mitigate the extra work that a game environment entails, the researcher should design the experiment in a way that is consistent with the "gameplay" that the game supports. For example, in the testbed we designed the pilot experiment to engage participants in team interaction, physical motion, object manipulation, interaction with computer controlled characters (NPC's), map-based navigation, and the use of virtual props. All of these were natural activities in the game before we began our work.

We posit that using a game-based testbed requires a somewhat new mindset towards experimental design. In traditional laboratory experiments the hypothesis drives the experiment design in an entirely top down fashion. (Though questions of cost, ethics, participant availability, and lab capabilities remain critical constraints). In field-based experiments, some portion of the real world is used and the experimenter must be acutely aware of how the experiment must fit into the world and take advantage of what is occurring there. Game-based experiments fall somewhere in between: the experimenter must design an experiment that fits into a virtual world where certain kinds of things can happen and others cannot, but the experimenter has a great deal of ability to tailor that "world" to his or her own design.

Pragmatically, we suggest that it is important that the experimental designer be deeply familiar with what is possible and natural in the underlying game world. Activities that are not well supported by the game or that exclude so much of the game world that the gameplay is disrupted, will not produce an immersive experience and will not leverage the testbed's strengths. For example, if we had wanted to design our pilot experiment to consist of only menu-based dialogs with NPC's, it may have been better to use a simpler experimental toolset. If experiments disregard the natural gameplay supported by the game, they will, at best, be awkward and expensive implementations that could have been done more simply from scratch. At worst, they will be so cumbersome for participants that any results will be meaningless.

6.4 Legal and Intellectual Property Issues

Game developers are increasingly creating games with improved capabilities for producing user-generated content, making it possible to customize games for experimentation and training applications. Despite the increased ease with which a user may develop content, there are often legal issues which must be resolved when commercial entities, rather than individuals, are involved. Even games designed for modification may contain clauses in the EULA (End User License Agreement) which are problematic for situations in which a developer is paid to produce a modification, and this may be the case even when the resulting modification will not be sold. Furthermore, the legal arrangement between game developer and game producer can be unclear to a third-party, making it difficult to ascertain who has legal authority over the issue in question. Make sure to acquire appropriate permissions early.

Further, a number of assessment instruments useful to researchers using the testbed may be copyrighted and therefore cannot be used without appropriate licensing arrangements. While we have currently integrated several licensable assessments instruments into the testbed (e.g., NEO-FFI, Cultural Styles Questionnaire) we will not release these assessments as part of the testbed. Only after appropriate permissions are granted to a research site will we provide the data enabling those questionnaires to be run from within the testbed.

7 FUTURE EXPERIMENTS IN CULTURE AND PERSONALITY

We are currently in the process of analyzing the results from the pilot experiment comparing behavioral differences between American and foreign teams. We are also in the process of designing and developing an experiment for the NATO Supreme Allied Command Transformation HQ, Futures and Engagement, Concept Development and Experimentation project entitled Leader and Team Adaptability in Multinational Coalitions (LTAMC), to be run at US and international sites.

The goal of the LTAMC project is to enable rapid formation of multination Joint Task Forces with the capability for seamless planning and execution, capable of operating across the full spectrum of Alliance operations more effectively and efficiently than they currently do. While there are difficulties inherent in these staffs (e.g., personnel experience, knowledge, and motivation) that NATO has taken steps to address through selection and training, the difficulties arising from characteristics and traits of individuals on those staffs that are rooted in national culture and individual personality continue to hinder effectiveness and efficiencies in team performance (Sutton and Pierce 2003).

The game-based experiment under design concentrates on cultural and personality effects on information sharing, situation awareness, and division of responsibilities in four-person teams. BBN developers are modifying game play to meet LTAMC objectives, something any researcher will be able to do with relative ease using authoring tools. Potentially, international sites may compare team communication patterns between non-native English speakers asked to communicate in English versus their native language.

It is possible that there are cultural assumptions built into the existing maps, text and action choices allowed to participants which may impact results. Pilot studies will be conducted in the individual national studies to identify participant reactions to the game to assess cultural assumptions that are unknowingly built into the game.

The research potential of this modified commercial role-playing game to assess the extent to which culture and personality affect planning, execution, feedback, interaction styles, tempo, information management, situation awareness and shared understanding, and decision making is extensive. For example, a two-person team variant could be developed. This would enable pair-wise exploration of culture cooperation and conflict situations. Manipulating group composition by nationality may reveal "best case" models for multicultural decision making teams. Social Network Analysis (SNA) could be employed as an analytical tool to study who communicates with whom in game play, and to what end, in order to determine whether there is a certain combination of personality traits that contribute to individuals who quickly become gatekeepers or conduits to communication in teams. One may want to investigate how team behavior changes when the nationality of the leader changes or study emergent leader behavior associated with the interaction of personality and culture.

To assist the reader in identifying research areas of interest that can be examined using the game-based research venue, the following are several research questions looking for answers: (a) How does culture or personality impact team adaptability (i.e., will initial strategies adopted by a team be altered based on performance or will the team maintain their strategy throughout the game)? (b) How does intent affect performance on multicultural teams (i.e., does a clear, explicit intent result in better performance)? (c) Are issues of trust and confidence related to nationality and to what extent do these variables affect information sharing and performance? (d) Do teams' subjective perceptions of their success or reasons for their success match with objective criteria and measures (i.e., do some cultures "undervalue" [from our cultural viewpoint] factors like communication, leadership, situation awareness?)? (e) What tasks are better done by culturally homogeneous teams or by culturally heterogeneous teams? (f) What are common or repetitive problems in the planning process that can be attributed to culture or personality? A recommended source of inspiration can be found in work done by the US Army Research Laboratory at Stabilization Force headquarters in Bosnia-Herzegovina (Sutton and Pierce 2003). Their findings on the impact of culture on teamwork at the operations level resulted in development of a framework for understanding cultural diversity in cognition and teamwork which can be used to guide research in the cultural domain.

8 CONCLUSIONS

Simulation of social scenarios using a commercial game-based testbed can be of significant value in research on cross-cultural effects on cognition and teamwork. Due to the immersive and action-demanding nature of the gameplay, the results can be very informative. Also, a simulation task can enable new research on mixed-culture teams. For example, it could be used to assess behavior at the operational level in the planning stage of the game and in the execution of the game. Research could produce ways of identifying individuals who might be particularly successful working in multinational team environments. It could further understanding of the expectations and role responsibilities of leaders, from both a leadership and team member perspective. By increasing understanding of communication patterns in teams of particular cultures, suggestions could be developed for how to adapt individual communications patterns to achieve better team performance. Finally, research results could be used to derive requirements for command and control systems.

ACKNOWLEDGMENTS

BBN Technologies' work on this project was performed for the Defense Modeling and Simulation Office (DMSO) under Contract FA8650–04–C–6437 administered by the Air Force Research Laboratory. Rik Warren (Air Force Research Laboratory) and Janet Sutton (Army Research Laboratory) are the Program Managers as agents of DMSO. All opinions are solely those of the authors and not official Department of Defense positions.

REFERENCES

Bioware Corp. [Computer Game Software]. 2004. *Neverwinter Nights* – Platinum Edition. New York: Atari Interactive.

Brouwers, S.A., D.A. Van Hemert, S. M. Breugelmans and F.J.R. Van de Vijver. 2004. A historical analysis of empirical studies published in the *Journal of Cross-Cultural Psychology* 1970–2004. *Journal of Cross-Cultural Psychology* 35: 251–262.

Diller, D.E., W. Ferguson, A.M. Leung, B. Benyo and D. Foley. 2004. Behavior modeling in commercial games. In *Proceedings of the 2004 Conference on*

Behavior Representation in Modeling and Simulation (BRIMS), May 17–20, 2004. Available online via <http://www.sisostds.org/index.php?tg=fileman&idx=get&inl=1&id=2&gr=Y&path=CGF-BR%2F13th+CGF-BR%2F13th+CGF-BR+Papers+and+Presentations&file=04-BRIMS-079.pdf>

Fischer, R. 2004. Standardization to account for cross-cultural response bias. *Journal of Cross-Cultural Psychology* 35: 263–282.

Matsumoto, D. 2000. *Culture and psychology* (2nd edition). Pacific Grove, CA: Brooks Cole.

Matsumoto, D. and LeRoux, J.A. 2003. Measuring the psychological engine of intercultural adjustment: the Inter-Cultural Adjustment Potential Scale (ICAPS). *Journal of Intercultural Communication* 6: 27–52.

McCrae, R. R. and Costa, P. T., Jr. 1987. Validation of the five-factor model of personality across instruments and ob-servers. *Journal of Personality and Social Psychology* 52: 81–90.

Sutton, J.L., and Pierce, L.G. 2003. A framework for understanding cultural diversity in cognition and teamwork. *Proceedings of the 8th International Command and Control Research and Technology Symposium*. Available online via <http://www.dodccrp.org/events/2003/8th_ICCRTS/Pres/track_1/1_1300sutton.pdf>

Warren, R. and G.E. Riccio. 1985. Visual cue dominance hierarchies: implications for simulator design. *Transactions of the Society for Automotive Engineering* 6: 937–951.

Warren, R., J. Sutton, D. Diller, W. Ferguson and A. Leung. 2004. A game-based testbed for culture & personality research. In *Proceedings of the NATO Modeling and Simulation Group — 037 Workshop on Exploiting Commercial Games for Military Use* held at The Hague, The Netherlands, October 2004.

AUTHOR BIOGRAPHIES

RIK WARREN is a Senior Research Psychologist in the Cognitive Systems Branch of the Human Effectiveness Directorate of the Air Force Research Laboratory. His re-search interests range from theoretical issues in the perception and control of self-motion and the design of flight simulation displays to understanding the role of culture and language in cognition and decision making. He serves on the editorial boards of Ecological Psychology and the International Journal of Aviation Psychology. His email address is <Rik.Warren@wpafb.af.mil>.

DAVID DILLER is a Senior Scientist at BBN Technologies in Cambridge, MA. He holds an M.S. in Computer Science and a joint Ph.D. in Cognitive Science and Cognitive Psychology from Indiana University. His current focus includes cognitive modeling, mixed-initiative agent-based systems, and simulation-based training applications. Recently, Dr. Diller has been involved in a number of projects utilizing commercial game technology for training applications. He is currently a Co-PI of the Cultural Modeling Testbed, the project described in this paper. His email address is <ddiller@bbn.com>.

ALICE LEUNG is a scientist in the Intelligent Distributed Computing Department of BBN Technologies. Her current interest is in harnessing games for behavior research. Previous projects include software for military logistics planning using distributed agent technology, small-scale simulation of information transfer economics, and infrastructure to support training through games. Her email address is <aleung@bbn.com>.

WILLIAM FERGUSON is a Division Scientist at BBN Technologies. His background is in artificial intelligence, simulation, computer-based training, and commercial game technology. He is currently Co-PI of the integration effort under DARPA's DARWARS Training program. He also serves as Co-PI of the Cultural Modeling Testbed, the project described in this paper. He worked for many years as the technical lead for the Analysis of Mobility Platform (AMP) project, USTRANSCOM's transportation programmatics modeling system. His email address is <wferguson@bbn.com>.

JANET SUTTON is a Research Psychologist in the Information Systems Branch of the Human Research and Engineering Directorate of the Army Research Laboratory. She offices at Fort Sill, Oklahoma where she maintains a program of research on developing adaptability in leaders and teams in addition to providing human factors support to the US Army Field Artillery School and the Depth and Simultaneous Attack Battle Lab. Dr. Sutton's primary research areas of interest are decision-making and teamwork in military command and control. Her focus is on development of methods and systems to promote rapid formation of effective teams performing command and control functions in Joint, Interagency, and Multinational (JIM) environments. Her email address is <janet.sutton@us.army.mil>.