

DESIGNING A SYNTHETIC TASK ENVIRONMENT

Nancy J. Cooke
New Mexico State University
Las Cruces, NM

and

Steven M. Shope
Sandia Research Corporation
Las Cruces, NM

INTRODUCTION

Synthetic tasks are "research tasks constructed by systematic abstraction from a corresponding real-world task" (Martin, Lyon, & Schreiber, 1998), p. 123). Performance on a synthetic task should exercise some of the same behavioral and cognitive skills associated with the real-world task. A STE (Synthetic Task Environment) provides the context for a suite of synthetic tasks. This environment offers a research platform that bridges the gap between controlled studies using artificial laboratory tasks and uncontrolled field studies on real tasks or using high-fidelity simulators.

In this paper we will first provide a brief overview of STEs, including how they differ from simulations and the costs and benefits associated with them. We then focus on the steps involved in designing an STE, based on our experiences designing a STE based on UAV (Uninhabited Air Vehicle) operation. We describe the design process that we followed, from using cognitive task analysis to understand the field of practice to the resulting STE and the research agenda surrounding it. Although there are undoubtedly other approaches to STE design, in the course of our design we felt the need for more guidance in this area, and so hope that our experiences can serve as a starting point or illustration for others.

STEs vs. Simulations

An STE can be considered a type of simulation, but philosophically differs from traditional simulations in terms of goals and resulting design decisions. Simulations typically recreate the work environment or the equipment or systems within that environment. An STE is "task centric" in that the goal is to recreate aspects of the task to differing degrees of fidelity. Thus, an STE may not have the "look and feel" of the operational environment, but instead requires the same thoughts and behavior of the operational task. Because tasks are often situated in rich environments, STEs often include simulations of systems required to support the task. However, the focus is on abstracting task features consistent with the purpose of the planned research associated with the STE and concomitant design objectives. As a result, several very different STEs can be based on the same real task by virtue of applying distinct filters, each associated with different objectives. Such is the case with the UAV task in which a variety of STEs have been developed that focus on various cognitive skills of individuals (e.g., Gugerty, Hall, & Tirre, 1998; Martin et al., 1998) and others, such as the one described in this paper, focusing on team cognition.

In addition, simulations often replicate the environment at the expense of the simulation's flexibility as a research tool. Researchers are limited in the degree to which they can alter or control the simulation and the measures that they can derive from it. STEs, on the other hand, typically incorporate multiple task scenarios, and often the ability to manipulate aspects of task scenarios, as well as flexibility in measurement. This increased flexibility is not so much inherent to the concept of an STE, as demanded by researchers who appreciate the benefit of trading off some aspects of fidelity for research flexibility (e.g. Fowlkes, Dwyer, Oser, & Salas, 1998; Cannon-Bowers, Burns, Sals, & Pruitt, 1998). Recently, researchers have cautioned against the use of simulations unguided by training principles or an understanding of the actual task requirements and have extolled the virtue of low-fidelity simulations that take such factors into account (Miller, Lehman, & Koedinger, 1999; Salas, Bowers, & Rhodenizer, 1998).

STEs and Fidelity

STEs, like high-fidelity simulations, can facilitate research in a safe and inexpensive setting and can also be used for task training and system design in support of tasks. They are also touted as providing a viable middle ground between overly-artificial lab research and uncontrollable field research (Brehmer & Dorner, 1993). In many ways, STEs seem like the best of both worlds -- the laboratory and the field. Alternatively, if they fail to meet the combined objectives of experimental control and sufficient representation of the task in question, they may instead capture the worst of both worlds.

Perhaps the biggest criticism levied against STEs does not concern lack of experimental control, but rather the issue of lack of fidelity. However, the labeling of traditional simulations as *high-fidelity* and of STEs as *low-fidelity*, makes little sense. STEs have been described as low-fidelity simulations, as opposed to traditional equipment-centric simulations. Indeed, in terms of replicating the features of the equipment, STEs may represent a lower fidelity mock-up. The lack of fidelity is tied to more general concerns about low face validity, but perhaps more importantly, low external validity and therefore generalizability to the situation of interest.

The face validity issue is addressed by Salas et al. (1998) who argue that face validity may dictate acceptance by users, but not necessarily success as a training or research tool. In addition, the low external validity concern may seem sound on the surface, but may break down if fidelity is considered more broadly. Fidelity is generally the match between the research environment and the specific environment to which results are assumed to transfer. The match, however, can be based on a number of dimensions including the equipment and the task requirements. Thus, fidelity is not necessarily a single feature that is high-or-low for a particular simulation, but rather a multidimensional feature that can ultimately result in contexts of mixed fidelity. That is, a simulation may be faithful to the equipment, but not to the task requirements. Kozlowski and DeShon (2002) make a similar point with regard to the importance of psychological fidelity. In light of the issue of external validity, it is important to determine the dimensions of the transfer situation that are relevant to the research questions to be generalized. A mixed fidelity simulation may have adequate external

validity, and thus generalizeability to the field of practice, if it is faithful to the relevant dimensions of the field of practice. Determining external validity then becomes a question of accurately identifying the relevant dimensions in the field of practice for the research questions. Generalizing results to other fields of practice amounts to identifying similar features along the same relevant dimensions in other fields of practice. It can then be assumed that the match is sufficient for research results to generalize to this environment. This enterprise of identifying and matching the features and dimensions among different work environments amounts to a theory of tasks or work environments.

Therefore, under this multidimensional view of fidelity, the labeling of traditional simulations as *high-fidelity* and of STEs as *low-fidelity*, makes little sense. Instead, STEs are typically high-fidelity with respect to the task and low-fidelity with respect to the equipment. Traditional simulations may more often be associated with the opposite pattern. External validity cannot be determined independent of the research questions. Research on cognitive aspects of a complex task such as decision making under stress, may best be addressed in a context that preserves the features of the task at the expense of the fidelity of the equipment. Alternatively, research directed at uncovering reasons for operational errors associated with a piece of equipment may generalize only in a context that faithfully replicates the equipment, perhaps at the expense of simplifying the overall task. The question of the external validity and extent of generalizeability of both traditional simulations and STEs needs to be addressed for each test-bed in the context of each research question.

DESIGNING AN STE

In the following sections we describe the STE design process, based on our own experiences in designing a UAV-STE (see also Cooke, Shope, & Rivera, 2001; Cooke & Shope, 2002). UAV operations, a task environment of increasing interest in both military and civilian aviation arenas, was selected as the field of practice, with the Air Force's Predator UAV operations serving as the specific work environment. The selection of the UAV work environment was also based on a need for a team STE, which was satisfied by the fact that this task requires team performance and is associated with interesting team issues involving rank differences, the team's interaction with automation, and the presence of hierarchical and distributed teams.

By way of brief summary, the Predator UAV is controlled by operators in a GCS (ground control station) who communicate with other groups concerning issues of data interpretation and airspace deconfliction. The major team members within the GCS include the AVO (Air Vehicle Operator) who operates the UAV, the PLO (Payload Operator) who operates the sensors, and the DEMPC (Data Exploitation, Mission Planning, and Communications Operator) who is responsible for mission planning. These individuals work together to accomplish the goal of navigating the UAV to a position to take reconnaissance photos of designated targets. Individual team members have access to information about the UAV flight system, sensor equipment, and the surrounding environment, by way of computer displays, hard copies, and communication channels.

The specific research questions to be addressed in this STE centered on an understanding and measurement of team cognition. Although a final test of the external validity of our STE awaits field trials in which the findings and measures derived in the STE context are applied to the field of practice, the design decisions that were made along the way took external validity and generalizability into account.

Step One: Gathering Information to Constrain Design

Before the abstraction of dimensions and features take place, it is necessary to acquire a thorough understanding of the field of practice and the research objectives. Together these constrain the abstraction process. In addition, other constraints may also surface, such as the expertise of the intended participant population and various technological constraints in replicating task features in the laboratory. This first step of our STE design process is depicted with the other steps that follow in Figure 1.

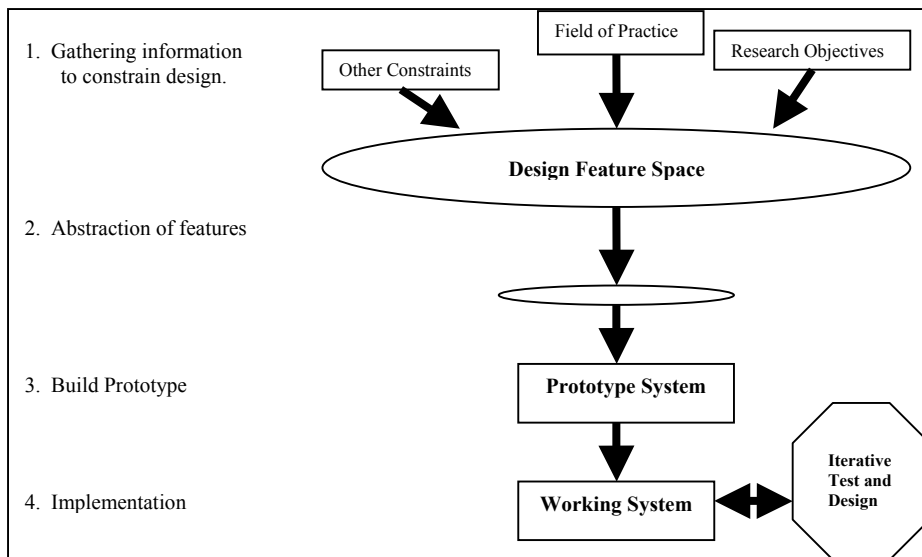


Figure 1. Our STE design process

Information from the field of practice. Paramount to the development of an STE is an understanding of the real task, based on documentation, interviews with experts, examination of other STEs for that work environment, and behavioral and cognitive task analyses. UAVs come in a number of varieties, but our focus was on the Air Force Predator, the first operational Air Force UAV, as well as one for which we had access to training data, operational specifications, and subject matter experts. In particular, we relied heavily on a cognitive task analysis (CTA) done on UAV operations at Indian Springs, NV (Gugerty, DeBoom, Walker, & Burns, 1999), information from actual Bosnia operations, an existing UAV-STE at Williams AFB in Mesa AZ (Martin, et al., 1998), and discussions with various investigators involved with these projects. Other information was obtained from various internet sites and unpublished reports, especially training documentation for the Predator UAV. From this information we began to develop an understanding of the UAV work environment.

Information from research objectives. The research objectives should also constrain the STE. Different features from the field of practice will be abstracted, ultimately resulting in a completely distinct UAV, given different research objectives. For instance, an STE that is developed for training UAV operators will appear quite different from one designed to develop models of team performance or one designed to test the effects of automation on team performance. In our case, the STE was designed to serve as a flexible task environment for the development and evaluation of measures of team cognition. It was also to be located in New Mexico State University's CERTT (Cognitive Engineering Research on Team Tasks) laboratory. This overall goal resulted in the following three objectives: 1) the STE should facilitate the measurement of team cognition, 2) the STE should provide a realistic task environment, and 3) the STE should provide an experimenter-friendly research test-bed. Each of these is discussed in depth below.

- 1) The STE should facilitate the measurement of team cognition. The *team* component of this objective means that the tasks and scenarios generated in the context of the STE should require a team of individuals. Team is defined as "a distinguishable set of two or more people who interact dynamically, interdependently, and adaptively toward a common goal/object/mission, who have each been assigned specific roles or functions to perform, and who have a limited life span of membership" (Salas, Dickinson, Converse, & Tannenbaum, 1992, pp. 126-127). Thus, it is important that each team member has a distinct role, and that without interdependence among individuals, the task cannot be completed.

The *cognition* component of this objective requires that the task environment pose cognitive demands both at the individual and team levels. For example, individuals may need to apply their own knowledge, and also share knowledge with fellow team members. In other cases, the individual situation assessments by team members may feed into the assessment of that situation by the team as a whole. In particular, the intended focus on the measurement of team knowledge requires that the STE support task that are knowledge- and information-intensive, with opportunities for variations in knowledge distribution across team members.

- 2) The STE should provide a realistic task environment. In general, this same objective motivates all STEs. It is based on a need to study behavior and cognition at a level in which it occurs in complex real-world tasks, yet to do so in the controlled environment of the laboratory. Specifically, for our purpose, it is important that the cognitive measures be developed and validated in a context that is relevant to real-world team tasks, and at the very least to the team tasks of the operational Predator UAV environment. This goal will be achieved to the extent that the STE exercises the same kinds of cognitive and team skills as the task in the field of practice.

Realism, like fidelity, is multidimensional and the dimensions of importance will vary with the research objectives. In our case, the STE needs to be realistic at the

level of team cognition and knowledge requirements. Under a different set of research objectives, realism in terms of face validity to actual operators or in terms of motor control of system devices may be more important.

- 3) The STE should provide an experimenter-friendly research test-bed. One of the purposes of collecting information on existing STEs and from investigators who use them was to situate the STE in a relatively unique niche, partially by benefiting from the lessons learned by others. Specifically, we uncovered a voiced need to make the STE more compatible with experimenter needs. In a sense, this objective involves moving away from a focus on simulation and toward a focus on research.

For instance, the STE should be flexible in supporting synthetic task variations of the type that may be required to manipulate a variable of interest or create a task scenario with slightly different cognitive requirements. It should allow, for example, rapid manipulation of task complexity, workload, cognitive demands, team process constraints (e.g., intrateam communication), as well as details of task scenarios. Further, it should allow the experimenter to intervene in the scenarios in order to insert measurement probes or to alter the task on-line (e.g., respond to varying levels of performance by increasing task difficulty during a mission). Similarly, the STE should be flexible in supporting a variety of cognitive and performance measures. Specific measurement capabilities should include embedded performance measures, logs of and post processing routines for computer and communication events, and rapid data access for immediate analysis.

Information from other constraints. There are typically other constraints associated with the limitations of the intended research environment. These limitations can be in terms of available space, budget, expertise of available participants, time constraints, and technology constraints. For example, there were some specific constraints associated with the fact that the STE was to be embedded as one of several STEs within NMSU's CERTT Laboratory. The CERTT hardware and software configurations provide a number of benefits in terms of overall flexibility and cognitive measurement, however, they also present some constraints. For example, the console displays have a general military-equipment look to them without using any hardware features uniquely associated with an actual UAV operator console. Instead, we migrated STE-specific functions to software displays. The hardware consoles have both computer and video monitors and general communications hardware interfaces that are appropriate and realistic to a wide variety of STEs.

The UAV STE was the first synthetic task designed for the CERTT Laboratory. The CERTT Lab's hardware design, development, and construction phases were being done in parallel with the UAV STE development. This permitted close developmental cooperation between this particular STE and the overall CERTT Lab design. This was a valuable feature not likely available to future STE development exercises. However, the

overall lab development goal was one that preserved a high degree of flexibility and adaptivity to meet the experimental goals of a wide range of STEs.

In addition, the CERTT Laboratory supports team tasks using up to four isolated participants and relies on a participant population of college students. Thus, the necessary STE skills and knowledge should be learnable with little or no background knowledge and without extensive training and practice. Time is also a factor here in that good team data should be obtainable within the time frame of a semester. This also puts limits on the amount of time that a team can spend in a training and skill acquisition phase.

Step Two: Abstraction of Features

The next step involves taking all of the information gathered from the field of practice, the research objectives, and other constraints and using it to abstract aspects of the task suitable for the STE. The process underlying such decisions, however, is difficult to explicate. That is, the identification of relevant dimensions in the field of practice and the nature of their replication in the STE, though critical to the external validity of the design, seems more like intuition or art, rather than the application of well-specified rules. For us, background knowledge in cognitive psychology and human factors and experience in experimentation, task analysis, and cognitive task analysis most likely fueled what looks like intuitive decisions. Although experts may disagree on some relevant dimensions, there is likely more consensus than not and the true success of this abstraction process cannot be determined without the field validation trials.

We identified those aspects of the actual task that we planned to emphasize (or maybe even exaggerate) in our STE. Due to our research objectives these were features of the task that centered on team cognition.

- 1) In the field of practice background knowledge and information relevant to UAV operations are *distributed* across team members. In other words, some UAV knowledge is uniquely associated with individual team members. For instance, the AVO has knowledge of the UAV's flight controls and equipment, whereas the PLO has knowledge of the sensor equipment. In a similar way, each team member is provided with information during the mission that is relevant to his or her role. This feature of distributed knowledge and information across a team was one that seemed relevant to team cognition and one that we chose to focus on in the STE version of the task.
- 2) Although information is distributed across team members in the field of practice, there are significant information interdependencies such that *knowledge and information sharing* among team members is required. For instance, the DEMPC determines the targets and provides navigation information to the AVO and PLO. The AVO, in turn, provides information relevant to the UAV that assists the DEMPC with route planning. The PLO communicates to other team members the navigation requirements for achieving target objectives. Thus, this feature of the

- task in which team members were required to share knowledge and information, not only filled the team constraint, but also the cognitive one, and was retained in the STE version.
- 3) UAV operations involve extensive *planning* on the part of the team in regard to the UAV's route. This planning occurs under multiple constraints. For instance, in order to plan the route, the PLO considers ground distance, altitude, depression angle, and slant range relevant to sensor operations; the AVO considers spatial layout and speed, and the DEMPC considers features of the bigger picture (i.e., restricted areas, ROZ boxes (Restricted Operating Zones), terrain height, weather, and speed). The planning process is a highly cognitive one and in the UAV task requires cognition also at the team level, for the various plans of individuals need to be integrated into a single plan. Thus, planning was retained as an integral component of the UAV-STE.
 - 4) Finally, due to the dynamic nature of the UAV operations environment (i.e., the presence of ad hoc targets, weather changes, UAV malfunctions, cloud cover) there is a need for *dynamic replanning* by individuals and the team as a whole. Team situation awareness is also required to plan and replan in an effective and efficient manner. Therefore we decided that the same kinds of dynamic events that occur in the field of practice needed to be preserved in the UAV-STE in order to provide the opportunity for dynamic replanning.

Step Three: Build a Prototype

The next step in our design process consisted of transforming the ideas, constraints, and abstracted features generated in the first two steps into a prototype STE. Although this step can be carried out using a number of prototyping techniques, tools, and software, we chose to use paper mock-ups, simple drawing software for screen designs, and extensive functional specifications. The difficulties that we encountered at this step were related to the fact that our design called for three independent, yet interconnected participant systems, as well as an experimenter system that would serve to monitor, manipulate, and measure aspects of the STE and even at times play a participatory role in the STE. Thus, a major part of our prototyping efforts involved determining interdependencies among the various systems.

We started this part of the design process by using the abstracted features and various design constraints to determine the minimum set of functions that each team member and experimenter would be required to perform. For instance, the PLO needed to monitor and adjust camera settings. We limited these to focus, shutter speed, aperture, zoom, and camera type. Each step of the way, we determined functionality with the goal of preserving task realism and especially functions that led to interdependencies (e.g., the shutter speed was tied to the airspeed of the UAV which was controlled by the PLO so this function was viewed as critical), but at the same time keeping in mind the background knowledge and training limitations of our participant population.

Once we settled on functions for each of the team roles we began the task of designing the interface. The CERTT Lab participant workstations each consist of two computer screens and so we decided to use one for primary control functions and the other for more secondary monitoring functions (i.e., warnings and alarms) of each team member. We made an intentional leap from realism where the interface is concerned in the interest of minimizing training time. The Predator interface is highly complex and can require a lengthy acquisition period (up to a year) for Air Force personnel to reach asymptotic performance. As our objectives centered on performance of a team with members already versed in their individual interface operation, the replication of the Predator interface was not only unnecessary (as functionality of the interface was preserved), but would be a hindrance to collecting data on teamwork and team cognition.

Next, we developed a list of software data structures needed to support the prototype design. These structures were further broken down into dynamic variables used to exchange information between the various computers on the network, startup variables, and data stores used to record performance information. An integrated system of data files was also defined to permit the experimenters to change startup parameters, record operational raw data, and to archive performance statistics derived from the various real-time measures. The initial DDE (network communications) architecture was considered at this time as well.

Functional specifications were drawn up in detail and provided information regarding interface displays and controls, measurement requirements, functional properties of subsystems, and information flow between start-up files and other subsystems. Our resulting prototype consisted of seven interconnected systems (two for each of three team members and one for the experimenter), and for each system, a representation of the screen and detailed functional specifications. The final version of the functional specifications was also key to the design of our training program for the UAV-STE.

Step Four: Implementation and Iterative Design

The next step involved presenting the functional specifications to the programmers who reviewed them and provided feedback on technological constraints and software architecture that would impact the basic design. This process began a series of discussions between STE designers and programmers and concomitant feedback-redesign iterations. The start of the implementation process further prompted additional feedback and redesign iterations that continued through the remainder of the UAV-STE development. Throughout this iterative design process it became clear that the successful implementation of the UAV-STE depended on effective teamwork skills (i.e., communication, coordination) among the designers and programmers. Hardware and budgetary constraints also played a major role in shaping the overall STE design. Reliability and straight-forwardness were two important design goals as well; not only for the developmental phase, but, for hardware/software maintenance and ease of future upgrades and or modifications.

THE RESULTING UAV-STE

In the resulting UAV-STE three team members work together to control and navigate the UAV and to take photographs of designated targets. The AVO is generally responsible for controlling the UAV's heading, altitude, and airspeed, as well as monitoring landing gear, flaps, and fuel consumption. The PLO takes photographs of designated targets under camera settings that are specified on the basis of destination waypoints and the current UAV speed and altitude that are broadcast from the AVO's computer to the other participants. The PLO also monitors the camera equipment such as battery, film, temperature, and lens. The DEMPC is responsible for mission planning and thus, has information that provides the "big picture" of the mission such as the world map that indicates all of the waypoints available. The DEMPC is presented with an initial route plan as the experiment begins, but has the capabilities to edit this plan as the experiment progresses.

Other characteristics of the UAV-STE support the objectives discussed previously. For instance, the interface requirements for the STE were based on the functions required to carry out the abstracted UAV tasks. Whereas the details of the displays and controls are somewhat dissimilar from the UAV interface in the field of practice, the new controls and displays support the new tasks and reduce the time needed to acquire proficiency. This change from the actual task reduces training time for the student participants and facilitates focus on the aspects of the task that involve team cognition. Flexibility is achieved through task parameters (e.g., mission time, fuel consumption rate, malfunction frequency) that can be easily changed by the experimenter via a set-up file. Spatial characteristics of the scenario (terrain features, weather obstructions, enemy activity, target priority, ROZ box location) are modified via a waypoint library that defines the scenario in terms of a world of waypoints, each with varying features. Also, in the course of developing the UAV-STE, performance measures were embedded in the task itself and other measurements were facilitated by the surrounding CERTT facility (i.e., communication flow, audio and video records, team process data).

As a result of the abstraction process, the UAV-STE, in comparison to the UAV task in the field of practice, exaggerates team cognition and interaction. Some aspects of the Predator UAV operations involve minimal or no team interaction and these have been eliminated or minimized in importance. For example, pre-flight planning that is DEMPC-focused and take-off and landing procedures that are AVO-exclusive have been eliminated in the UAV-STE. In addition, the UAV-STE retains the distributed nature of the task information and in fact, increases the compartmentalization of this information so that there are more cases in which information is uniquely associated with a single individual, thereby increasing interdependencies. To illustrate, in the STE, unlike the field of practice, the DEMPC is solely responsible for route planning from target-to-target and has sole access to a "big picture" map. In a similar way, information interdependence and the need for team coordination and knowledge sharing have been exaggerated by adding or modifying some rules that depend on information integration across two or three team members. For instance, the AVO and PLO both have altitude

constraints concerning the UAV flight path and sensor operation respectively. Both have to work together to simultaneously satisfy these constraints.

RESEARCH AGENDA FOR THE UAV-STE

The current research in NMSU's CERTT Laboratory using the UAV-STE focuses on the development and evaluation of measures of team cognition (Cooke, Salas, Cannon-Bowers, & Stout, 2000). We assume that a better understanding of team cognition, coupled with a suite of valid measures of team cognition, will facilitate the development of team training programs, selection processes, and system design for applications in which team tasks involve high levels of cognitive complexity.

Currently, we are tackling the issues surrounding a subset of team cognition, team knowledge. What does it mean for a team to think or have knowledge and how should that be measured? Figure 2 represents a simplified version of our view of some of the theoretical underpinnings of this construct. A review of the literature indicates that most measures of team knowledge have focused at the collective level. That is they measure the knowledge of individuals on the team and then aggregate these individual results to generate a representation of the collective knowledge of a team. Although we believe that collective knowledge should be predictive of team performance, it is also devoid of the team process behaviors (e.g., communication, coordination, situation awareness), analogous to individual cognitive processes that transform the collective knowledge into effective knowledge. This effective knowledge is what we describe as the holistic level and is associated with actions and ultimate team performance. One of our research goals is to test this model and identify ways to measure team knowledge at the holistic level.

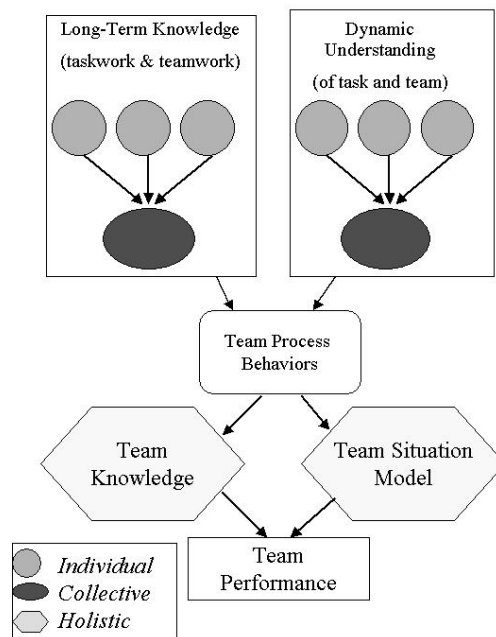


Figure 2. A framework for understanding team knowledge.

Other measurement issues have also captured our attention including 1) measures of team cognition suited to teams with different roles (i.e., heterogeneous teams), 2) methods for aggregating individual data to generate collective knowledge, 3) measures of team knowledge that target the more dynamic and fleeting situation models, 4) measures that target different types of team knowledge (e.g., strategic, declarative, procedural knowledge or task vs. team knowledge), 5) the extension of a broader range of knowledge elicitation methods to the problem of eliciting team cognition, 6) the streamlining of measurement methods to better automate them and embed them within the task context, and 7) the validation of various measures that we develop.

While measurement methods are developed and validated, we also apply these methods to better understand team cognition in tasks like the CERTT-UAV task. We are interested in various theoretical issues including the effects of factors such as workload, distributed vs. co-located teams, intrateam trust, communication barriers, and technological advances on team cognition. We have conducted several experiments using the UAV-STE (e.g., Cooke, Kiekel, & Helm, 2001).

The UAV-STE provides an ideal environment for the development of knowledge measures, because it offers experimental control over the information presented to the team members, the information acquired in training (i.e., knowledge), and team experience. Thus, for the purposes of measure development, a team can be constructed in which knowledge and information are completely and nonredundantly distributed across team members or in which all team members are trained on, and presented with, identical information. The ability to characterize, with any certainty, a team's experience and knowledge base is a luxury unique to the synthetic task environment. Characterizing tasks and teams in regard to knowledge distribution or experience in the field is challenging at best, and artificial lab tasks fail to provide the information richness needed for work of this kind. With this kind of information in hand, newly developed knowledge measures can be applied to team tasks in which the answers (i.e., content and distribution of knowledge) are known. A good measure should be capable of identifying those known features and distinguishing teams that have different knowledge profiles.

Thus, much of the research using the UAV-STE is directed at evaluating various knowledge measures in terms of their ability to identify and discriminate teams and tasks on the basis of their known knowledge bases. In addition, the measures can be evaluated with respect to their ability to predict team performance or team process behaviors. Other factors such as task complexity and training strategy can also be easily manipulated in the context of the UAV-STE, enabling specific hypothesis testing regarding the effects of such factors on team knowledge and performance.

SOME LESSONS LEARNED

We have now had several years to conduct research in the UAV-STE environment and have learned valuable lessons about the design process. In general, we feel that we did a good job at capturing the aspects of the Predator UAV task on which the STE is based. Empirical results as well as several judgments of face validity made by Predator

operators support this conclusion. We attribute much of our success to a highly collaborative design process that involved programmers, as well as researchers. It was often a struggle to come up with a design that met all of the task and research constraints while also satisfying software implementation constraints and hardware availability.

We also recommend that any initial STE be designed with future expansion in mind. Additionally, planning for potential heavy useage and rapid maintenance cycles is also desireable. For instance, the CERTT Lab is presently being used in excess of 14 hours per day, six days a week – far exceeding our original design criteria. The use of off-the-shelf hardware, wherever possible, is highly recommended to meet both of these issues. By using industry standard software designs, we are able to expand CERTT software modules without the need for massive system modifications. Keeping up with industry change is also important. For example, this particular STE requires 12 unique applications to be integrated to function as a single synchronous application. Although we originally chose DDE messaging to accomplish this objective, we are now upgrading to the DIS/HLA intercommunucations protocol in line with industry/military simulation standards

Finally, we have learned that there is a price to pay for the design of a tool that has maximum research flexibility. With increased flexibility comes increased complexity. As a result, there is a learning curve for researchers and experimenters to reach the level of competency that allows them to make use of the inherent flexibility of the system.

CONCLUSION

In summary, the UAV-STE, like STEs in general, contributes to the methodological arsenal available to researchers by facilitating controlled experimentation, while preserving many of the cognitive and behavioral characteristics of the task. The STE design process described in this paper is meant as an illustration, rather than a standard or prescription for design. However, this illustration does highlight the need for a consideration and integration of multiple simultaneous constraints originating from the field of practice, research objectives, and technology. We believe that the careful balance of these various, often competing constraints, is key to the design of a useful and valid STE.

The UAV-STE described in this paper emphasizes those characteristics of the actual UAV task domain that pertain to team cognition. Thus, it provides an environment in which cognitive team skills can be exercised to provide rich opportunities for the measurement of team cognition. The built-in flexibility of the UAV-STE allows for continuous refinement of scenarios and knowledge characteristics, as data are collected within the environment.

ACKNOWLEDGEMENTS

We are grateful to Erin Helm, Steve Hottman, Steve Jackson, Preston Kiekel, Harry Pedersen, and Merrill Sapp, who provided valuable input to this paper in the form of comments during a seminar taught by the first author on Team Cognition. Thanks also to Sara Caulkwell and Krisela Rivera who worked with the authors during the design process and to Ellen Hall and Leo Gugerty for providing us with their cognitive task analysis of the Predator operations task.

The work described in this paper was supported by a DOD DURIP grant F08671-9700815 and AFOSR grants F49620-98-1-0287 and F49620-01-0261.

REFERENCES

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, 171-184.

Cannon-Bowers, J. A., Burns, J. J., Salas, E., & Pruitt, J. S. (1998). Advanced technology in scenario-based training. In J. A. Cannon-Bowers & E. Salas (Eds.), *Making Decisions Under Stress*, (pp. 365-374). Washington, DC: American Psychological Association.

Cooke, N. J., Kiekel, P. A., & Helm E. (2001). Measuring team knowledge during skill acquisition of a complex task. *International Journal of Cognitive Ergonomics: Special Section on Knowledge Acquisition*, 5, 297-315.

Cooke, N. J., Salas, E., Cannon-Bowers, J. A., & Stout, R. (2000). Measuring team knowledge. *Human Factors*, 42, 151-173.

Cooke, N. J., & Shope, S. M. (2002). The CERTT-UAV Task: A Synthetic Task Environment to Facilitate Team Research. *Proceedings of the Advanced Simulation Technologies Conference: Military, Government, and Aerospace Simulation Symposium*, pp. 25-30. San Diego, CA: The Society for Modeling and Simulation International.

Cooke, N. J., Shope, S.M., & Rivera, K. (2000). Control of an uninhabited air vehicle: A synthetic task environment for teams. *Proceedings of the Human Factors and Ergonomics Society 44th Annual Meeting*, 389.

Fowlkes, J., Dwyer, D. J., Oser, R. L., & Salas, E. (1998). Event-based approach to training (EBAT). *The International Journal of Aviation Psychology*, 8, 209-221.

Gugerty, L., DeBoom, D., Walker, R., & Burns, J. (1999). Developing a simulated uninhabited aerial vehicle (UAV) task based on cognitive task analysis: Task analysis results and preliminary simulator data. *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting* (pp. 86-90). Santa Monica, CA: Human Factors and Ergonomics Society.

Gugerty, L., Hall, E., & Tirre, W. (1998). Designing a simulation environment for uninhabited aerial vehicle (UAV) operations based on cognitive task analysis. *Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting* (p. 1609). Santa Monica, CA: Human Factors and Ergonomics Society.

Kozlowski, S. & SeShon, R. (2002). ***chapter in this book***

Martin, E., Lyon, D. R., & Schreiber, B. T. (1998). Designing synthetic tasks for human factors research: An application to uninhabited air vehicles. *Proceedings of the*

Human Factors and Ergonomics Society 42nd Annual Meeting (pp. 123-127). Santa Monica, CA: Human Factors and Ergonomics Society.

Miller, C. S., Lehman, J. F., & Koedinger, K. R. (1999). Goals and learning in microworlds. *Cognitive Science*, 23, 305-336.

Salas, E., Bowers, C. A., & Rhodenizer, L. (1998). It is not how much you have but how you use it: Toward a rational use of simulation to support aviation training. *The International Journal of Aviation Psychology*, 8, 197-208.

Salas, E., Dickinson, T. L., Converse, S. A., & Tannenbaum, S. I. (1992). Toward an understanding of team performance and training. In R. W. Swezey & E. Salas (Eds.), *Teams: Their Training and Performance* (pp. 3-29). Norwood, NJ: Ablex.