

Bringing Analytical Rigor to Joint Warfighting  
Experimentation  
Joint Forces Staff College  
Norfolk, Virginia  
3-5 October 2006

Introductory Remarks:

Working Group 4  
*People as Experimental Assets: Applying SME and  
Human-in-the-Loop (HITL) Information*

John Rice  
Navy HPC / COMOPTEVFOR Detachment  
757 318 0671  
Ricej@cotf.navy.mil

## Problem

- There is a dearth of validated knowledge of human capacity and performance in contemporary military working environments.
  - Revolutionary tools in the work space
  - Changing work
  - Pressure for reduced manning
    - Changing work force
    - Redistribution of work and responsibility/authority
    - Assumptions about human capacity to “suck it up”

## WHY

- There is a lack of contemporary tools and infrastructure for observing human behavior in the work space.
  - Little invested in behavioral science tools
  - Assembly line “green card” experts not available
  - Trade-off of Human Centered studies for technology centered studies
  - Work spaces too complex and expensive to replicate in controlled “lab” environments

## Limitations (or Myths?)

- Human data is all subjective
- Human data is only qualitative
- Humans are too unpredictable
- Machines and Software can replace human
- Situational awareness can be measured by measuring machine characteristics
- Inability to distinguish between human performance studies and evaluation of a human’s performance exacerbated by military culture

## Limitations

- Human data collection is often an after thought to other material focused experiments
- Lack of consideration of the range of DOTMLPF components that effect HP
- Need to protect Human Subjects in Research

## Some Issues

- Assumptions that human behavior is not subject to scientific methods of assessment and evaluation
- Over reliance on “SME” data and “survey” data collection instruments
- Number crunching = analysis (?)

## Limitations to value of SME sources

- SMEs seen as cheap human data collection method of choice for system focused experiments
- All participants of opportunity considered equal when viewed as SME sources of individual (or system) performance data.
- Lack of systemic SME selection strategies to achieve “creditable” SME pool output.
- Lack of post event SME accreditation process.
- Panels, formal focus groups and post event lessons learned analysis too costly or time consuming to conduct, collect or analyze.

## Need to augment SME sources

- Need real-time observation methods to collect independent event data for post event association with SME reports.
  - Video and voice recording
  - Keystroke and other human interaction capture
  - Timeline observations of critical events and associated human activity by behavioral science and subject expert pairs.

## Opportunities for Improvements

- Acquisition reform focused on mission capability
  - Machine/software capability is an enabler
  - Human System Integration emphases in DOD and Service directives
  - Human as a component in total system
- Could mandate humans as system components to be tested inside the mandatory interoperability test boundaries

## Opportunities for Improvements

- Could leverage system test environment for basic and applied human performance studies and evaluations
  - Event concurrent studies of opportunity
  - Post system/engineering event HP events leveraging networks and facilities used in engineering tests.
- Bayesian analytical methods
- Increase synergy of human across service and joint organizations.

## Potential Discussion Areas\*

- (a) How do you classify and compare the impact of SME knowledge and experience to the impact of other assets on the quality of your experimental results?
- (b) How do you manage the controlled and uncontrolled aspects of SME knowledge and experience WRT the design of your experiment?
- (c) What types of unique influences on performance that live operators have on an experiment add value to the results?
- (d) Do we post-validate the HITL results? How do you account for different aspects of DOTMPLF (e.g. level of training of participants versus actual anticipated users)?
- (e) What types of qualitative data should be collected to characterize human performance? What are the social and cognitive domain metrics that should be captured? How do you report these results?
- (f) How is HITL data captured and fed back into mission- and campaign-level warfighting models to enhance their own representation of the battlespace?
- (g) By conducting HITL experiments, does that slow down the analytical process? If so, what are the costs (i.e., time, effort, etc.) and benefits?

\*From conference Terms of Reference