

Technical Paper

SUBMITTED TO:

**6th INTERNATIONAL COMMAND AND CONTROL RESEARCH AND
TECHNOLOGY SYMPOSIUM**

Department of Defense Cooperative C4ISR Research Program (CCRP)

Paper Title:

**The Evolution of the Headquarters Effectiveness Assessment Tool
(HEAT) and Its Applications to Joint Experimentation**

Authors:

**Dr. Richard E. Hayes Ph.D.
RADM Gary Wheatley USN (Ret.)**

**Evidence Based Research, Inc.
1595 Spring Hill Road Suite 250
Vienna, VA 22182**

**Phone: (703) 287-0383
Fax: (703) 821-7742
E-mail: wheatley@ebrinc.com**

Introduction

The Headquarters Effectiveness Assessment Tool (HEAT) was developed over 20 years ago by Dr. Richard E. Hayes. The system has proven to be an effective and robust method of evaluating Command and Control effectiveness and has been used in over 200 military exercises and Experiments over the past 20 years. In this paper we described the HEAT methodology and its evolution. We then further describe how the HEAT methodology will be used to apply in the US Joint Forces Command experiment Unified Vision 2001 (UV01).

Background

The U.S. Joint Forces Command (USJFCOM) has the primary responsibility to transform the DoD for the 21st century. Attendant to transforming DOD is the development of a new Command and Control (C2) architecture to support joint operations.

Experimentation is an integral part of the process in developing a new joint command and control architecture and critical to its success. These experiments provide opportunities to assess hypotheses for joint Command and Control processes and the means to determine efficiencies gained from new technologies. Indeed, they are the main tools that will enable JFCOM to transform the force and achieve the goals of Joint Vision 2020 (JV2020).

In support of USJFCOM Campaign Plan 2001, the Joint Concept and Experimentation Directorate (J9) has developed an experimental campaign plan that is presently building toward two major live validation experiments: Millenium Challenge 02 (MC02) and Olympic Challenge 04 (OC04). Supporting these are other experimental events such as the Unified Vision series. Unified Vision 2001 (UV01) is a concept refinement experiment that will integrate the Rapid Decisive Operations (RDO) concept and its supporting functional concepts. As part of the UV01 assessment process, Evidence Based Research (EBR) was asked to provide Headquarters Effectiveness Assessment Tool metrics. These were integrated into the UV01 Assessment Plan, utilized throughout the experiment, and analyzed for input into the Experimental Assessment Report. At the time of this paper's preparation the UV01 experiment had not yet commenced (UV01 ran

from 5-27 May 2001). Preliminary results will be briefed at the 6th International Command and Control Research and Technology Symposium (6th ICCRTS) but because of the Proceedings printing deadline, This paper was submitted before the experiment actually took place.

HEAT

HEAT was developed to render quantitative, objective, and reproducible effectiveness scores in order to assist those charged with; designing and running higher level headquarters. No existing tool allowed the measurement of effectiveness for such headquarters in alternative configurations (mobile, distributed, unitary, underground, airborne, etc.). Without such a tool, changes such as those designed to improve survivability could have the effect of "doing the enemy's job for him" by rendering the headquarters so small that it was ineffective. The tool has the added advantages that it will permit the collection of key information over time to improve our capability to predict theater headquarters effectiveness and lend insight into performance of command centers at all levels.

HEAT was designed for application to those headquarters that are primarily responsible for the planning, supporting and coordination of fighting forces, not direct warfighting. For US forces the effort focuses on echelons above corps, numbered fleets and major air commands. Application experience indicates that HEAT may also prove to be expandable to other types of subordinate headquarters. So far, however, neither the theory nor the experience of HEAT applications has indicated that such an expansion is an obvious growth direction.

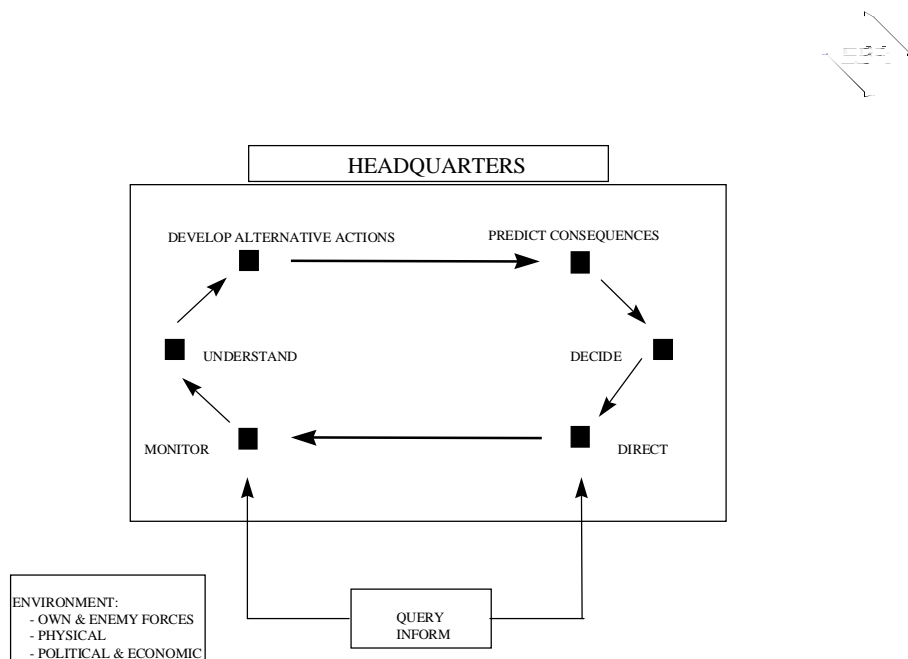


Figure 1. The HEAT Cycle

As displayed in the HEAT figure (Figure 1), HEAT treats a headquarters or command center as an adaptive control system that seeks to influence key elements of its environment by means of the plans (directives) that it issues to its subordinates. A full HEAT control cycle begins with the receipt of information about the environment, and continues with evaluation of the status of the situation *vis a vis* the current plan or situation, achieving an understanding or recognition of a need for change(s) to the plan, exploration of alternatives, a comprehensive and detailed plan development, and concludes with preparation and issuance of a directive. Note that HEAT control cycles will not always involve all of the above steps. As shown by earlier work conducted by Klein Associates and EBR beginning in 1989, there is a direct path (in many cases) from understanding to decision (to understand is to decide) either because of severe time pressure, or in the case of activation of a contingency, or because of the commander's knowledge and experience. Hence, measurement systems must be prepared to trace whatever decision cycle the unit follows, not fit its actions into a fixed structure. Metrics have been developed and applied throughout all phases of the HEAT cycle.

Following early work in US Navy SECOND Fleet exercises, HEAT was adapted for the US Army Battalion Commander Training Program (BCTP) as the Army Command and Control Evaluation System (ACCES). In numerous live exercises the number and variety of HEAT metrics evolved to over 250 different measures. These measures fall into several categories as discussed below.

Types of Performance Measures

In the Early 1980s, a Military Operations Research Society (MORS) Workshop set the standards for the metric taxonomy. They are still in use today, as reflected in the NATO Code of Best Practice for C2 Modeling published this year. These metric relationships are illustrated in Figure 2.

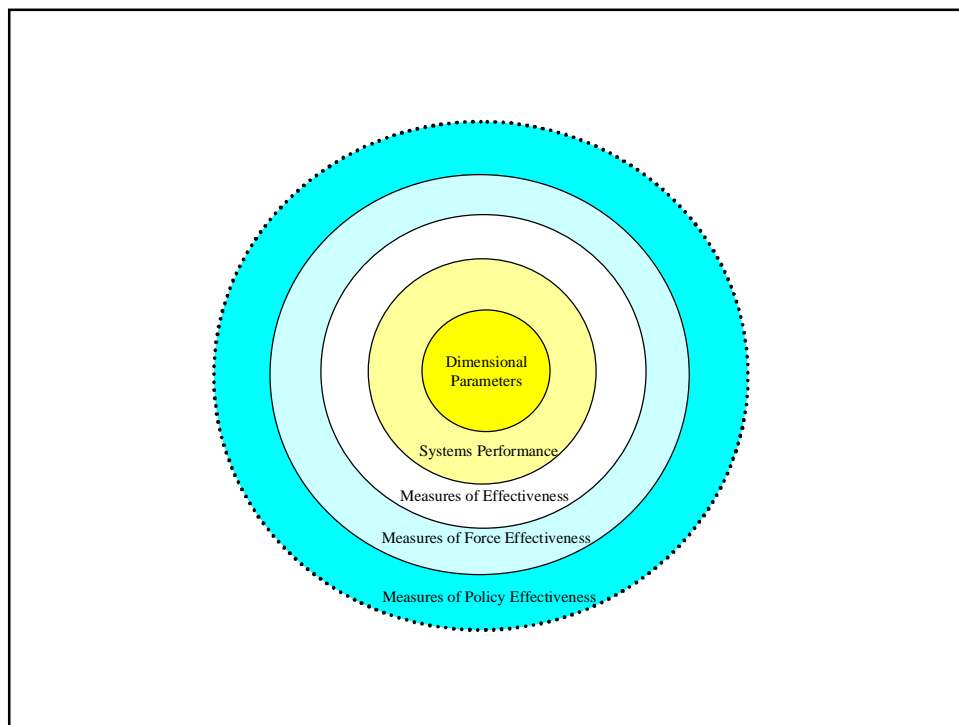


Figure 2. Measure of Merit Conceptual Relationships

At the most fundamental level, the measures may deal with the dimensional parameters of the systems themselves and whether these parameters have been achieved (e.g., bandwidth, memory, MIPs, etc.). At the second level, performance of the C2 system (C2

MoP) itself may be at issue (e.g., range, accuracy, reliability, etc.). Finally, measures of system and force effectiveness (MoSE & MoFE) deal with whether the system provides requisite capability for the force (a system of systems) to achieve its mission. For example, a C2 system (e.g., AWACS) may correctly identify a hostile air track and direct missiles or fighters to intercept. The C2 system can function perfectly but alone cannot destroy the target; the missile or interceptor must complete the end game. Representative the metrics include loss ratios, casualties inflicted, area taken, or other surrogates for military mission accomplishment. Note that the outer ring of the figure "Measures of Policy Effectiveness" may be outside the scope of the Joint Experimentation Program (JEXP), however in the context of Rapid Decisive Operations (RDO) and Effects Based Operations (EBO), this may require further discussion and analysis.

Measurement. Formal experimentation depends on precise measurement. Hence, these conceptual hypotheses must be "operationalized"; i.e., some observable phenomenon must be identified that allows assignment of values to the concepts in the hypotheses. This measurement must be both reliable (using tools and metrics that consistently provide the same values regardless of the observer or the observation situation) and valid (measure the concept under study). Measurement is often a major challenge, particularly when cognitive variables are involved. For example, measurement of "battlespace visualization" and "quality of military planning" are challenging on both validity and reliability grounds.

In UV01 the RDO Joint Force Headquarters will be composed of forward and rear elements that work together in a collaborative environment. Hence collaboration metrics are essential in determining the quality and effectiveness of that process. In MC00, for example, surveys were used to assess the collaboration process with little or no supporting quantitative data. But, collaboration can be accurately and quantitatively measured by focusing on information sharing and efforts to work together toward common purposes.

Collaboration metrics address team products, processes, and coherence. The product metrics concern the quality and timeliness of the collaboration product and the efficiency with which the product is produced. Here one can measure the time required to develop the product and the product's usefulness to the commander. For example, the time required for a CCIR response, and the accuracy of that response. The time recorded by itself is not particularly useful. This time metric becomes valuable when compared to a baseline, which could be the time required for our current process. In the absence of a baseline, the original time metrics become the baseline for further experimentation.

Process metrics concern how well the staff functions as a team. Representative process metrics are displayed in Table 1.

Table 1. Example Process Metrics

- Time after information is needed from other team members that is provided
- Time required by product provider to adjust product as requested by recipient
- Time needed to disseminate messages
- Percent of messages received that are relevant
- Number of instances where problems were not recognized prior to problem impacting team product
- Time required by team to modify tasks to adapt to new circumstances
- Time needed to disseminate the revised plan

Coherence metrics measure the staff's cognitive coherence and alignment. Table 2 lists some Team Coherence Metrics.

Table 2. Example team coherence metrics

- Similarity of interpretation of commander's intent among team members
- Number of centers of gravity that all team members identify; number that some but not all team members identify
- Accuracy of team member's knowledge of roles and responsibilities of other team members.
- Accuracy of identification of team members able to obtain specified information elements

Baselines

Quality experimentation requires appropriate baselines. Indeed, running experiments to evaluate new concepts, approaches, and structures without having a baseline makes it very difficult to demonstrate value added from the innovation. The overall design of UV 01 does not allow for collection in traditional headquarters structures and processes that can be compared with those used in the innovative headquarters that will be played. Hence, baselines for key functions will need to be drawn from other research and experimentation, or UV01 will itself become the baseline for further exploration of the experimental concepts.

UV01 Assessment Concept of Operations

The key to successful assessment is getting the right data. In UV 01 we will use a combination of methods to insure to insure that we obtain proper data and not miss any important events. These methods will include:

- Observers
- Participants surveys
- Screen captures
- Data scoring and analysis
- After Action Review (AAR) team

Observers

Trained data collection observers will be "joined at the hip" with key experiment participants to follow their actions and record key events. The key participant work stations will have "Y" audio connections installed so that the observers can plug in and hear exactly what the participant is listening to. Observations will be entered utilizing pre-formatted data collection sheets. An example of one of these sheets is illustrated in Figure 3.

In the event the observer wishes to record is not covered by a pre-formatted data collection sheet they will enter their results in a journal and later fill out the appropriate sheet for that observation.

The data collection sheets will be summarized in data scoring sheets that will facilitate further analysis of the collected data

Observer: <input style="width: 100%;" type="text"/>	Location: <input style="width: 100%;" type="text"/>	Cell/Function: <input style="width: 100%;" type="text"/>
DTG Start: <input style="width: 100%;" type="text"/>	<-- To -->	DTG End: <input style="width: 100%;" type="text"/>

Participant:	Role:	Sketch Layout:
1. <input style="width: 100%;" type="text"/>	<input style="width: 100%;" type="text"/>	<div></div>
2. <input style="width: 100%;" type="text"/>	<input style="width: 100%;" type="text"/>	
3. <input style="width: 100%;" type="text"/>	<input style="width: 100%;" type="text"/>	
4. <input style="width: 100%;" type="text"/>	<input style="width: 100%;" type="text"/>	
5. <input style="width: 100%;" type="text"/>	<input style="width: 100%;" type="text"/>	
6. <input style="width: 100%;" type="text"/>	<input style="width: 100%;" type="text"/>	
7. <input style="width: 100%;" type="text"/>	<input style="width: 100%;" type="text"/>	
8. <input style="width: 100%;" type="text"/>	<input style="width: 100%;" type="text"/>	

Team vs. Individual: <input checked="" type="radio"/> Team <input checked="" type="radio"/> Individual	Team Members Tasked? <input type="radio"/> Yes <input type="radio"/> No	Purpose of Goal Met? <input type="radio"/> Yes <input type="radio"/> No
--	---	---

Collaboration Mode: <input type="checkbox"/> Same Place <input type="checkbox"/> Different Place <input type="checkbox"/> Same Time <input type="checkbox"/> Different Time	Follow Up Meeting Scheduled? <input type="radio"/> Yes <input type="radio"/> No
--	---

Collaboration Method: <input checked="" type="radio"/> Brainstorm <input checked="" type="radio"/> Review/Evaluate <input checked="" type="radio"/> Negotiate <input checked="" type="radio"/> Consolodate <input checked="" type="radio"/> Integrate	Collab Purpose/Goal <input checked="" type="radio"/> DM <input checked="" type="radio"/> SA <div style="text-align: center;">Planning</div> <input checked="" type="radio"/>	Collaboration Tool Used? <input checked="" type="radio"/> Yes <input checked="" type="radio"/> No Tool Used: <input style="width: 100%;" type="text"/> Tool Features Used: <input style="width: 100%;" type="text"/>
---	---	--

Observed Technical Difficulties: <div style="height: 40px; border: 1px solid black;"></div>	Other Observations: <div style="height: 40px; border: 1px solid black;"></div>
---	--

Figure 3. Collaboration Data Collection Sheet.

Participants surveys

Each day participants will be required to fill out automated survey sheets (web-based) which will be entered into a special system allowing correlation of results.

Screen captures

Ground truth for each event in the experiment will be available at the location of the Joint Exercise Control Group (JECG). This is essential because it will be compared with the Blue and Red forces' view of the situation. The Common Relevant Operational Picture (CROP) display has been designed to permit instantaneous screen captures of the ground truth and the ability to compare it with other screen captures so that we can precisely measure the exact degree of situation awareness from both the Blue perspective and the Red perspective. In this manner we will be able to show a precise quantitative measure of situational awareness.

Data scoring and analysis.

As discussed under the observer section each observation sheet will be coded in such a manner that its results can be entered on data scoring sheets which will summarize related events and observations, and thereby permit extensive further analysis.

After Action Review.

In addition to the observations, screen captures, participant surveys, and data scoring we will have a special After Action Review team that will observe and evaluate the operational and tactical significance of each event as the experiment unfolds. While not a part of the formal data analysis their observations will be included in the overall assessment program. The AAR will conduct a debrief on each Friday afternoon during the experiment.

Summary

We expect that the combination of data collection methods and analyses will result in an accurate assessment supported by not only survey data but also by quantitative evidence that will clearly show the degree of efficiency, collaboration, and effectiveness of the

conceptual headquarters organization. The final results will not be available until several months following the experiment because of the large quantity and scope of the data that will require analysis. We hope the results can be rolled forward to support the Millennium Challenge 02 live experiment that will be conducted in June 2001. Unlike UV01 which will be conducted in a laboratory environment, MC02 will utilize large numbers of Joint Forces actually distributed across a large geographic area of the United States. The assessment and lessons learned from UV01 will be essential to the success of MC02.