DMT-RNet: An Internet-based Infrastructure for Distributed Multidisciplinary Investigations of C2 Performance

Eileen B. Entin
Daniel Serfaty
Aptima, Inc.
12 Gill Street, Suite 1400
Woburn, MA 01801
ebe, Serfaty@aptima.com

Linda R. Elliott
Veridian, Inc.
San Antonio, TX 78235
Linda.Elliott@brooks.af.mil

Samuel G. Schiflett
Air Force Research Laboratory
Brooks AFB, TX 78235
Sam.Schiflett@brooks.af.mil

Abstract[1]

We describe the first phase of a collaborative research program that leveraged emerging capabilities of the internet to connect a distributed PC-based synthetic task environment system for the purpose of research in distributed team training and decisionmaking. Using an internet-based version of the dynamic distributed decision making (DDD) paradigm, 16 participants from four multidisciplinary research groups (Brooks AFB, University of South Florida, University of Central Florida, and Aptima, Inc.) were connected and performed in a team-on-teams C2 scenario. Observers at the sites viewed scenario play in real time, and communicated with one another via voice and email. Immediately following the end of scenario play, team performance assessments were collected, integrated, and distributed back to all participants for use in an after-action review. This infrastructure will provide the means to investigate issues related to effectiveness in military distributed mission training (DMT) systems including: integration of training goals; use of constructed forces; training content, sequencing, and delivery; scenario fidelity; distributed online performance assessment; and interventions to improve learning and performance in mission planning, execution, and debriefing. The internet-enabled DDD provides a collaborative training space for investigation of these DMT issues while permitting experimental control, unrestricted data analysis, and cooperative research across distributed sites.

Introduction

There is no doubt that distributed mission training (DMT) will be critical to 21st century military training and performance (Kreisher, 2000). DMT links multiple high-fidelity simulators to allow operational personnel in distributed locations to train together in complex battlespace scenarios. However, high-fidelity simulators are limited in availability, and the scenarios are so realistic they run in classified mode, which impedes investigation and reporting from a researcher's perspective. Indeed it impedes the participation of any researcher who does not have a security clearance. A need exists for an infrastructure that enables investigations of DMT that capture the requisite complexity while enabling systematic, controlled, and reportable studies. Such an infrastructure

[1] Opinions, interpretations, conclusions, and recommendations are the authors and are not necessarily endorsed by the United States Air Force. This research was supported by work units 2313HA07 and 1123A002 through funding by AFOSR and AFRL.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **2001**	2. REPORT TYPE	3. DATES COVERED **00-00-2001 to 00-00-2001**
4. TITLE AND SUBTITLE **DMT-RNet: An Internet-based Infrastructure for Distributed Multidisciplinary Investigations of C2 Performance**		5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Air Force Research Laboratory,Brooks AFB,TX,78235**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
The original document contains color images.

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**		**12**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

would provide cost-effective and highly germane training experiences at moderate costs, with little or no danger to physical health or environmental consequence.

In the first phase of a collaborative research program we successfully leveraged emerging capabilities of the internet (i.e., internet II) to connect distributed team players and distributed observers via a PC-based distributed synthetic task environment (STE) to investigate issues related to effectiveness in military distributed mission training (DMT) systems. The infrastructure we have developed provides a DMT research net (DMT-RNet) that will allow us to investigate issues related to distributed training performance assessment, and feedback, and to investigate the merits of alternative interventions to improve learning and performance in mission planning, execution, adaptive performance, and debriefing. In this paper we describe the distributed STE we have developed to conduct research on the "team of teams" scale. We then discuss issues related to scenario fidelity raised by use of a STE. Following that, we describe the scenario we have developed, and the procedures we developed and used for the collection and dissemination of performance data. We close with a discussion of theoretical and applied research issues enabled by this infrastructure.

The Distributed Dynamic Decisionmaking Network (DDDNet)

The Distributed Dynamic Decisionmaking (DDD) paradigm (Kleinman and Serfaty, 1989; Kleinman, Young, and Higgins, 1996) is a unique distributed multi-person simulation and software tool for understanding command and control (C2) issues in a dynamic team environment. It provides a distributed real-time simulation environment implementing a complex synthetic team task that includes many of the behaviors at the core of almost any team task: assessing the situation, planning response actions, gathering information, sharing information, allocating resources to accomplish tasks, coordinating actions, and sharing or transferring resources. Research applications using the DDD have demonstrated the paradigm's flexibility in reflecting different domains and scenarios to study realistic and complex military team decision-making, and to explore issues related to team training. The DDD provides a substantial degree of control over scenario design, team organization, and data collection, while engaging the team players in a low-to-moderate degree of realism. The abstract, game-like nature of DDD scenarios makes them well suited for use by teams who have not yet acquired the level of knowledge and proficiency required to operate and benefit from a high fidelity simulator.

In the first phase of a collaborative program of research on DMT, we developed an internet-based version of the DDD, DDD Network (DDDNet). The DDDNet is an integrated internet-enabled, collaborative gaming space that connects players to each other and to others, such as observers, confederates, trainers, or researchers. It allows players in distributed locations to connect and perform a distributed mission in real time. Observers and/or team trainers at any location in the network can observe the scenario play in real time. They can view the screen display and electronic communications of any player, and communicate to one another via email or voice. In addition, the DDDNet can connect players to one another for interactive mission planning, debriefings and after-action reviews (AARs).

The DDDNet provides the basis for the DMT Research Network (DMT-RNet). An integral part of the DMT-RNet is a suite of team performance and process measures that provide feedback on and evaluation metrics for team performance. These measures are available from three sources: the DDD simulation, distributed observers, and distributed players. They can be targeted at three levels: individual performance; team performance, and system-level, team-of-teams performance. The DDD-based performance outcome measures focus on effectiveness, timeliness, efficiency, speed, and latency. The DDD also provides real time individual and team performance measures that are calculated dynamically and visible throughout the performance of a scenario.

The DDDNet provides an effective training environment because it collects all events and actions that occur over the course of a scenario run. Post-processing modules can parse the output files generated in a DDD run, building a database of the results for rapid manipulation, analysis, and feedback to players and/or use in performance evaluation. The output files provide the capability to play the game back after the team members have completed a scenario run, which can be particularly useful in an instructional environment.

In successful tests of the DDDNet participants from four C2 research groups participating in the research program—Brooks Air Force Base, University of South Florida (USF), University of Central Florida (UCF), and Aptima, Inc. —were connected and performed in a team-on-teams C2 scenario. The DDDNet achieved and maintained a synchronized connection involving 16 participants. During the execution of the scenario, players were able to see dynamically changing measures of effectiveness. Observers were able to focus their observations at any level of detail they chose, ranging from the entire mission team to individual players on one team. At the end of play, both players and observers responded to an assessment questionnaire via the internet. As soon as that data was collected, it was tabulated and summarized, and made available to both players and trainers. At the same time, performance and process measures derived from the DDD simulation were calculated and distributed. A debriefing session was conducted in two phases. In the first phase observers provided performance feedback to the individual nodes and in the second phase observers and players at all the sites were interconnected for feedback and discussion.

Scenario Development: Issues Regarding Fidelity

A PC-based STE system such as the DDDNet will not be as realistic as a high-fidelity DMT system. These systems have a high degree of physical fidelity, in that the system replicates actual operational equipment, procedures, and task demands. DMT systems run in a highly secure classified mode within their own network and allow operators of different simulations systems to operate together in realistic battle scenarios while located at distributed sites. DMT systems allow highly realistic mission rehearsal practice in a multi-system multi-operator environment.

PC-based STE systems can also create complex environments for multi-operator training and performance research that network in unclassified mode on the internet. This allows the STEs to be deployed to almost any setting. It also allows university-based research to occur, based on systems that reflect essential components of operator expertise. STE systems should enable more controlled investigations while capturing performance phenomena in complex, multi-operator, expert-based operational performance (Driskell & Salas,1992).

Effectiveness of an STE depends on appropriate scaling of the operational performance domain to assure that core functions and task characteristics are maintained at proper levels of complexity. STE platforms and scenarios are "synthetic" in that they are developed through functional and cognitive task analysis of an operational performance domain, and are thus analogues with regard to particular aspects of the performance domain. As operational analogues, they can be distinguished from other PC-based gaming environments.

Physical Versus Psychological Fidelity

The traditional approach to simulation systems is to maximize the realism, or physical fidelity, of equipment and procedures. This approach was assumed to maximize transfer of training to operational systems, based on the "identical elements theory" of transfer proposed at the turn of the 20^{th} century by Thorndike and Woodworth (1909). Assuming that scenario content is also fully realistic, these systems will provide task-specific findings with the most accurate extrapolation for performance models.

Fidelity is a difficult and complex multi-dimensional issue. Realistic equipment will certainly facilitate training and transfer of knowledge and procedural skill in equipment use. However, there are higher-level knowledge and skills that are independent from equipment procedures per se, such as knowledge of goals, roles, responsibilities, taskwork, tactics, strategies, teamwork, and/or dynamic problem solving. These aspects of expertise in a domain can be elicited, assessed, and/or trained without full equipment realism. Indeed, the fidelity and validity of a fully realistic simulation system is also limited by the degree of realism and operational relevance of its scenarios.

Aspects of Psychological Fidelity

Physical fidelity refers to the extent to which the system looks and performs like the "real thing." Psychological fidelity is a broad term that refers to the degree to which the system captures functional and cognitive aspects of the performance domain. As such, it is related to issues of content validity and/or construct validity, depending on the purpose of the STE. Systems based on content/construct validity should also generalize to operational systems, despite limitations in physical fidelity (Berkowitz & Donnerstein, 1982; Bowers, et al., 1992; Cronbach, 1989; Mook, 1983).

Table 1 lists several aspects of fidelity and their impact on three kinds of validity. Each aspect listed in Table 1 contributes to overall content validity, construct validity, and/or external validity (e.g., generalizability, training transfer) of the STE system, and can also serve as a starting point for a more detailed analysis of the operational domain. For example, physical fidelity may affect content and external validity but is relatively independent of construct validity. Functional fidelity and cognitive fidelity are aspects contributing to content validity, and should therefore also relate to external validity. A platform may be high or low in each aspect of fidelity. For example, a platform may have high construct fidelity/validity but low functional fidelity, and vice versa. The research and/or training goals will help identify relevant task characteristics, levels of performance, and the level of expertise required for scenario performance. *It should be noted that systems high in physical fidelity may still be low in content, construct, and external validity, depending on scenario content and training/research goals.*

Table 1. Dimensions of Fidelity: Impact on Validity

DIMENSION	DEFINITION	IMPACT ON VALIDITY
PHYSICAL Fidelity	Degree to which equipment and procedures are similar to operational features.	Content, External
FUNCTIONAL Fidelity	Degree to which operational goals and functions are representative, (e.g., mission / tactical goals, subgoals, operator roles, interdependencies, and specific decision / task events).	Content, External
COGNITIVE Fidelity	Degree of match in cognitive complexity of individual and team information processing demands. (e.g., information monitoring, perception, interpretation, and decision making.)	Content, Construct, External
CONSTRUCT Fidelity	Degree to which performance constructs of interest (e.g. planning, teamwork, situation awareness, decisionmaking, problem solving, etc.) are inherent in both operational and STE settings.	Construct, External

Scenario Characteristics

Scenario Description

In developing the scenario for the first phase of this collaborative research, a great deal of effort was focused on capturing important aspects of functional and cognitive fidelity. The scenario developed involves three nodes or teams, representing the functions of an Air Force AWACS team (AWACS), a Navy Carrier Battle Group (CVNBG) and an Air Force Combined Air Operations Center-Forward (CAOC-F) team. The Joint mission was designed to motivate intra- and inter-team collaboration and coordination of activities.

The primary constructs in this scenario are aspects of C2 teamwork. Teamwork in any setting, in its essence, is the effectiveness by which team members manage their interdependencies in order to achieve a shared goal. They may do this by specifying coordination mechanisms in advance (e.g. tactics, strategies, contingencies, etc.) or by responding effectively to unexpected events through general mechanisms of problem recognition and problem solving/decision making. Here, we wanted to ensure capture of both aspects of team performance as they affect (a) mission planning, (b) sequencing of events, according to plan, (c) sharing of information and/or assets, according to plan, and (d) dynamic problem solving. These aspects were identified through various cognitive task analysis methods that were focused on individual performance (Klinger, Andriole, Militello, Adelman, Klein, & Gomes, 1993) and team performance (MacMillan, Serfaty, Young, Klinger, Thordsen, Cohen, & Freeman, 1998; Elliott, Schiflett, Hollenbeck, & Dalrymple, in review; Coovert, Gordon, Foster, Riddle, Miles, Hoffman, & King, 1999) for the purpose of scaling STEs (Schiflett & Elliott, 2000).

The overall context of the scenario is a coalition defense of a small friendly country against invasion from a large enemy country. The scenario unfolds over three phases: (a) seize air superiority; (b) defend against response; and (c) halt invasion. Each phase has tactical goals, such as "destroy the enemy AWACS asset." Each tactical goal has subgoals, often accomplished by coordinated action. For example, the destruction of enemy AWACS is executed through a coordinated sequence of events by each team. The CAOC-F team draws out enemy fighter assets, while the AWACS team sends more powerful assets to destroy the enemy AWACS. Each tactical goal and subgoal is made explicit in the initial Air Tasking Order (ATO). In addition, further coordination and contingency planning is explicated during mission planning. These events and plans form a solid basis for the assessment of mission plan execution. The scenario enables event-based assessment of performance in mission planning, team coordination (execution of the plan), and adaptive problem solving, using an event-based measurement process similar to the approach described by Fowlkes, et al. (1994).

Implementation on the DDDNet

Figure 1 shows the structure and communication paths within and across teams that were used in the Phase I work. Team members at the individual nodes were co-located, while the team of teams was distributed across the continental U. S. At the start of planning, each team was given an ATO that described the current tactical situation and laid out the tactical goals and plans at an overall level. During mission planning the teams worked out details with regard to assets and targets. Team members worked out details with regard to which particular resources they would use to verify and prosecute targets. In planning target sequences, priority codes are considered, along with target window of opportunity (there may be a particular time when it should be attacked) and asset window of opportunity (location, fuel, etc). The CAOC-F node may have to assign resources to verify targets that are not listed as "certain"—the ATO may list locations where a target, such as a mobile missile launcher, is thought to be located. Then surveillance assets would have to go first to verify the target before sending assets to destroy it.

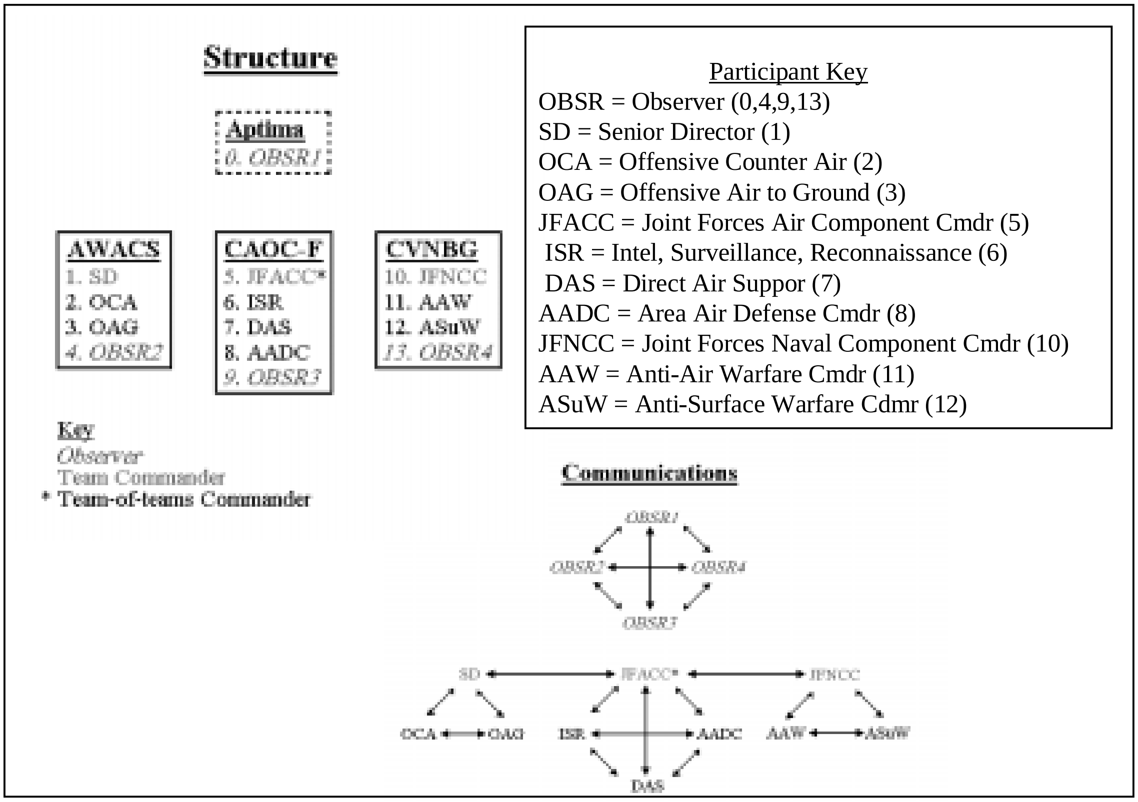

Figure 1. Team Organization and Communication Network

Explication of the mission plan, tactical goals, subgoals, and contingencies accomplished during mission planning served as a basis for investigation of mission performance. It is relatively straightforward to assess achievement of tactical goals and identify errors and/or deviations from the plan. At the same time, we wished to allow investigation of dynamic and adaptive performance as a response to unexpected events. In the scenario many of the planned targets either did not exist or were decoys. If the surveillance function identified these targets correctly, plans had to change in rapid response in order to avoid wasted effort and assets.

Figure 2 shows a screen shot of the DDD-Net display captured during the final Phase I DMT-RNet demonstration. Training in use of the DDDNet was provided prior to the Phase I demonstration. Figure 2 highlights some of the features of the DDDNet simulation testbed and user interface. Scenario events are portrayed in the window on the left hand side of the display. Display controls are located in the middle of the right-hand side. Identification and prosecution of potential targets is accomplished with mouse clicks. Real-time measures of individual and team performance are displayed in the upper right-hand side.

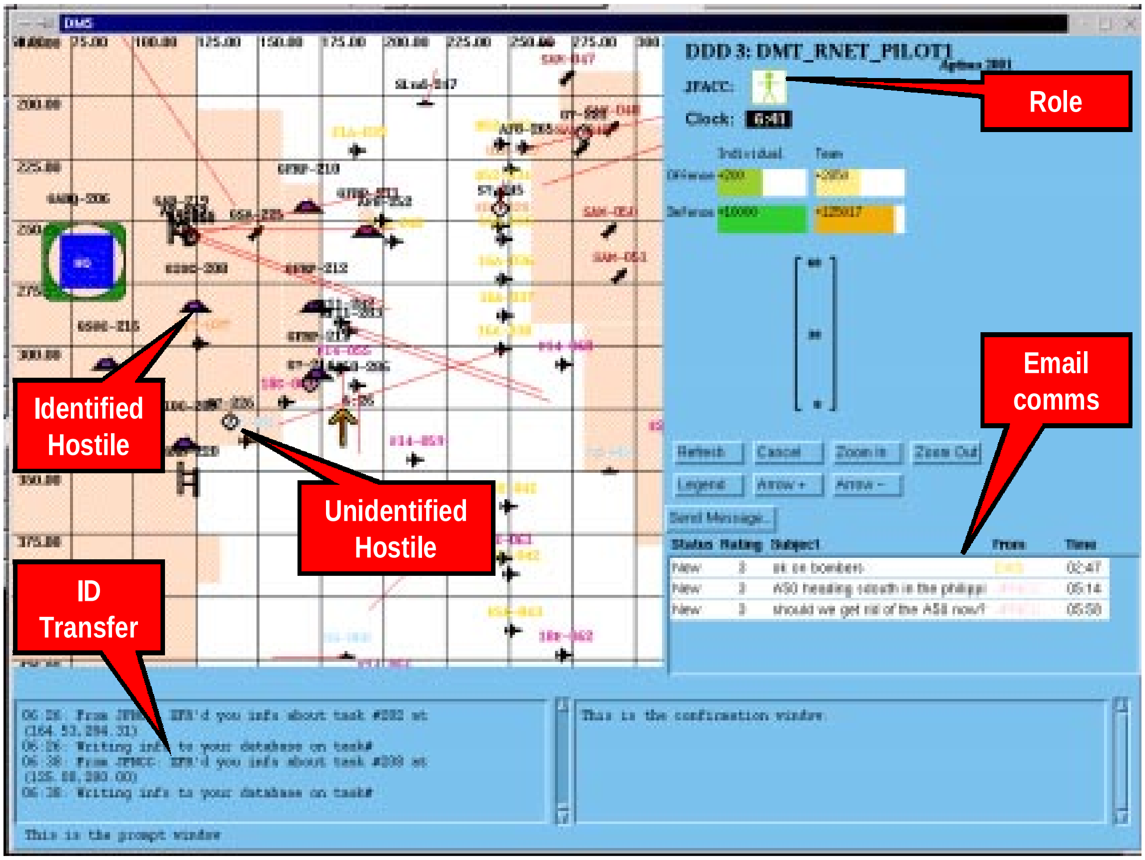

Figure 2. DMT-RNet Screen from distributed scenario execution involving four nodes

Collecting and Distributing Team Performance and Process Measures

Measures of team performance and team processes were collected from three different sources: the DDD simulation, players, and observers. A goal of the research team was to be able to provide feedback to the participants in the scenario immediately following the end of the simulation, and this was successfully achieved in the Phase I demonstration.

Simulator Data: DDD-based measures were derived from the performance and process data collected by the simulator during scenario play. At the end of the scenario, a post-processor at the Aptima DDD node aggregated the measures at three levels: individual, team, and team-of-teams. Table 2 shows some examples of measures that were obtained, the levels at which they were measured, and the type of feedback to players that they provided. These measures were sent via the internet to a node at the USF, where they were posted on an internet site accessible to all participants and researchers involved in the Phase I demonstration.

Table 2. DDD-Based Team Performance and Process Measures

Level of Measurement	Definition	Type of Feedback
Task, team-of-teams performance	Gains/Losses (Displayed Dynamically) Offense: Value of enemy assets destroyed (increase from zero at start) Defense: Diminished value of own assets (decrement from total value at outset)	Role specific Individual performance score; Overall mission offense/defense performance
Task, team, team-of-teams functioning	(Note: friendly = role owned) # friendly assets not destroyed # friendly assets destroyed # friendly assets lost to hostile fire # friendly assets lost to fuel out # hostile assets destroyed by friendly action kill ratio air refuelings completed	Role specific Individual performance score; Offense/defense actions – summed across individuals in team, Overall mission performance
Team Processes	Number of transfers of resources # transfers in # transfers out Number of emails # emails sent # emails received	Individual, node-level, overall (team-of-teams) teamwork

Player Evaluation: At the end of the scenario run, players at each node responded to a questionnaire developed by researchers at the USF. This questionnaire tapped the players' assessments of their and their teammates' performance during the simulation, as well as their assessment of the effectiveness of team processes. The players' responses were aggregated and posted at a USF internet site accessible to all participants.

Observer Evaluation: At the end of the scenario run, the observers at each site also responded to a questionnaire that focused on the performance and processes of the team they observed. The observers' responses were aggregated at USF and posted to the same internet site as the players' responses.

At the conclusion of the scenario execution, team performance was discussed and feedback provided at the individual team sites. Following that a distributed AAR involving players and researchers at all nodes was conducted. The ability to collect, aggregate, and make feedback available to all participants immediately following scenario play demonstrated a fundamental requirement for distributed online training and evaluation. As well as providing post-scenario feedback, the data and measures collected during and following the scenario run remain available to researchers for additional analysis. We also note that the replay feature of the DDDNet simulation provides a useful feature for training feedback and after action reviews.

Summary and Future Directions

Phase I Accomplishments

In the Phase I work we achieved several critical goals, including:

- Demonstrating the ability to interconnect four DDD team simulators over the internet;
- Running a distributed team-of-teams scenario in synchronous mode;
- Collecting and integrating evaluation data over the internet immediately following the end of scenario play, and;
- Rapid internet-enabled presentation of evaluation measures to all participants in distributed locations.

Much of the Phase I work focused on technical connectivity issues and issues related to the coordination and communication of distributed players and observers. Planning for and developing the Phase I demonstration also illuminated some of the challenges that occur when a distributed group of researchers are involved, such as efficient methods for distributing scenario and evaluation materials.

Future Research Goals

A major goal of the DMT-RNet program is to explore efficient and effective ways to conduct distributed training. Issues that must be resolved include: coordinating co-present and distant trainers; establishing appropriate observation stations at distributed locations; effective methods for communication among distributed observers, trainers and trainees; optimal methods for presenting and discussing training evaluations in a distributed environment, and visual display of performance feedback. Other fundamental research issues that will be investigated in future work include scenario complexity and realism, scenario-driven measures of performance, and carryover of DMT using STEs to distributed high-fidelity simulations and to on-the-job performance. A related avenue of research involves the use of constructed forces for DMT. The infrastructure for the DMT-RNet we have developed in the first phase of our research provides a flexible and cost-effective basis for exploring these issues in a systematic research program that can be conducted in a nonclassified mode.

Acknowledgements

We thank Mathieu Dalrymple at Veridian, Inc. for his work in designing the scenario. We thank Dawn Riddle at the University of South Florida for her work in posting the performance measures on the internet.

References

Berkowitz, L., & Donnerstein, E. (1982). External validity is more than skin deep. Some answers to criticisms of laboratory experiments. *American Psychologist, 3*7 (3), 245-257.

Bowers, C., Salas, E., Prince, C., & Brannick, M. (1992). Games teams play: A method for investigating team coordination and performance. *Behavior Research Methods, Instruments, and Computers, 24* (4), 503-506.

Coovert, M. D., Gordon, T., Foster, L. L., Riddle, D., Miles, D. E., Hoffman, K. A., & King, T. S. (1999). *AWACS weapons directors cognitive/behavioral categories: The final report of USF's AWACS research group.* University of South Florida.

Cronbach, L. (1989). Construct validity after 30 years. In R. L. Linn (Ed.) *Intelligence: Measurement, theory and public policy*. Urbana and Chicago: University of Illinois Press.

Driskell, J., & Salas, E. (1992). Can you study real teams in contrived settings? The value of small group research to understanding teams. In R. Swezey & E. Salas (Eds.) *Teams: Their training and performance* (pp. 101-126). Norwood NJ: Ablex.

Elliott, L. R., Hollenbeck, J. R., Schiflett, S. G., & Dalrymple, M. (In Review). Investigation of Situation Awareness and Performance in Realistic Command and Control Scenarios. In Review for M. McNeese, E. Salas, & M. Endsley (Eds.) *Group Situational Awareness: New Views of Complex System Dynamics* . LEA.

Fowlkes, J., Lane, N., Salas, E., Franz, T., & Oser, R. (1994). Improving the measurement of team performance: the TARGETS methodology. *Military Psychology, 6*, 47-61.

Kleinman, D. L. & Serfaty, D. (1989). Team performance assessment in distributed decision-making. *Proceedings of the Symposium on Interactive Networked Simulation for Training* (pp. 22-27), Orlando, FL.

Kleinman, D. L., Young, P. W., and Higgins, G. W. (1996). The DDD-III: A tool for empirical research in adaptive organizations. *Proceedings of the 1996 Command and Control Research and Technology Symposium* (pp. 827-836), NPS, Monterey, CA.

Klinger, D. W., Andriole, S. J., Militello, L. G., Adelman, L., Klein, G., & Gomes, M. E. (1993). *Designing for performance: A cognitive systems engineering approach to modifying an AWACS human-computer interface* (AL/CF-TR-1993-0093). Wright-Patterson AFB, OH: Armstrong Laboratory.

Kreisher, O. (2000). The Wild VR yonder. *Military Training Technology, 5*, 3.

MacMillan, J, Serfaty, D., Young, P., Klinger, D., Thordsen, M., Cohen, M., & Freeman, J. (1998). *A system to enhance team decision making performance: Phase 1 Final Report.* Brooks AFB: Air Force Research Laboratory, Warfighter Training Research Division.

Mook, D. G. (1983). In defense of external invalidity. *American Psychologist*, 379-387.

Schiflett, S. & Elliott, L. (2000). Development of Synthetic Team Training Environments: Application to USAF Command and Control Aircrews. In H. O'Neil & D. Andrews (Eds.) *Aircrew Training: Methods, Technologies, and Assessment*.

Thorndike, E. L., & Woodworth, R. (1909). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review*, *8*, 247-261.