Synthetic Task Environments for Teams: CERTT's UAV-STE

Nancy J. Cooke
Department of Psychology, MS 3452
New Mexico State University
Las Cruces, NM 88003

cooke@crl.nmsu.edu
(505) 646-1630

Steven M. Shope
US Positioning LLC
Arrowhead Research Park, Box 5040
Las Cruces, NM 88003

shope@uspositioning.com
(505) 646-2221

Background and Applications

Team tasks such as air traffic control, emergency response, and military command-and-control can be characterized as cognitively complex and embedded in a socio-technical environment. Recent views that are opposed to studying behavior and cognition apart from the natural context in which it occurs (e.g., Hutchins, 1995; Zsambok, 1997) have also created a need for a new research paradigm that preserves task richness and complexity of work domain, yet provides more experimental control than typical field settings. STEs (synthetic task environments) or "research tasks constructed by systematic abstraction from a corresponding real-world task" (Martin, Lyon, & Schreiber 1998, p. 123) offer a solution. The objective of STEs is to be able to reproduce behavior and cognitive processes associated with these complex settings in the laboratory where some experimental control and measurement capabilities can be preserved. An STE is a task environment in which a number of different task scenarios can be simulated. Compared to simulations, STEs tend to be more task-centric (i.e., faithful to behavioral and cognitive dimensions of task) and less equipment-centric.

A number of team STEs have been recently developed and funded primarily by the US military. In this chapter we describe a specific STE for teams and use it to illustrate the methodology involved in developing STEs. The STE we will describe is based on the task of ground operations of a UAV (uninhabited air vehicle) by three interdependent individuals: the AVO (air vehicle operator), the PLO (payload operator), and the DEMPC (data exploitation mission planning and communications operator) or navigator and mission planner. These individuals work together to accomplish the goal of navigating the UAV to a position to take reconnaissance photos of designated targets. We use this STE at New Mexico State University's CERTT (Cognitive Engineering Research on Team Tasks) Laboratory to study team cognition in this rich context.

Procedure

The procedure for designing the UAV-STE involved five steps, which we believe generalize to the development of STEs for other task domains:

1) Understanding the task in the field of practice

2) Understanding other constraints on the STE

3) Abstraction of task features

4) STE design

5) Validation of the STE

Step 1: Understanding the Task in the Field of Practice

The first step in designing a synthetic task is to acquire a full understanding of the task in the field-of-practice. There are a number of methods that can be used to achieve such an understanding, including interviews with subject matter experts, cognitive task analysis, naturalistic observation, and perusing technical manuals and other documentation. More generally, knowledge elicitation techniques (Cooke, 1994) used to elicit and model the knowledge of domain experts, can be helpful in this regard. It is likely that some research constraints will provide a filter for the information acquired under this step. For instance, because we were interested in developing a STE for the purpose of understanding team cognition, we attended most closely to aspects of the UAV task that were especially relevant to team performance or cognition.

For our project, information from the field-of-practice was gleaned largely through a cognitive task analysis done on UAV operations (Gugerty, DeBoom, Walker, & Burns 1999), information from actual operators, observation of 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 UAV.

Step 2: Understanding Other Constraints on the STE

Exactly what features of the actual task are abstracted for the STE version of the task depends not only on understanding the task and work domain, but also an awareness of the objectives of, and constraints on, the STE. As mentioned under the previous step, research objectives (e.g., understanding and measuring team cognition) provide one filter for task understanding and similarly serve as a filter for further selection of those features that will be represented in the STE. Another objective in our case was that the STE provide an experimenter-friendly research test-bed. In other words, STE scenarios should be easy to modify, experimental manipulation and control should be facilitated, and cognitive and behavioral measurement should be supported. These kinds of objectives also serve as filters for the abstraction of task features.

Some constraints are more practical in nature. Some of ours included the expertise and scheduling constraints of the university participant population and various technological constraints in replicating task features in the CERTT laboratory. The CERTT facility was developed in parallel with the STE for the purpose of providing a state-of-the-art environment for STEs that are focused on the measurement of team cognition. In total, the lab contains five interconnected participant consoles and a large experimenter control center. Each participant console contains two Windows NT machines connected to other CERTT computers through a local area network (Figure 1 displays the six screens). Participant consoles contain video monitors, headsets, and an intercom. The experimenter control center contains a variety of monitoring and data recording equipment including audio recording equipment, an intercom, headsets, video recording, and performance monitoring software. Thus, this hardware configuration provided additional constraints on the STE design.

Step 3: Abstraction of Task Features

Once constraints imposed by research objectives and pragmatic considerations of the research setting are identified, features are abstracted from the actual task which are within the bounds of these constraints. To this end, we identified those aspects of the UAV ground control task that we planned to emphasize (or maybe even exaggerate) in our STE. For instance, 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. This feature of distributed knowledge and information across a team was one that seemed relevant to team. Other team task features that we abstracted included knowledge and information sharing requirements, extensive planning under multiple constraints, and dynamic re-planning. In other cases, we deliberately chose to alter aspects of the original task due to constraints. For example, in the interest of minimizing training time we altered the human-computer interface substantially. The original interface is highly complex and can require a lengthy acquisition period (up to a year) to reach asymptotic levels of performance. As our objectives centered on performance of a team with members already versed in their individual interface operation, the replication of the 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 in our university setting.

Step 4: STE Design

The next step in the design process consists of transforming the ideas, constraints, and abstracted features generated in Step 3 into a prototype STE. We chose to use cardboard and paper mock-ups, graphics software for screen designs, and extensive functional specifications. We started by determining the minimum set of functions that each team member and experimenter would be required to perform and then began the task of designing the interface. Next, we developed a list of software data structures, dynamic variables, startup variables and files, and data stores needed to support the design. 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. Functional specifications were presented 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.

Step 5: Validation of the STE

We view STE validity as a separate, but related issue, to the validity of the measures and experimental results developed in its context. However, demonstrations of the validity of the latter, support the validity of the UAV-STE. The validity of an STE really concerns the issue of fidelity to the field-of-practice. Our view is that fidelity is multidimensional in that a task can be faithful to some aspects of the field, but not others. Thus, the UAV-STE should be faithful to the behavioral and cognitive aspects of the team task of operating a UAV.

Validity of this type can be assessed in various ways. First, to the extent that the cognitive task analysis and its translation into the STE are valid, we can infer that the UAV-STE that results is valid. So, any test validation of the cognitive task analysis reflects on the STE validity. Second, face validity can be determined through expert assessment. Assessment of the face validity of the CERTT UAV-STE by one expert UAV operator and a host of other individuals knowledgeable in UAV operations has provided preliminary support for STE validity. Third, to the extent that the STE is faithful to a particular aspect of the actual task, then domain experts should excel on those aspects of the synthetic task. To this end, we are interested in benchmarking expert UAV operator performance on our UAV-STE to determine the extent to which the STE is faithful to team cognition.

Advantages

- STEs allow for experimental control.

- STEs facilitate measurement capabilities.
- STEs can be faithful to dimensions of the task.
- STEs provide a rich test bed.
- Connecting STEs over the Internet enhances scope.

Disadvantages

- STEs are more complex than traditional lab settings and so some of challenges of field research remain.
- STEs are effective research tools to the extent that they are valid representations of the task.
- STEs can be expensive to build and maintain, though less so than traditional high-fidelity simulators
- STEs without the experimental infrastructure are ineffective research tools.

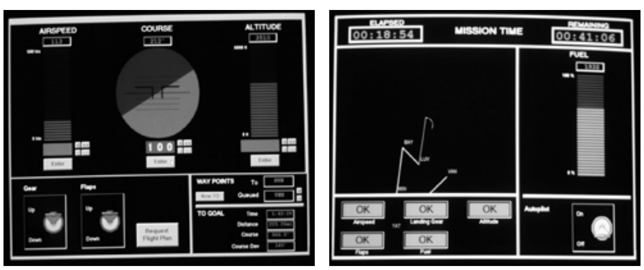

Figure 1a. Air vehicle operator screens.

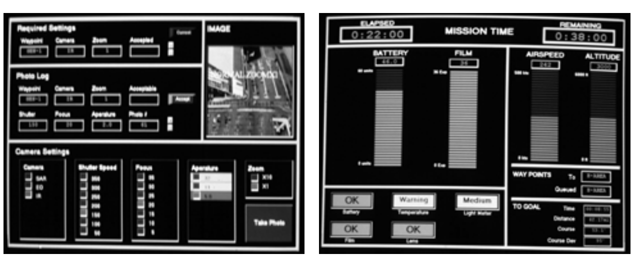

Figure 1b. Payload operator screens.

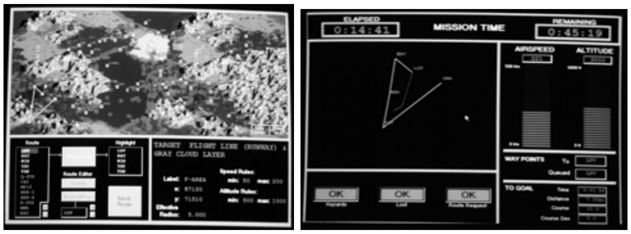

Figure 1c. DEMPC screens.

Figure 1. Two displays for each of the three participants in the CERTT UAV-STE.

Related methods

Because STEs abstract aspects of a task that are congruent with research objectives and other constraints, several very different STEs based on the same real task can emerge by virtue of these distinct filters. 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. Other related STEs exist that are based on other military command and control tasks such as ARGUS (Schoelles & Gray, 1998).

Standards and regulations

Communications Protocols

The current CERTT Lab uses a combination of dynamic data exchange (DDE) for intra-lab communications and distributed interactive simulation (DIS) protocols for interconnecting the CERTT Lab to outside entities. Both are industry standard protocols. In next generation CERTT designs, we plan to use DIS almost exclusively. DIS seems ideally suited for interconnecting remote STE sites as well as intra-lab communications. HLA (high level architecture) is also an acceptable protocol but perhaps more stringent than can be justified for STE operations presently.

Software

We presently use custom user-defined-objects in the task software which was developed in the Rapid™ Development Environment. This makes it difficult for outside parties to develop their own version of a participant console. One of our next-generation objectives is to use more ActiveX and other industry-standard objects to allow others to readily develop their own software modules for participating in the STE.

Approximate training and application times

The hardware for the UAV-STEs was completed in one year and the software in another year, however most of this time was spent in initial design. Given off-the-shelf software and hardware, the entire configuration could be set up in 90 days. Training experimenters to use the STE takes approximately 40 hours (most of it on-the-job training). Conducting experiments requires approximately 2 hours of participant training, followed by a series of 40-minute missions per team in which teams typically reach asymptotic levels of performance after the fourth mission.

Reliability and validity

Determining STE validity is an integral part of the procedure of designing an STE. Methods for determining validity were discussed in the earlier section. Thus far, validity of UAV-STE has been assumed based on its development using a cognitive task analysis of the actual task and several reports of sufficient face validity. In addition, research results established in this context have demonstrated predictive validity. Validity in terms of benchmarked expert performance remains to be evaluated.

Tools needed

In the initial CERTT Lab design, we used a combination of custom-built and off-the-shelf hardware. We used custom hardware when cost and/or product capability did not meet our design objectives. Currently, our philosophy is to use off-the-shelf hardware almost exclusively, especially for the participant consoles. The driving force behind this line of thinking is to facilitate outside and remote research groups participating in CERTT Lab experiments and operations.

In the initial CERTT Lab design, we used Windows-based software that would run on any PC using Windows NT or higher operating systems. This was not only on the STE task and participant applications, but also on the experimenter control station software and embedded measures. We plan to stay with Windows-based software, especially in view of the availability of image generation and virtual reality software components for Windows machines.

References

Cooke, N. J. (1994). Varieties of knowledge elicitation techniques. *International Journal of Human-Computer Studies, 41,* 801-849.

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). Human Factors and Ergonomics Society, Santa Monica, CA.

Hutchins, E. (1995). *Cognition in the Wild.* MIT Press, Cambridge, MA.

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). Human Factors and Ergonomics Society, Santa Monica, CA.

Schoelles, M. J., & Gray, W. D. (1998). *Argus, a system for varying cognitive workload.* Proceedings of the Human Factors and Ergonomics Society 42nd Annual Meeting. Santa Monica, CA.

Zsambok, C. E. (1997). Naturalistic Decision Making and Related Research Issues. In C. E. Zsambok and G. Klein (Eds.), *Naturalistic Decision Making* (pp. 3-16). Lawrence Erlbaum Associates, Mahwah, NJ.