The Utilization of Synthetic Task Environments in C2 Domains

Scott M. Galster
W. Todd Nelson
Robert S. Bolia

Air Force Research Lab
WPAFB, OH

ABSTRACT

A description of several Synthetic Task Environments that are used in command and control research programs is offered. The STEs as well as new augmenting capabilities are currently being used at the Air Force Research Lab to examine technologies, procedures, and concepts that will enhance air battle manager capabilities and situation awareness while decreasing the workload associated with these environments. A simple taxonomy is given for the selection of synthetic task environments in this domain.

INTRODUCTION

Synthetic task environments (STEs) are often thought of as places where abstractions of real world tasks can be researched and are often employed in research programs that address command and control (C2) issues (Cooke & Shope, 2004). These authors recently outlined the costs and benefits of utilizing STEs and the authors provided a step-by-step process for designing appropriate STEs for use in research programs. The purpose of this paper is to explore a continuum of STEs that are appropriate for use depending on the type of research that is conducted and the research questions that are being asked. In basic research, which might be conducted to flush out theoretical propositions, STEs are afforded the luxury of higher levels of task abstraction. Conversely, STEs for advanced research projects that fall just short of operational fielding should minimize abstraction and rely more profoundly on ecological validity or realism. We propose that the mediating factor in the choice of which STE to use is a function of the ability to transition the results to real or higher fidelity systems and the proposed return on the investment made by choosing a particular STE.

The criteria for STE selection in C2 environments depends heavily on the type of research that is being conducted. Selection of STEs in C2 programs is unique in that characteristics of C2 environments constrain the number of candidate systems. C2 environments are inherently dynamic, have a need for communication and/or interoperability, and the STEs used for C2 research programs must possess these same characteristics. A simple model of the trade space between research objectives, user expertise, and fidelity will be proposed that will facilitate the selection process of STEs within the C2 domain. Some of these factors have been discussed at length (see Schiflett, Elliott, Salas, & Coovert, 2004). Our framework will focus on three dimensions; tractability, realism, and experimental control.

Tractability (Ehret, Gray, & Kirschenbaum, 2000) refers to the ability of the STE to answer the research question that is being asked. It is notionally aligned to the issue of complexity described by Brehmer, and Dörner (1993) in that it is concerned with the granularity of the data collected and the skill level or training required to learn how to use the STE. We propose that tractability also refers to the level at which the STE addresses a theoretical continuum. This continuum is anchored on one side as theoretical, and on the other side as applied. It is noteworthy to point out that most STEs occupy a range on this continuum rather than a single point.

Realism refers to the maintenance of the functional relationship between the STE and the real-world system (if applicable) to which the STE research would be applied. Realism can vary as tasks are manipulated in their level of abstraction. High levels of abstraction are more appropriate for more fundamental research questions while low levels of abstraction lead to tasks that are designed to mimic those tasks operators perform in real systems.

Experimental control refers to the longstanding issue of the allowance of variability in an experiment that has the potential to directly affect the behaviors/responses of participants. The decision that has to be made by the experimenter is whether to accept the risk associated with lower levels of experimental control. Higher risk levels are usually indicative of higher potential payoffs in terms of data interpretation and applicability. However, if the risk is not judiciously accounted for, the potential exists that the data, and therefore the applicability of the data, will not be interpretable with respect to the research question that is being asked.

This paper describes several STEs in use at the Air Force Research Laboratory (AFRL), Human Effectiveness Directorate, Collaborative Interfaces Branch (henceforth referred to as AFRL). The STEs are used to examine C2 issues as they relate to current or potential Air Force initiatives. All of the STEs share the goal of eliciting feedback in the form of human performance differences and changes in situation awareness and workload as experimental variables are manipulated. The STEs range in complexity from a very basic gaming environment to a complex multi-role air battle management platform simulator. The STEs are used for very different purposes which will be expounded in the text as each STE is presented.

One tenet that must be followed for all research activities conducted at AFRL is the notion that research programs are best designed and executed with the purpose of providing more capability to the end-user. This can be accomplished by transitioning technology, processes, and procedures that have been evaluated and have been determined to provide an increased capability with limited or acceptable repercussions. All of the research conducted at AFRL, from the most basic to the most complex, is required to have a transition path identified which will result in increased capability to the warfighter.

BASIC RESEARCH

Most of the basic research conducted at AFRL under the Battle Management Command and Control (BMC2) domain takes place in the Decision-Making and Automation Research Testbed (DART) Laboratory. The DART Lab is host to the RoboFlag STE, the Multi-Modal Immersive Intelligent Interface for Remote Operation (MIIIRO) STE, and the Dynamic Distributed Decision making (DDD) STE. Some common characteristics of the STEs in this lab include: theory-driven research applicable to nearly all C2 environments; higher task abstraction levels; rapid evaluations of theory driven constructs; and moderate to high levels of experimental control. These STEs, in general, evaluate C2 concepts taken directly from theory driven constructs. These constructs are applied to a C2 environment, evaluated for both applicability and utility, and then identified as either worthy or not of transition to STEs with higher levels of scrutiny. This process is iterative in nature as many aspects of one theoretical construct may be evaluated recursively. Because of the low cost and flexibility associated with these STEs they are used often in research environments.

RoboFlag

The RoboFlag STE was developed at Cornell University to develop and evaluate algorithms for hierarchal control of multiple autonomous vehicles (see http://roboflag.mae.cornell.edu/ for further information). It is based on the game of "Capture the Flag" and offers a wide degree of flexibility and the ability to explore numerous control issues (D'Andrea & Babish, 2003; D'Andrea & Murray, 2003). The RoboFlag platform has been utilized recently by AFRL, Cornell University, The Catholic University of America, George Mason University, and Smart Information Flow Technologies, Inc. (SIFT) to evaluate human interaction with multiple autonomous vehicles via a delegation control architecture.

This approach is driven primarily by the fact that the supervision of multiple unmanned vehicles (UVs) is currently labor intensive and control architectures are mapped directly onto single tasks. For example, several humans are typically required to operate and supervise a single UV. The longstanding goal for developers of these systems has been to reduce the number of operators while concurrently increasing the number of UVs controlled. Increasing the vehicle-to-operator ratio by increasing UV autonomy is only one of the methodologies that should be considered. Another method to consider is the evaluation of operator interface design types that also increase the probability of mission success (Army Science Board, 2003).

One type of interface design applicable to the control of multiple UVs is a delegation-type interface. These interfaces can permit adaptable automation to be implemented in a flexible and variable fashion. In general, delegation-type architectures provide highly flexible methods of implementing human-declared goals (Sheridan, 1987). An example of such an delegation architecture is the Playbook™, which has been described elsewhere (Miller & Parasuraman, 2003; Miller, Pelican, & Goldman, 2000)-so named because it is based on the metaphor of a sports team's book of approved plays and the selection of these plays by the team leader (e.g., the quarterback in American football) and executed by the team members (the other players).

Three studies have been conducted to investigate the system performance effects of delegation-type interfaces on human supervision of multiple UVs. Participants supervised up to eight UVs using automated behaviors called "plays", manual point-to-point "waypoint" control, or a combination of these to capture the flag of an opposing team with an equal number of UVs. A typical RoboFlag user interface is shown in Figure 1.

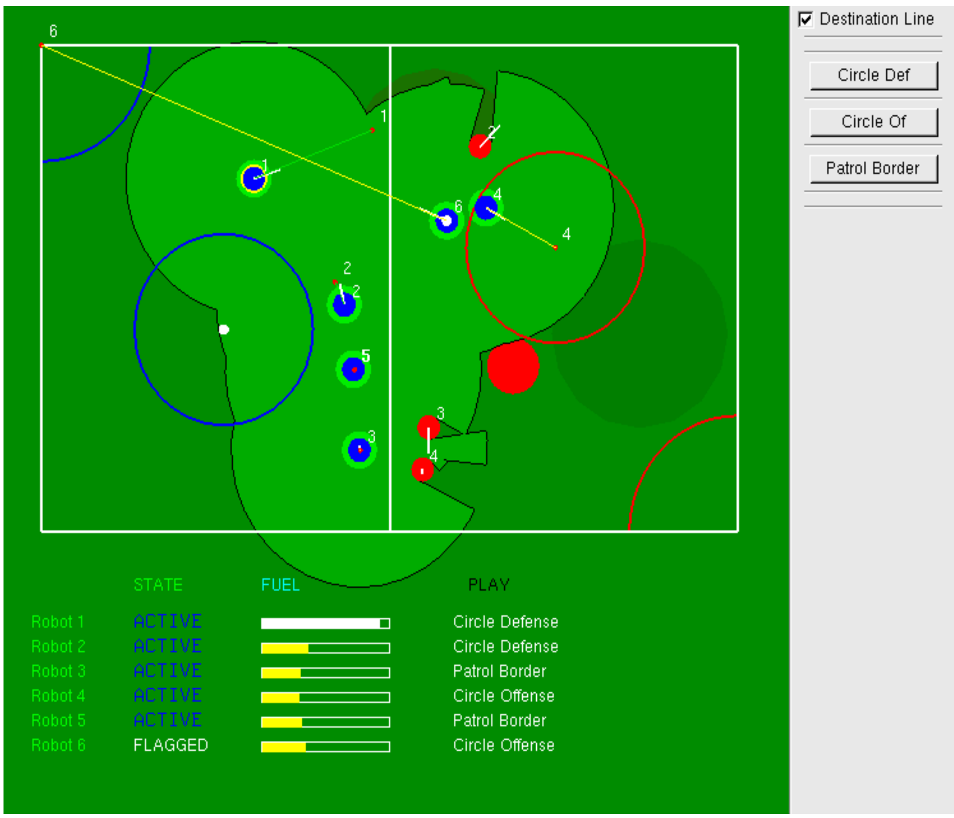

Figure 1. A typical RoboFlag interface showing a blue UV heading back to the home area with the red team flag.

The first experiment (Parasuraman, Galster & Miller, 2003) demonstrated that the delegation-type interface was effective at increasing mission effectiveness while concomitantly reducing mission completion times when the opposing team adopted a purely offensive or defensive stance. The second experiment (Squire, Galster & Parasuraman, 2004) demonstrated that mission effectiveness was increased when operators could choose between plays or waypoint (flexible) control rather than having to rely on either play or waypoint control exclusively. The downside to this finding was that operators reported a small increase in subjective mental workload while utilizing the flexible control. The third experiment (Squire, Furukawa, Galster, Miller & Parasuraman, 2004) compared delegation-type interfaces to restricted interfaces. The relevant factors included the level of abstraction and the level of aggregation (single or multiple UVs) that control could be utilized. The results suggested that performance was enhanced when operators could control UVs with flexible interfaces. However, the performance benefit diminished as operators had to supervise an increasing number of UVs.

The RoboFlag platform has also been used to identify and develop measures of decision quality as it relates to mission effectiveness in an examination of the strategies employed by operators when they were provided unreliable information with regard to the opposing team strategy (Galster & Bolia, 2004a; Galster & Bolia, 2004b). The results of the experiment indicated that sub-optimal strategy usage was prevalent in games that were both won and lost. However, inaction by the participant was indicated in a higher percentage of the games that resulted in a loss. Further, for the games that the participant did make a re-tasking of the UVs to counter the unreliable information, the re-tasking action was initiated faster in games that resulted in wins compared to games that resulted in losses. Additionally, those re-tasking orders involved the use of fewer UVs in the games that resulted in wins.

Most recently, the results of the RoboFlag experiments have been used to model the changes in subjective mental workload as a function of interface manipulations (Parasuraman, Galster, Squire, Furukawa, & Miller, in press). A computational analysis using task-network modeling and Monte Carlo simulation provided results that aligned with the empirical data from the third experiment on flexible and restricted interfaces (Squire, *et al*., 2004).

MIIIRO

The Multi-Modal Immersive Intelligent Interface for Remote Operation (MIIIRO) is an operator interface for planning and controlling unmanned aerial vehicles (UAVs), unmanned tactical aircrafts (UTAs) and other remote systems (see http://www.ia-tech.com/miiiro/ or Tso *et al*., 2003). Per the website description;

> "MIIIRO consists of (1) a community of intelligent agents which are aimed at reducing work and information overload, (2) an immersive environment which induces a sense of presence in the engagement area, and (3) multi-modal inputs, including head tracker and joystick, which enable efficient interactions. Intelligent agents are used to integrate, assimilate, and present data, and interact with the operator to plan and control the remote systems and mission payloads. An innovative open multi-agent architecture is being developed to facilitate communication and coordinate activities among the intelligent agents. The intelligent agents are implemented in Java while the virtual worlds in Java3D and VRML.

> MIIIRO can also support human factors experiments on UAV missions. It provides the planning and trial capabilities for conducting experiments on the effects of multi-modal interfaces, multiple vehicle supervisory control, levels of automation, levels of system fidelity, and levels of time delay (information update rate) on human performance in supervising a system that locates and identifies ground-based targets during a hypothesized future multiple UAV mission scenario."

MIIIRO has been used by AFRL to examine the effects of levels-of-automation and automation reliability on the number of UVs that could be supervised by a single operator (Ruff, Calhoun, Draper, Fontejon, & Guilfoos, 2004). It has also been used to evaluate the effectiveness of target symbology as a function of increasing number of UVs supervised (Nelson, Lefebvre, & Andre, 2004). AFRL's research plan is to utilize MIIIRO to evaluate models of human-interaction with automated systems, the efficacy of automation aiding, human performance metrics in multiple task engagements, and task re-engagement strategy implementations invoked by operators after primary task interruptions. Due to its reliance on scripted behavior, MIIIRO has a high level of experimental control. It is mid-range on the realism scale, although this is tenuous because interfaces do not yet exist to control more than one UV at a time. The tractability level is flexible and can vary depending on the script and tasks the operators are asked to perform

DDD

DDD was originally developed by Aptima Corporation, in conjunction with the Office of Naval Research to study how teams operate in complex and dynamic environments (see http://www.aptima.com/Projects/Distributed_Dynamic_Decision_making.html for more information). The original task simulated a military command and control context, where the decision makers own and operate various vehicles, such as helicopters, jets, tanks, and radar planes (Kleinman and Serfaty; 1989; Kleinman, Pattipati, Luh, & Serfaty, 1992). The DDD is a unique distributed multi-person simulation and software tool for understanding how high-performance teams operate in complex decision-making environments. Unlike typical platforms that focus on specific and highly structured task domains, the DDD was designed to capture the essential elements of many different team tasks, and to allow the experimenter to vary team structure, access to information, and control of resources. One type of user interface for the DDD is shown in Figure 2.

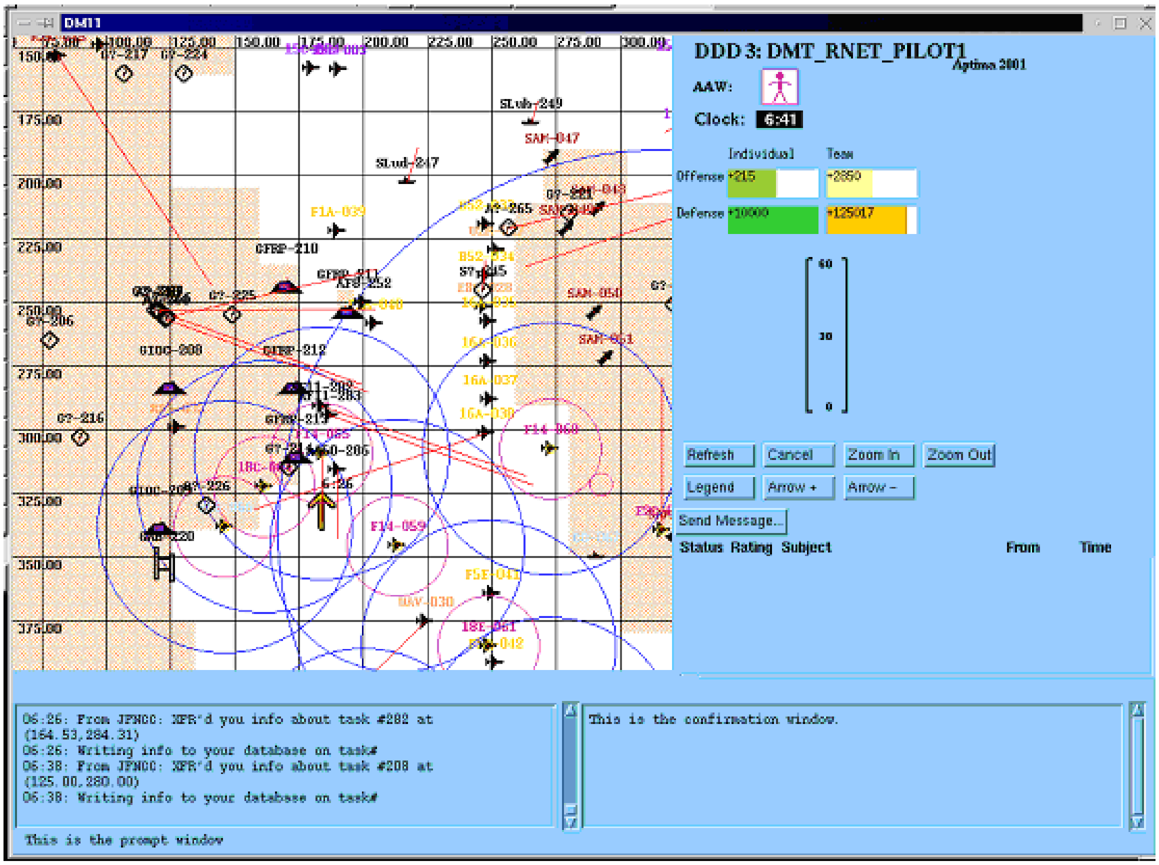

Figure 2. One type of user interface used with DDD.

The DDD allows for a substantial degree of experimental control while maintaining a low to moderate degree of realism. Like the MIIIRO STE, DDD is flexible on the tractability scale depending on the required tasks. The task loads in DDD scenarios can easily be manipulated by changing the number, type, timing and uncertainty associated with the tasks that need to processed. Additionally, organizational structures can be manipulated by changing authority levels, ownership of assets, communication variables, information availability variables and team membership variables (MacMillan, Entin, Hess, & Paley, 2004). This flexibility has allowed for widespread use of DDD in a number of varying research domains (see link above for a partial list of publications in various domains).

Aptima, Inc. is currently engaging in the development and delivery of a visual scenario generator (VSG) under a Phase II Small Business Innovative Research (SBIR) to AFRL. The principal goal of this effort is to refine the methodology and design the supporting tools that will improve the utility of DDD in simulation-based team research and training. Specifically, one objective is to enhance the Air Battle Management (ABM) capabilities of DDD. Another objective is to increase the capability of the DDD to collect performance data and generate after action reviews. The payoff for this effort will be the ability to generate ABM scenarios quickly, without having to delve deep into the source code. Once delivered, this will enable the ABM team at AFRL the ability to examine issues relating to time critical targeting, team formation and cohesiveness properties, and team decision-making quality and the resultant effectiveness of the decisions that were generated. In addition, tools that promote collaboration among teams will also be evaluated.

ADVANCED RESEARCH

BMC2

The BMC2 lab (formerly known as the Multi-sensory Overview Large-scale Tactical Knowledge Environment (MOLTKE) lab) is a medium-fidelity simulation of an Airborne Warning And Control System (AWACS) environment. The laboratory consists of six workstations arranged in two rows of three facing each other, similar to a console arrangement on the AWACS E-3 aircraft. Each workstation consists of two 900 MHz computers running Microsoft Windows 2000 and the Solipsys Prototype AWACS Display (PAD) software, one 19" flat panel-display, keyboard, mouse, programmable keypad, audio control panel, and two footswitches. The six workstations are connected to an experimenter's control station. The experimenter's control stations is composed of several computers running the AuSIM audio system, A/D and D/A converters, the Multi Source Correlator Tracker Lite (MSCT) software, and the experimental control software.

The primary purpose of the BMC2 lab is to examine the readiness of potential technologies in ABM. Operators on these platforms typically use interfaces that are manually intensive, cluttered, and require a significant amount of verbal communication. The BMC2 lab has the capability to portray high degree of realism while maintaining a suitable degree of experimental control. This combination usually restricts the tractability to applied research efforts.

Figure 3. The BMC2 lab configuration.

Given the problems listed above in ABM, the BMC2 lab has been instrumental in evaluating the effectiveness of spatial audio displays (Nelson, & Bolia, 2003), and speech recognition (Guilliams *et al*., 2004) in simulated air battle management environments (see also Vidulich *et al*., 2004). Spatial audio, mission phase and chatter level served as experimental factors in the evaluation of spatial audio displays. Ten trained Air Weapons Officers participated in an experiment that emulated a Close Air Support mission. Speech intelligibility was measured as the dependent variable. The results indicated that speech intelligibility was degraded during the more demanding experimental conditions and that spatial audio moderately alleviated this degradation. Additionally, faster response times for the correct identification of critical call signs were demonstrated when spatial audio was present.

A similar ABM scenario was used to evaluate the maturity and appropriateness level of speech recognition technology to offset some of the workload experienced by current operators. Twelve trained Air Weapons Officers participated in the scenario and the results suggested that speech recognition significantly reduced the amount of time operators took to complete their set-up, initial target transmissions, and strike package repairings. Additionally, the subjective workload ratings by the operators suggested that the speech recognition control condition engendered lower workload ratings in all phases of the task and were pronounced during the retargeting phase. The results of these two applied evaluations suggest that these technologies may be mature enough to start planning the transition to field operations.

The BMC2 lab is currently examining the utility of using advanced display technologies (head-mounted displays and multi-layer displays) in ABM environments to help ameliorate the primary display clutter problem experienced by most ABM platforms. A series of experiments are planned that will require that participants engage in information retrieval activities to get applicable data in order to re-task strike aircraft. The information sources will become more complex and dynamic as this research progresses.

FUTURE CAPABILITIES

CTT

Often STEs are augmented with new capabilities to determine if those capabilities provide an additional benefit to the operator. AFRL is committed to the development of these capabilities and the execution of programs that evaluate the potential benefits they may provide.

The purpose of the Collaborative Technology Testbed (CTT) is to permit the systematic evaluation of collaborative interface technologies (e.g., instant messaging (IM) and chat, virtual whiteboards, automated workflow management) and their effects on team performance, communication effectiveness, shared situation awareness, and decision effectiveness. The CTT is designed to support a program of basic and applied human factors research in the context of network-centric BMC2. Work domains such as these are characteristically communication-intense, fast-paced, rapidly changing, and replete with information that is often incomplete, inaccurate, and uncertain. Accordingly, one of the primary challenges involves the identification of interface concepts and technologies that will enable teams of operators to execute efficient and effective tactical problem solving and decision making. Of particular relevance are those interface technologies that facilitate group communication and the rapid acquisition, maintenance, and sharing of tactical situation awareness. Collaboration technologies that may be particularly effective in this domain include:

Chat and Instant Messaging – real-time text messaging with the ability to form mission teams, share information (pictures, data, applications), and review chat sessions

Video and Tele-Conferencing – real-time video and tele-conferencing to provide face-to-face interaction between operators

File and Application Transfer – ability to share files and applications within and between platforms and operators, providing seamless interoperability

Large Scale Shared Video Display – large, flat panel video displays for use by groups of operators working as a team where have a common battlespace picture is key

Data Capture and Replay – real-time video and data capture that can be reviewed, marked-up, and shared with operators throughout the battlespace

Data Visualization and Manipulation Tools – allows the manipulation, augmentation, and shared visualization of the battlespace content, especially by operators who are geographical distributed throughout the battlespace

Interactive Whiteboards – shared, interactive virtual whiteboards that are integrated with other displays and interface technology

Broadcast and Alerts – real-time broadcasts and alerts, and the ability to subscribe and publish these types of messages

Automated Workflow and Mission Timelines – real-time, mission-specific tracking of time-critical events, opportunities, and relevant assets; includes shared graphical interfaces and associated alerts, warnings, and broadcast capabilities

Interactive Intelligent Agents – intelligent agents that are used to find and gather information, alert operators to important events, provide briefings, and serve as a personal assistant for team coordination and workflow management

Opinion and Polling Tools – real-time assessment of operator opinion, also can be use for quick post-mission and effectiveness assessment for aiding decision makers

Automated Decision Support Tools – recommendations and alternatives for real-time individual and team decision making.

The CTT comprises numerous operator workstations, portable workstations (laptop PCs, Tablet PCs and PDAs), shared large displays, interactive whiteboards, head-mounted displays, spatial audio intercoms, and various collaborative software tools. In order to accommodate rapid reconfiguration of workstation layout and team structure, the CTT employs a commercially-available wireless local area network (WLAN) using the standard IEEE 802.11 protocol. In addition, various devices throughout the CTT take advantage of Bluetooth technology, which provides an additional means of achieving wireless connectivity between PDAs, laptops, workstations, and printers.

The CTT is outfitted with several COTS collaboration software packages, including InfoWorkSpace (IWS) and Microsoft Office Live Communication Server 2005 (LCS2005). Both software suites provide numerous collaboration capabilities including, but not limited to, instant messaging (IM) and chat rooms, tele-conferencing, video-conferencing, application and file sharing, and shared workspaces such as bulletin boards and virtual whiteboards. Additional software packages provide capabilities as such as automated workflow, content and knowledge management, advanced data visualization, intelligent agents and data mining utilities, and knowledge locators and opinion polling tools.

As described above, the CTT is designed to enable operator-in-the-loop simulations involving network-centric air battle management scenarios, in which collaborative tools enable teams of operators. This concept is illustrated in Figures 4a-c, which depicts three different levels of collaboration – face-to-face, local remote, and distant remote.

Face-to-face collaboration (Figure 4a) involves team configurations in which operators are physically collocated and are able to engage in face-to-face communication with other members of the team. This is indicated in the figure by same-color teams. Clearly, one of the advantages of face-to-face communication is that it permits the use of non-verbal cues such as facial expressions, body language, and emotion, which may be important for assessing level of common agreement, team situation awareness, and even the suitability of team decisions.

Local remote collaboration involves communication between operators who are separated by physical distance (i.e., not collocated), but share a physical environment (see Figure 4b). Local remote communication will be facilitated by collaboration tools such as IM and chat, video- and teleconferencing, shared displays and interactive virtual whiteboards, as well as file and application sharing. Dynamic work domains that require frequent temporary participation by operators may greatly benefit from these technologies, especially if primary roles and responsibilities mandate that operators stay at their workstations. For example, as depicted in Figure 4b, it may be necessary for individuals or teams to temporary join other groups to aid in problem solving or decision making. In this case, operators belonging to the blue and gray teams join the green team using collaborative technologies.

Distant remote collaboration involves communication between geographical distributed teams. This situation is illustrated in Figure 4c, which represents a network-centric battlespace scenario, involving real-time synchronous communication and collaboration between teams on the air battle management platform, AOC, UAVs, and ground forces. Collaboration technologies believed to enable such a scenario include automated workflow tools, intelligent agents, decision support aids, automated content and knowledge management systems, IM, tele- and video-conferencing, and shared interactive situation displays.

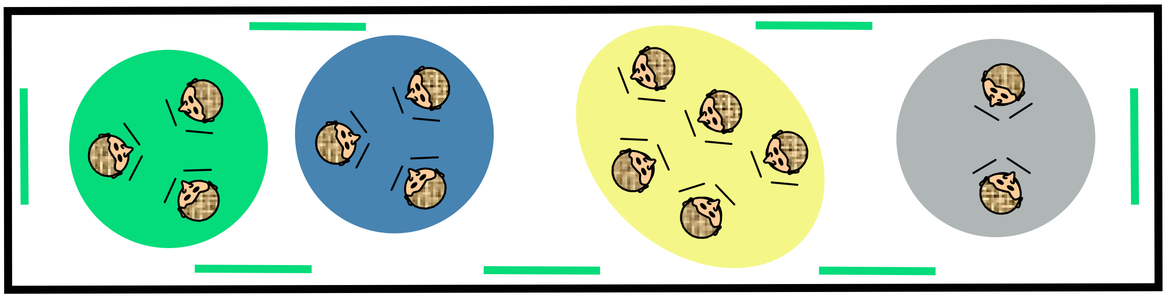

Figure 4a. Face-To-Face Collaboration. Conceptual layout of generic multi-mission air battle management platform, in which teams are arranged to leverage **face-to-face** collaboration and augmented by advanced collaborative interface technology such as instant messaging, workflow management, shared large displays (indicated by thin green rectangles), virtual whiteboards, etc.

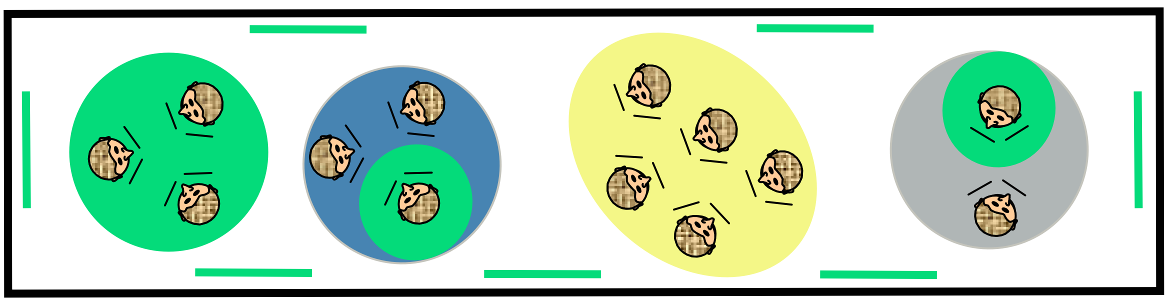

Figure 4b. Local Remote Collaboration. **Local remote** collaboration permits operators to remain at their primary workstations while synchronously collaborating with another team. In this case, operators belonging to the blue and gray teams have temporarily joined the green team.

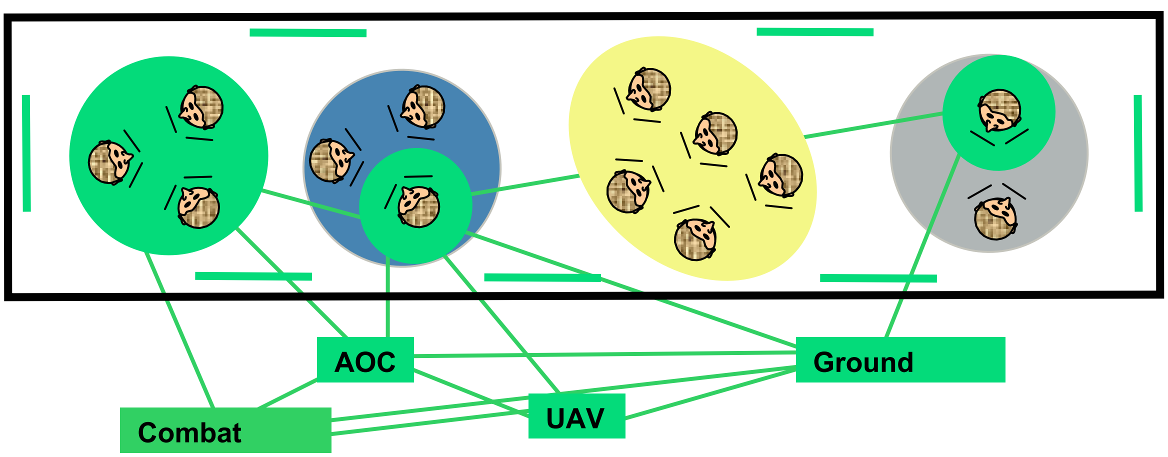

Figure 4c. Distant Remote Collaboration. **Distant remote** collaboration enables synchronous, parallel, multi-team communication and information exchange. In this case, collaboration technologies permit shared situation and collaboration decision making across the battlespace constellation in support of network-centric concepts of operations such as time-critical-targeting.

DISCUSSION

The STEs described here are all used in AFRLs effort to evaluate presently available technology and theoretical concepts in the C2 domain. The transition of these technologies and concepts usually follows a prescribed path; start at the lowest level of tractability and the highest level of experimental control and realism. After examination, if those technologies or concepts can potentially provide the end-use operator a new or enhanced capability they will be tested in a STE that is increased in realism and is more applied, forgoing the reduction of experimental control for the potential payoff. This transition often includes a demonstration or trial period in an actual operational environment to refine the experimental design that is executed in the STEs used for advanced research.

Because there is usually a large manpower investment in the acquisition of STEs, it is prudent to evaluate all of the potential capabilities of the STE under consideration. One of the potential capabilities that was considered before the investment in each of the STEs described above was the ability to have connectivity with other STEs. The STEs described above all have the obvious capability to function as stand-alone research environments. Some however have the ability to operate in conjunction with other STEs. For example, the BMC2 lab may want to include the MIIIRO STE in an experiment on the ability of ABM operators to supervise UVs remotely to gather and evaluate information relevant to a potential re-tasking order. This capability should not be overlooked when considering a potential STE.

C2 environments have unique requirements that need to be considered in research programs. These requirements do not need to overly restrict the opportunities researchers have in choosing appropriate STEs to use in conducting their research. However, careful and prudent evaluations of the amount of experimental control that is desired, the amount of realism that will suffice, and the tractability of the research questions will be helpful in determining if a particular STE will be suited for individual research programs.

REFERENCES

Army Science Board (2003). *Ad Hoc Study on Human Robot Interface Issues* Washington Headquarters Services, Directorate for Information Operations and Reports, Arlington, VA, 2002.

Brehmer, B., & Dörner, D. (1993). Experiments with computer-simulated microworlds: Escaping both the narrow straits of the laboratory and the deep blue sea of the field study. *Computers in Human Behavior*, 9(2-3), 171-184.

Cooke, N.J. & Shope, S.M. (2004). Designing a synthetic task environment. In S. Schiflett, L. Elliot, E. Salas, & M. Coovert (Eds.) Scaled worlds: Development, validation and applications. Aldershot, England: Ashgate Publishing.

D'Andrea, R., and R. Murray, R. (2003), The RoboFlag Competition. In *Proceedings of the American Controls Conference*, June, 2003, pp.650-655.

D'Andrea, R., and Babish, M. (2003). The RoboFlag testbed. In *Proceedings of the American Controls Conference*, Denver, CO, June 2003, pp. 656-660.

Ehret, B. D., Gray, W. D., & Kirschenbaum, S. S. (2000). Contending with complexity: Developing and using a scaled world in applied cognitive research. *Human Factors, 42(1)*, 8-23.
Galster, S.M., & Bolia, R.S. (2004a). Exploring the relationship between decision appropriateness and mission effectiveness in a simulated command & control task. Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting (pp. 448-452). Santa Monica, CA: Human Factors and Ergonomics Society.
Galster, S.M., & Bolia, R. S. (2004b). Decision quality and mission effectiveness in a simulated command and control environment. In D. A. Vincenzi, M. Mouloua, & P. A. Hancock (Eds.), *Human performance, situation awareness, and automation: Current research and trends*, Vol. I (pp. 264-268). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Guilliams, N. M., McLaughlin, A. B., Vidulich, M. A., Nelson, W. T., Bolia, R. S., & Donnelly, B. P. (2004). An evaluation of speech recognition technology in a simulated air battle management task. In D. A. Vincenzi, M. Mouloua, & P. A. Hancock (Eds.), *Human performance, situation awareness, and automation: Current research and trends*, Vol. II (pp. 230-235). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
Kleinman, D. L. & Serfaty, D. (1989). Team Performance assessment in distributed decision-making. In *Proceedings of the Symposium on Interactive Networked Simulation for Training*, Orlando, FL
Kleinman, D., Pattipati, K. Luh, P., and Serfaty, D. (1992). Mathematical models of team performance: A distributed decision-making approach. In R. Swezey and E. Salas, (Eds.) *Teams: Their Training and Performance* . Norwood, NJ: Ablex.
MacMillan, J., Entin, E.B., Hess, K.P., & Paley, M.J. (2004). Measuring performance in a scaled world: Lessons learned from the Distributed Dynamic Decision-making (DDD) synthetic team task. In S. Schiflett, L. Elliot, E. Salas, & M. Coovert (Eds.) Scaled worlds: Development, validation and applications. Aldershot, England: Ashgate Publishing.
Miller, C., & Parasuraman, R. (2003) Beyond levels of automation: An architecture for more flexible human automation collaboration. In *Proceedings of The Human Factors and Ergonomics Society,* Denver, CO.
Miller, C., Pelican, M., & Goldman, R. (2000) 'Tasking' interfaces for flexible interaction with automation: Keeping the operator in control. In *Proceedings of the Conference on Human Interaction with Complex Systems.* Urbana-Champaign, IL.
Nelson, J.T., Lefebvre, A.T., & Andre, T.S. (2004). Managing multiple uninhabited aerial vehicles: Changes in number of vehicles and type of target symbology. In *Proceedings of the 2004 Interservice/Industry Training, Simulation & Education Conference: Human System Interaction*. Arlington, VA: NTSA(Paper No. 1536).
Parasuraman, R., Galster, S.M., & Miller, C.A. (2003). Human control of multiple robots in the RoboFlag simulation environment. *Proceedings of the 2003 IEEE International Conference on Systems, Man & Cybernetics*. Washington, DC: IEEE. (3232-3238).

Parasuraman, R., Galster, S.M., Squire, P.N., Furukawa, H., & Miller, C.A (in press). A flexible delegation-type interface enhances system performance in human supervision of multiple robots: Empirical studies with RoboFlag. *Special issue of the IEEE Systems, Man & Cybernetics: Human Robotic Interfaces*. Washington, DC: IEEE.

Ruff, H.A., Calhoun, G.L., Draper, M.H., Fontejon, J.V., & Guilfoos, B.J. (2004). Exploring automation issues in supervisory control of multiple UAVs. In D. A. Vincenzi, M. Mouloua, & P. A. Hancock (Eds.), *Human performance, situation awareness, and automation: Current research and trends*, Vol. II (pp. 218-222). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Schiflett, S.G., Elliott, L.R., Salas, E., & Coovert, M.D. (2004). Scaled worlds: Development, Validation and applications. Aldershot, England: Ashgate Publishing.

Squire, P.N., Galster, S.M., & Parasuraman, R. (2004). The effects of levels of automation in the human control of multiple robots in the roboflag simulation environment. In D. A. Vincenzi, M. Mouloua, & P. A. Hancock (Eds.), *Human performance, situation awareness, and automation: Current research and trends*, Vol. II (pp. 48-53). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Squire, P.N., Furukawa, H., Galster, S.M., Miller, C.A., & Parasuraman, R. (2004). Adaptability and flexibility are key! Benefits of the "Playbook" interface for human supervision of multiple unmanned vehicles. *Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting* (pp. 61-65). Santa Monica, CA: Human Factors and Ergonomics Society.

Sheridan, T. (1987). Supervisory control. In G. Salvendy (Ed.) *Handbook of human factors*. Wiley, New York, NY, pp. 1244-1268.

Tso, K.S., Tharp, G.K., Tai, A.T., Draper, M.H., Calhoun, G.L. & Ruff, H.A. (2003). A human factors testbed for command and control of unmanned air vehicles. In *Proceedings of the Digital Avionics Systems 22nd Conference*, 8.C.1-1-12

Vidulich, M. A., Nelson, W. T., Bolia, R. S., Guilliams, N. M., McLaughlin, A. B., Donnelly, B. P., Collier, J. S., Fix, E. L., Miller, B., Brown, B. D., & Poole, M. R. (2004). *An evaluation of speech controls for AWACS weapons directors.* (Tech. Rep. No. AFRL-HE-WP-TR-2004-0019).