# Hostile HCI: Research Methods for Unobservable Conflict Environments

**Katie Minardo**
kminardo@mitre.org
781-271-3244

**MITRE Corporation**
**Center for Air Force Command and Control Systems (CAFC2S)**
MITRE Corporation is a not-for-profit, federally funded research and development center (FFRDC) that provides technology advisement to sponsoring government organizations

**Biography:** Since joining MITRE, Katie Minardo has worked on flexible intranet applications, a room-sized visualization for resource-planning, an intelligent news service and a hand-held map application for the Air Force. Katie's current research interests include: the rise of search as a navigation method and field methods in adverse environments.

Prior to joining MITRE, Katie worked at Mayaviz (now General Dynamics) on flexible visualization of transportation logistics for U.S. Transportation Command, at Language Computer Corporation in ontology visualization for the Advanced Research and Development Agency (ARDA) and at Siemens Corporation on portal usability and cardiology-imaging software.

Katie holds a Masters in Human Computer Interaction and a BFA in Industrial Design, both from Carnegie Mellon.

Tactical Air Control Party (TACP) is composed of Air Force specialists who plan, request and direct air strikes at enemy targets. TACP specialists are typically embedded with an army unit and provide targeted strikes when enemies and friendly forces are in close proximity (additional information is available at: http://www.globalsecurity.org/military/agency/usaf/tacp.htm).

The TACP software allows ground-based TACPs to communicate with incoming aircraft, the command center and other units during conflict to coordinate air support. During field movements, TACP operators repeatedly troubleshoot connectivity and confirm communications to ensure support is available when it is needed. Additionally, the physical environment of the operator is challenging; they carry heavy pack loads and may also be in motion or wearing thick protective gear (gas mask, gloves, etc.).

As part of a software modernization effort, a human factors evaluation of the existing software was planned to build requirements for future software versions.

Accurate evaluation of the TACP software was a major research challenge. First, the software is used in a truly hostile environment where enemy fire is an imminent threat. Unlike routine field communications, such as requesting re-supply or submitting status reports, air support requests are most needed during the most critical moments. Users must be able to complete their task with high accuracy during the worst-case scenario.

For obvious reasons, the system cannot be observed for evaluation during actual field use. The task context, and its resulting level of distraction and adrenaline, cannot be accurately replicated in a laboratory. Additionally, the software requires communication with an intricate web of external systems and collaborators that may include command centers, other units and varied types of aircraft.

Research methods that are typically used in extreme environments are not applicable on the TACP problem space. Activities cannot be captured or documented for later review, deployed users are not available to provide detailed walkthroughs, and similar systems are difficult to identify. To best understand the overall context and the system details, a phased multi-method approach was chosen to work around the environmental challenges. While no single method provided perfectly accurate data, using data from the suite of methods built a realistic picture of the TACP system in use.

First, several expert walkthroughs were completed with operators to understand the system and tasks. A detailed Heuristic Evaluation (HE) was completed on both the hardware and software to identify potential problem areas and provide detailed feedback on the task design. After preliminary HE findings were presented to TACP program management and software contractors, additional field studies were planned to better understand the extended system context.

Second, two operator-training exercises were observed to gain relevant contextual information. While the conditions did not exactly mirror actual conflict scenarios, they provided an opportunity to observe the system in use. Operators completed key tasks, like troubleshooting connectivity, collaborating with the command center and communicating with the aircraft just as they would in the field. Because interruptions to actual operators were impossible, additional expert operators assisted during the observation to explain events as they occurred.

As an additional challenge, the operators involved in the training exercise were motivated to closely follow system procedures to ensure their (re-)certification on the equipment; this by-the-book behavior may not accurately reflect their field use. To reduce the training bias, the operators were interviewed after critical activities were completed. Each of the seven operators had recently returned actual combat so they were able to correct and clarify the accuracy of the training observation. These contextual interviews increased the applicability of our field observation. Detailed task models were developed and verified by the operators to ensure accuracy of the observation.

Lastly, our team evaluated similar Army software used to request air support. The Pocket Forward Entry Device (PFED) (additional information available at: http://www.defense-update.com/products/p/pfed.htm) is similar to a PDA and is used by the Army to perform support request tasks; both the conditions of use and the primary tasks were parallel in the two systems. The PFED system was evaluated using the same heuristic evaluation methodology as the TACP system; the systems were compared to assess advantages and disadvantages in their respective designs.

By using multiple methods, accurate usability assessment was possible for a hostile, rare and unobservable environment. By necessity, this research extends the methods typically used in difficult field environments. Since no methods were perfect, this method distributed inherent bias risk across multiple evaluations to build the most accurate evaluation possible for this difficult environment.