TASP: A Flexible Alternative for Team Training and Performance Research

Amy E. Bolton, Gwendolyn E. Campbell, Wendi L. Buff
NAVAIR Orlando Training Systems Division
Orlando, Florida
amy.bolton@navy.mil, gwendolyn.campbell@navy.mil, wendi.buff@navy.mil

Ling Rothrock, Damodar Bhandarkar, Hari Thiruvengada, Urmila Kukreja
Pennsylvania State University
State College, Pennsylvania
lrothroc@psu.edu, dnb133@psu.edu, hzt103@psu.edu, ulk100@psu.edu

ABSTRACT

Recognizing the critical role of teams in the U.S. military, the Department of Defense has been a strong supporter of research into team training and performance. In this paper we will describe a newly developed simulation-based testbed suitable for supporting a wide variety of programs within this research area. The freely-available software suite includes: (a) a scenario authoring tool, (b) a simulation testbed for a complex, three-role team task, (c) three synthetic agents, each capable of playing one of the roles in the testbed, and (d) tools for performance recording and assessment. The functional specifications of the software suite were developed collaboratively by a group of industrial engineers, computer scientists, and research psychologists. The task is a complex, dynamic team task designed to simulate air defense operations. The three roles are distinct, with separate interfaces and responsibilities, and are organized into a simple command hierarchy of one leader and two supporting team members. Scenarios created in the authoring tool are deterministic, allowing for a high degree of experimental control. Scenario difficulty can be controlled along multiple dimensions, including raw data volume, resource allocation complexity and information ambiguity. The synthetic agents may be used as teammates for human participants during data collection. These agents are capable of performing competently on previously un-encountered scenarios within a broad (but not infinite) problem space. In addition, the experimenter may degrade the agents' performance on an event-by-event basis through flags set in the scenario authoring tool. The testbed captures and records multi-modal performance data at multiple levels of detail ranging from time-stamped keystrokes, mouse actions and speech interactions to an analysis of actions within a "Windows of Opportunity" assessment framework. Cooperative use of this testbed within government, academia, and industry will generate a more reliable and valid understanding of team training and performance.

ABOUT THE AUTHORS

Amy E. Bolton is a Research Psychologist at NAVAIR Orlando Training Systems Division. She holds a M.S. in Cognitive Human Factors and is a doctoral candidate in the Applied Experimental and Human Factors Psychology Ph.D. program at the University of Central Florida. Her research interests include the application of intelligent agents in individual and team training for dynamic tasks.

Gwendolyn E. Campbell is a Senior Research Psychologist at NAVAIR Orlando Training Systems Division. She holds an M.S. and Ph.D. in Experimental Psychology from the University of South Florida and a B.A. in Mathematics from Youngstown State University. Her research interests include the application of human performance modeling techniques within training systems and the development of a cognitively based science of instruction.

Wendi L. Buff is a Research Psychologist at NAVAIR Orlando Training Systems Division. She is currently a doctoral candidate in the Applied Experimental and Human Factors Psychology Ph.D. program at the University of Central Florida. Her research interests include applications of human performance modeling and intelligent tutoring systems (ITS) to individual and team training.

Ling Rothrock is an assistant professor at the Harold and Inge Marcus Department of Industrial and Manufacturing Engineering at Penn State University. He received his Ph.D. in Industrial Engineering from Georgia Institute of Technology in 1995 and served as an officer in the U.S. Army until 2000. His research areas include judgment and decision making, human-machine performance modeling, and human-in-the-loop simulations. He is the director of the Human Performance Assessment and Modeling (HPAM) Laboratory.

Damodar Bhandarkar is a graduate student in the Harold and Inge Marcus Department of Industrial and Manufacturing Engineering at Penn State University. He received his M.S. from the Human Factors Department at Wright State University in 2001. He is a research assistant in the HPAM laboratory. His research interests are behavioral decision making and cognitive modeling.

Hari Thiruvengada is a graduate student in the Harold and Inge Marcus Department of Industrial and Manufacturing Engineering at Penn State University. He received an M.S. in Computer Science from the University at Buffalo. He is a research assistant in the HPAM laboratory. His research interests include applications of computer science, man machine systems, human performance modeling and human in the loop simulations to individual and team training.

Urmila Kukreja is a Research Assistant at the Human Performance and Assessment Modeling Lab at Penn State University. She is pursuing her M.S. in Computer Science and her current research involves modeling human behavior to improve human-computer interaction.

TASP: A Flexible Alternative for Team Training and Performance Research

Amy Bolton, Gwendolyn Campbell, Wendi Buff
NAVAIR Orlando TSD
Orlando, Florida
amy.bolton@navy.mil, gwendolyn.campbell@navy.mil, wendi.buff@navy.mil

Ling Rothrock, Damodar Bhandarkar, Hari Thiruvengada, Urmila Kukreja
Pennsylvania State University
State College, Pennsylvania
lrothroc@psu.edu, dnb133@psu.edu, hzt103@psu.edu, ulk100@psu.edu

INTRODUCTION

Recognizing the critical role of teams in the U.S. military, the Department of Defense has been a strong supporter of research into team performance and training. However, conducting useful research in this area requires testbeds suitable for supporting a wide variety of team training experiments. Towards this end, a group of industrial engineers, computer scientists and research psychologists, funded by the Office of Naval Research, set out to design such a testbed. The purpose of this paper is to describe the resulting testbed and discuss the specific features designed to support the conduct of team performance and training experiments. It is hoped that other researchers in government, academia and industry may find that this freely available testbed will be useful in supporting their research.

REQUIREMENTS

Before designing and programming the experimental testbed, our team invested a significant amount of time and effort into identifying the requirements for the testbed. It should be noted that we were not only designing a computer program, but we were also designing the actual team task (i.e., the roles and responsibilities of the team members) to be instantiated in the simulation-based testbed. Knowledge of basic experimental methodology, as well as knowledge of theories of team performance and training, helped to establish many of the task and experimental requirements. In addition, members of the project team have experience studying individual decision-making performance in a single person simulation-based testbed, which proved to be invaluable when identifying requirements for the team testbed.

Task Requirements

Probably the most fundamental and important requirement for the task is that it be a rich, complex and challenging team task. According to a standard definition of a "team" (vice a "group"), successful mission accomplishment requires that each team member be responsible for a set of tasks that are both independent and interdependent (Fleishman & Zaccaro, 1992). In other words, each team member should have a set of unique responsibilities that partially contribute to the success of the team's overall mission and accomplishing the team mission successfully should require coordination among the team members. Beyond these two requirements, we decided to model our team task on a real world, military team task, which imposes the additional requirement that the team be organized with a hierarchical command structure.

Experimental Requirements

While we decided to fashion our task on a military domain, practical issues associated with being able to conduct many research studies relatively easily led us to add the requirement that the task be simplified enough to make it feasible to train college students as experimental participants. As experimenters, we also wanted significant flexibility and control over as many aspects of the scenarios as possible. However, we decided to use deterministic scenarios, so that each team working through a scenario would be exposed to the same set of events. In other words, the behavior of an entity in the scenario would be scripted during scenario generation, and would not depend on the behavior of a team member during a simulation run (with only a few limited exceptions).

One class of studies that we are pursuing include investigations of the capability of Intelligent Agents (or Human Behavioral Representations) to serve as synthetic teammates during a team training exercise. Another class of studies that we are pursuing include investigations of the capability for Intelligent Agents to represent expert performance and assess the performance of novice trainees in an attempt to provide effective training feedback. These two research goals imposed the requirement that the watchstations for teammates incorporate software "hooks" necessary to allow an independent software module to communicate bidirectionally with the simulation.

Finally, in order to conduct our research, we need a system that automatically collects a rich, multifaceted, multileveled set of performance data. Ideally, there should also be some automated analyses built into the system. However, the lowest level data necessary to completely recreate every aspect of a team's scenario run must be preserved in case a researcher generates a new way to analyze the data at some future date.

System Requirements

A large part of teamwork is conducted via verbal communications. This created a requirement for the system to have the capability to capture and "understand" speech based on grammar templates constructed using Nuance Grammar Builder® tool. However, as this is a relatively risky undertaking, we also needed the capability to allow for a portion of each team member's tasking to be accomplished via mouse and keyboard. In addition to interaction requirements, there are other system requirements associated with robustness, reliability and processing speed of the overall testbed, which are met by adopting the correct software and hardware architectures.

Based on the above listed requirements we have developed the Team Aegis Simulation Platform (TASP). Next, we will describe how each requirement was realized in the design of the system.

TEAM AEGIS SIMULATION PLATFORM

Task Description

To meet the requirement of having a rich, complex and challenging task that requires a team of individuals to perform, and the requirement to model this task after a military task, we chose to loosely represent the duties performed by a Naval Air Defense Warfare (ADW) team stationed in the Combat Information Center (CIC) of a Navy ship. Because we did not want to simulate the actual ADW team that is relatively large, we simplified the duties to allow the task to be performed by a team consisting of only three individuals. We were able to distinctly define the individual roles and provide each team member with a separate interface and separate sets of primary responsibilities. The team was organized into a command hierarchy of one leader and two supporting team members. The overarching team responsibilities are protecting ownship and other friendly assets in the battlegroup by monitoring the airspace, identifying unknown air contacts, and taking both defensive and offensive actions as prescribed by the Rules of Engagement (ROEs).

The team leader, the Anti-Air Warfare Coordinator (AAWC), is responsible for monitoring a radar scope and identifying all air contacts that appear on the screen. The radar scope along with the air contacts are pictured on the right side of the interface displayed in Figure 1. To make identifications, the AAWC may use a variety of information about the unknown air contacts displayed in the Character Readout (CRO) located in the upper left-hand corner of the screen. Examples of this information include altitude, speed, range from ownship, point of origin, and direction of travel. The AAWC uses either the keyboard or a mouse to interact with the menu displayed across the bottom of the interface.

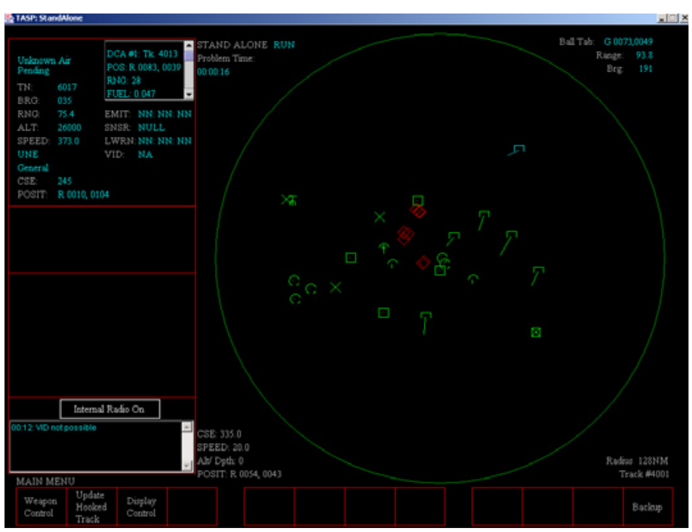

Figure 1. AAWC Screen Shot

The AAWC coordinates with two supporting team members, the Sensor Operator (SO) and Air Intercept Coordinator (AIC), to obtain additional information prior to making identifications. The SO is responsible for correlating, cataloguing, and transmitting sensor information about various air contacts using the panel below the CRO as depicted in Figure 2. The AAWC may then use the sensor information provided by the SO to help make the final aircraft identification.

The AIC is responsible for monitoring, managing, and protecting friendly air assets called Defensive Counter Aircraft (DCA). DCA may be ordered by the AIC to vector to an unidentified air contact and make a definitive Visual Identification (VID). The control of DCA is accomplished using the panel under the CRO in Figure 3. For more detail on the individual responsibilities of the team members, see Bolton, Dorsey, and Campbell (2004).

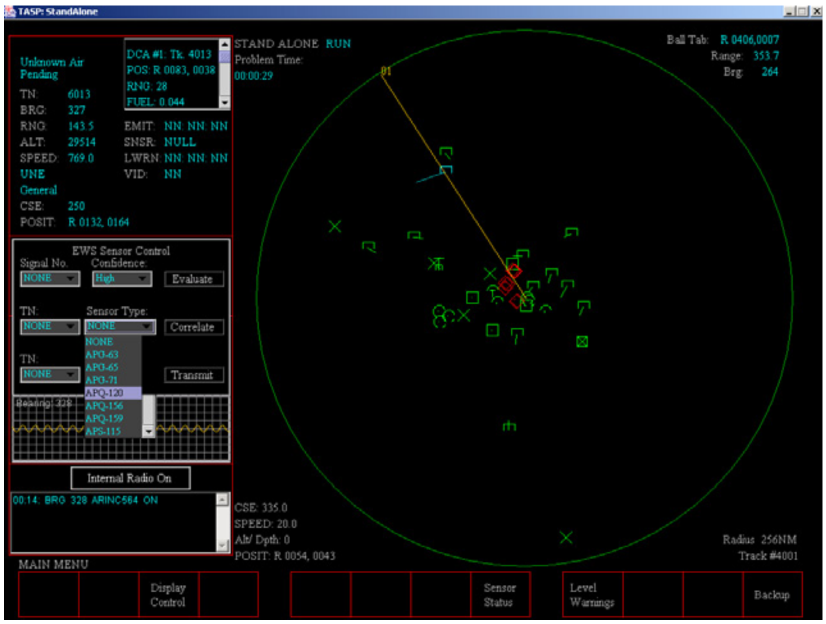

Figure 2. SO Screen Shot

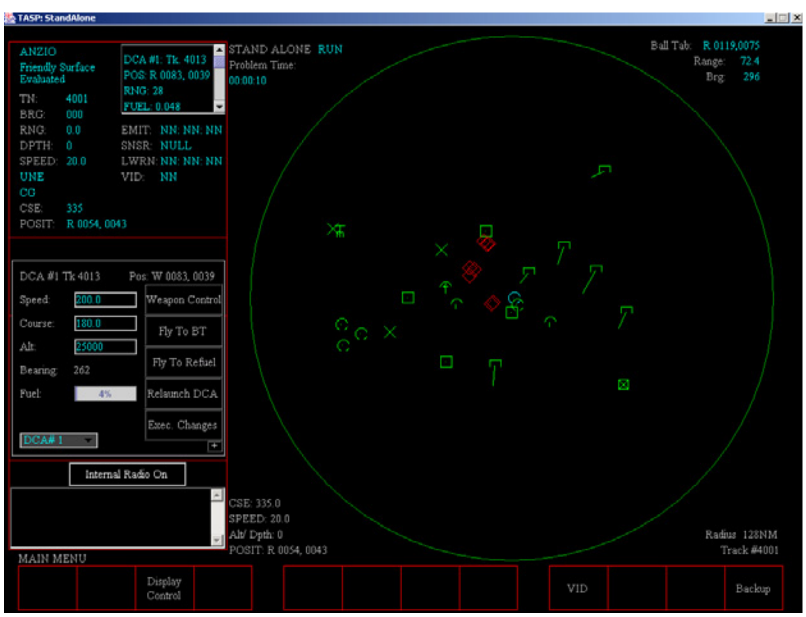

Figure 3. AIC Screen Shot

It should be noted that, while these tasks bear a superficial resemblance to real responsibilities of an ADW team, they have been dramatically simplified, and are suitable for use with college students. Although, of course, training will still be required.

Experimental Capabilities

Scenario Generation

To meet the experimental requirements associated with scenario generation capability, a scenario development authoring tool was developed which maximizes authoring flexibility while providing a set of limitations or constraints to maintain a degree of realism. This tool, ScriptMaker, provides the capability for any researcher to develop their own scenarios in order to meet the specific requirements of their research objectives. In addition, scenarios created in the authoring tool are deterministic, allowing for a high degree of experimental control. Scenario difficulty can be controlled along multiple dimensions, including raw data volume, resource allocation complexity and information ambiguity. For instance, pacing, workload, stress, interpositional conflict, and ambiguity can be manipulated during scenario development.

Connectivity for Intelligent Agents

Another experimental requirement was for the capability to allow testbed connectivity with Intelligent Agents. The TASP simulation has software hooks at each of the three workstations that are designed to support the connection of any software module, not simply Human Behavioral Representations (HBRs). Moreover, because a bidirectional communication capability exists for each hook, software modules may either run simultaneously with the human operator (e.g., serving as an over-the-shoulder instructor) or replace the human operator entirely (i.e., serve as a teammate).

Synthetic Teammates

As we anticipate conducting research in the area of synthetic teammates, we are funding the development of HBRs for each of the three roles in the TASP testbed. These HBRs will be able to perform competently on previously un-encountered scenarios within a broad (but not infinite) problem space. As part of the scenario development process, the behavior of these teammates can be controlled through a feature embedded in ScriptMaker. For example, the experimenter may degrade the agents' performance on an event-by-event basis through flags set in the scenario authoring tool. This ScriptMaker feature allows the scenario developer to indicate exceptions to competent HBR performance on particular events. Specifically, the flags inform the HBR, when faced with the flagged event, to either make a performance error, or simply fail to take any action at all. From a team training standpoint, this provides the opportunity for a human trainee to practice supporting team behaviors, such as backup and error correction, during the exercise with the simulated teammates. From an experimental control standpoint, this flagging feature provides the ability for the experimenter to have control over the teamwork skills exhibited by the agents (i.e., backup and error correction). See Buff, Bolton, and Campbell (2003) for a more complete discussion of this approach.

Data Collection and Analysis Tools

Finally, a key experimental requirement for the testbed focused on its data collection and recording capabilities. TASP includes the capability for automated data capture, performance measurement, and performance assessment. Raw data are saved in text files, and the TASP software suite includes a program called Converter, which translates these files into a set of linked Microsoft® Access tables. At the lowest level, every keystroke and mouse action is recorded and time stamped. While some of these user actions are abortive, many of these mouse and keyboard inputs are components of meaningful sequences. Thus, a second level of data is stored, summarizing the meaningful actions taken. Some actions can only be recorded as "taken", but the validity or accuracy of other actions (such as the identification assigned to an air contact) can be assessed. For these cases, validity is assessed and recorded. Finally, all of the characteristics of all of the entities in the environment are recorded at six-second intervals, which is the frequency with which this information is updated by the system. This allows the researcher to recreate the context in which each action was taken.

Figure 4, presenting real performance data from a longitudinal study currently underway at NAVAIR Orlando, provides a typical, easy to generate graph. Performance on many different actions could be represented. In this case, the dependent variable being plotted is the number of correct identifications completed per scenario. The independent variables are amount of practice (represented in days) and the presumed level of difficulty of the scenarios. The results suggest, for this team, that the scenarios that were designed to be of moderate difficulty were actually not much more difficult than the easy scenarios. The graph also shows the team becoming more capable until, approximately half way through the experiment, its performance on the difficult scenarios becomes indistinguishable from its performance on the easy and moderate scenarios. Finally, these data suggest that there is something unusual about the easy scenario run on day 7.

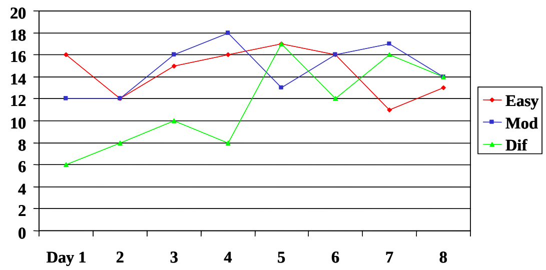

Figure 4. Number of Air Contacts Correctly Identified by Scenario Difficulty and Day

In addition to this inclusive data collection capability, TASP incorporates some powerful automated performance measurement capabilities. Specifically, the Windows of Opportunity (WOOP) measurement paradigm originally applied to taskwork assessment in the Georgia Tech Aegis Simulation Platform (GTASP) (Rothrock, 2001) has been embedded into TASP, and extended to assess teamwork. WOOPs are a mathematical way to evaluate performance based on a task analysis of the environment. They are continuously updated to determine what actions are appropriate at any given time. WOOPs identify the "purest," most complete set of opportunities or environmental affordances independent of actual human capabilities, and it is within this set of environmental affordances that WOOPs evaluate human performance (i.e., actions taken on time, early, late or missed).

For example, a ROE states that any hostile air contact must be given a level 1 warning if it flies within 50 and 41 nm from ownship. The SO has the primary responsibility for giving level warnings. When an enemy air contact comes within 50 nm, a level 1 warning WOOP opens. If the SO gives the level 1 warning before the air contact comes within 40 nm, that action is marked as an "on time" action. When the air contact comes within 40 nm of ownship, the level 1 warning WOOP closes, and if the SO gives a level 1 warning at that time, it is considered a "late" action. If the SO gives a level 1 warning to a hostile air contact that is not within 50 nm of ownship, then it will either be counted as an "early" action (if the air contact eventually comes that close) or a "false alarm" (if the air contact never breaches the 50 nm tripwire). In this way, actions are assessed against affordances in the environment.

As previously mentioned, the WOOP paradigm, originally applied only to taskwork behavior, has been extended in TASP to include teamwork behaviors. Specifically, teamwork WOOPs may be driven both by actions taken via the keyboard/mouse or by actions generated and taken through speech. For example, when the identification WOOP opens for the AAWC (indicating the appearance of a new unknown air contact to be identified), supporting team WOOPs open for the AIC and the SO to backup the AAWC in making the identification. If the AAWC identifies the track incorrectly, team WOOPs open for the SO and AIC to provide error correction (and for AAWC to correct his or her own error as well). Team WOOPs may be satisfied by taking the correct keyboard/mouse actions as well as by speech. The speech recognition and recording capabilities of the testbed will be discussed in the next section.

System Capabilities

The TASP simulation is part of a Team-in-the-Loop (TITL) simulation system that includes a scenario generation tool (ScriptMaker), an environmental simulation (of ADW), and a database repository that stores and manages both system and human performance data (Bhandarkar, Thiruvengada, Rothrock, Campbell, & Bolton, 2004). TASP was programmed in Java to run on a client-server architecture. Three clients representing each team member's workstation are connected to a single server. System requirements include Windows® 2000 and a Pentium 4 processor.

As mentioned previously, the TASP system captures and records speech interactions in order to support automated teamwork performance assessment with WOOPs. These speech interactions are recognized via the use of grammar templates embedded in Nuance®. This additional requirement for a speech recognition system forced the inclusion of three additional client computers to support the required processing speed for speech recognition capabilities at each workstation. It should be noted that Nuance® is a COTS product, and thus is the only component of the testbed that is not government-owned.

Because automated performance measurement and assessment via a speech recognition system is a relatively risky undertaking, we have included redundant keyboard/mouse actions for each team member's primary tasking. Additionally, to support teamwork interactions, a limited set of supporting team behaviors may be conducted via the mouse and keyboard as well (e.g., the SO and AIC can backup the AAWC by entering identifications).

USE CASE EXAMPLE

To help illustrate the TASP testbed and its capabilities, we will now briefly describe a hypothetical use case. Imagine that a researcher is interested in investigating the characteristics that a synthetic agent-based teammate would need to have in order to successfully support the training goal of improving teamwork skills. In particular, the researcher is interested in determining whether it is the act of modeling effective teamwork skills or the act of providing opportunities for the trainee to practice teamwork skills that leads to the most effective training experience for a human trainee. This research question will guide the development of the training scenarios and the design of the experimental and control conditions.

First, the researcher would use ScriptMaker to create a set of scenarios that require extensive use of teamwork skills. For example, incorporating a large number of events that are likely to overwhelm one of the team members creates a situation that requires the teamwork skill of backing up your teammate. Similarly, designing a scenario in which the context dramatically changes (perhaps from routine to threatening) requires the use of effective leadership and communication to help the team maintain a common situational awareness.

Next, the researcher would plan the nature of the experimental and control conditions. For example, if the synthetic teammate who was overworked in the scenario was "programmed" (via the flags in ScriptMaker described earlier) to miss many required actions and make some mistakes, this provides the opportunity for someone to provide support to that agent (i.e., error correction or backup). Using a second HBR-based teammate to perform many of the supporting behaviors creates a condition in which good teamwork skills are modeled for the trainee. Alternatively, using the flags in ScriptMaker to prevent a second synthetic teammate from performing the supporting behaviors establishes a condition in which there are many opportunities for the human trainee to support the original agent.

A control condition could be established by running teams with 3 human trainees, which is quite realistic, but gives up control over the specific nature of the learning experience. In other words, there are no guarantees that there will be any opportunities to provide supporting behaviors, or that there will be any instances of good supporting behavior being modeled in this team.

After creating the scenarios and establishing the three conditions (2 experimental & 1 control), the researcher would recruit participants, randomly assign them to one of the conditions, and run them through the experimental procedure. The procedure would undoubtedly start with some overview training of the TASP system and its interface, as well as the roles and responsibilities of each team member. Then, each team would run through a small number of the planned scenarios. It is likely that each scenario would be followed by a debrief or after-action-review, which would include a discussion of mission outcomes and the relevant taskwork and teamwork skills. One third of the participants would complete these scenarios with other human trainees. One third would experience one synthetic teammate consistently demonstrating good teamwork skills by backing up or correcting the other synthetic teammate. The last third of the participants would experience scenarios in which they repeatedly had the opportunity to backup or correct a synthetic teammate who made a lot of errors.

The procedure would end with some test of the effectiveness of those varied learning experiences. An obvious test would be to run each human trainee through a final scenario with two human teammates. Those teammates would be experimental confederates who deliberately made a number of errors and missed taking certain required actions. The dependent variable would be the extent to which the trainee effectively backed up and/or corrected the confederates.

The data collection component of the TASP testbed would provide information at many different levels to help the researcher figure out what happened in this experiment. At the highest level, the automated performance measurement system, based on Windows of Opportunity (as described earlier) will provide counts of the number of times the human trainee appropriately backed up and/or corrected his/her teammates during the final test, relative to the number of opportunities for supporting behavior (it should be noted that some of these behaviors would have been expressed through actions taken via the keyboard, while other supporting behaviors would have occurred through natural language interactions, and thus would have been captured with Nuance®). This allows a direct comparison of the outcome of the three training conditions, and thus a direct answer to the researcher's initial question: what behaviors should a synthetic teammate exhibit in order to effectively train teamwork skills?

A finer grained analysis, however, would most likely provide insight into exactly what happened and why during the training exercises and final test. For example, it would be important to determine exactly how many instances of good teamwork skills were modeled during the training of the control group, as well as how many opportunities each team member had to support a teammate. In all probability, these numbers varied highly from team to team within this condition (as compared to the two experimental conditions with synthetic teammates, in which these events were carefully controlled). It would also be important to determine exactly what each trainee was doing during those times when he/she missed opportunities to provide backup or correct a teammate. For example, those instances in which the trainee was busy with primary responsibilities should be interpreted differently from those instances in which the trainee is not heavily tasked. This can be determined by referring to the time-stamped log of all sequences of meaningful actions, which also indicates which team member took each action.

The results of this research should be guidance for the development of synthetic agents which could be used during simulation-based training exercises when human teammates are not easily available and/or (depending on the results of the study, of course) when there are specific training goals that cannot be met as efficiently when training with a team of all humans.

LIMITATIONS

While the TASP testbed was designed to maximize flexibility to support a wide variety of research efforts on team performance and training, there are some limitations that should be noted. First, the nature of the simulated task in the TASP testbed relies heavily on cognitive capabilities involving information processing and decision making. Therefore, the testbed would not be well suited for all applications. For example, it would not be the appropriate testbed to study the development of perceptual-motor skills. Additionally, the automated team performance measurement capability is supported by a commercially-available software package, Nuance®, that must be purchased in order to study team communications. While the testbed and supporting task performance assessment system will run without the Nuance® software support, the automated team performance measurement system will not. Finally, the TASP system will not run faster than real-time, so this poses a limitation to those researchers using HBRs in lieu of human participants to conduct experiments.

SUMMARY

Teams play a crucial role in military operations, thus the U.S. Department of Defense needs to understand how they perform a mission and how to best train individuals to operate as a distributed, yet cohesive and effective unit. Research and Development (R&D) in the area of team performance and training will help shed light on these issues. To provide a flexible testbed to support these R&D efforts, an interdisciplinary team of researchers defined team testbed requirements and developed a solution, the Team Aegis Simulation Platform (TASP). Some of the many strengths of the TASP testbed include: free access, the inclusion of a flexible scenario design tool, the capability to communicate in real time with intelligent agents, and the inclusion of a comprehensive data collection system. We hope that researchers will find TASP to be a valuable tool for their research toolbox and that it will provide a significant contribution to the body of knowledge on teams.

ACKNOWLEDGEMENTS

This project was funded by Dr. Harold Hawkins at the Office of Naval Research (ONR) under award number N0001403WX20275. The synthetic teammates are being developed by CHI Systems, Inc. with a portion of the funding received from Dr. Hawkins. The authors would also like to thank Cheryl Johnson, Gabriel Lopez, and Allison McCreary for their support on the data collection effort performed under this program.

REFERENCES

Bhandarkar, D., Thiruvengada, H., Rothrock, L., Campbell, G., & Bolton, A. (2004). TASP: A toolkit to analyze team performance in a complex task environment. Presented at the *IIE 13th Industrial Engineering Research Conference*, May 15-19, 2004, Houston, TX.

Bolton, A. E., Dorsey, D. W., & Campbell, G. E. (2004). Application of time-sensitive computational models to training within a scenario-based team training system. *Proceedings of the 13th Conference on Behavior Representation in Modeling and Simulation* [CD-ROM].

Buff, W. L., Bolton, A. E., & Campbell, G. E. (2003). Providing an integrated team training capability using synthetic teammates. *Proceedings of the American Society of Naval Engineers Human Systems Integration Symposium* [CD-ROM].

Fleishman, E. A. & Zaccaro, S. J. (1992). Toward a taxonomy of team performance function. In R.W. Sweezy & E. Salas (Eds.), *Teams: Their training and performance*. Norwood, NJ: Ablex.

Rothrock, L. (2001). Using time windows to evaluate operator performance. *International Journal of Cognitive Ergonomics, 5*(2), 95-119.