Evaluating Human Performance in Command and Control Environments

Susan G. Hutchins, William G. Kemple, Gary R. Porter, Michael G. Sovereign
Naval Postgraduate School
589 Dyer Road
Monterey, CA 93943
(831) 656-3768
shutchins@nps.navy.mil

Abstract

Performance assessment, in the broadest terms, includes the quantification of any human-system interaction which impacts the accomplishment of a mission. This chapter will cover the challenges inherent in developing measurement instruments and techniques for assessing perfor-mance in command and control settings based on experience from several major U.S. Navy research programs. The path by which a question of interest drives the development of a measurement instrument will be presented and examples of instruments that have been successful will be described. Measuring complex cognitive decisionmaking, the advantages and disadvantages of collecting various types of data, and lessons learned will also be discussed. Measurement instruments developed for the Adaptive Architectures for Command and Control program, and other current Navy research programs, will be used as examples.

Extended Summary

Developing performance measures for command and control (C2) settings (both experimental and operational) is challenging for three major reasons. Humans are ultimately responsible for C2 activities and in many cases the commander relies upon or delegates portions of his C2 responsibilities to other humans. The second is that there is often no single "right" way to accomplish a task due to the latitude that is doctrinally provided to the on-site tactical commander to encourage initiative. The third reason is the variety of missions facing modern armed forces under peacetime conditions. For missions other than war, complex political arrangements and sensitivities can color all decisions. Even in war, joint and coalition doctrine make decisionmaking vastly complicated for the C2 commander, and thus researchers who are trying to sort relevant data from the fog of war. In short, performance measures are interde-pendent with the mission environment. Given the constraints outlined above, measurement issues are very difficult. For example, traditional outcome measures, such as the number of "kills," are not particularly useful with current scenarios, such as information warfare, peacekeeping and humanitarian assistance missions.

The difficulties of measurement, as described above, argue for a judicious blending of theory, modeling, and experiment with frequent iteration between the steps so that participation of experienced subjects can be applied to sequentially build a solid linkage of the scenario, the experimental treatment conditions, the hypotheses and the conclusions.

Instruments useful for performance measurement require a balance between experimental control on the one hand and operational realism on the other. Experimental control refers to the ability to structure the environment so that the data obtained will be clearly interpretable. It means the environment presented to the participants needs to be controlled so that extraneous factors (intervening variables) do not cloud the picture by influencing performance in ways that are not intended while, at the same time, ensuring that the scenario is not so sterile that operational realism is missing. Experimental control includes the idea that the measurements need to be valid within the experimental setting while reliably capturing data that will provide answers to the question/s of interest. Reliability refers to the idea that if the same events occur, the scores should be relatively similar. Validity refers to the degree to which the measurement instrument actually measures what it was designed to measure. Validity tests how well the measurement instrument fulfills its function. Operational realism requires that the instruments for measurement not intrude upon the decisionmaker's process.

Measurement of complex cognitive constructs requires more than traditional time and accuracy measurements. Current multifaceted military objectives require more complex behaviors such as developing situation awareness and acting as a supervisor and monitor of complex systems (vice performing a single decisionmaking task) within a hierarchical and Joint/Coalition command structure. Automated systems have produced significant changes in the types of decisions required by humans. A shift has occurred from jobs that once depended on sensory-perceptual and motor tasks to jobs that now emphasize cognitive tasks. Observing the "real" world is often replaced with a totally synthetic/auditory scene depicted on various sensor and decision support system displays.

One aspect of the changing nature of performance requirements driven by this new technology is the idea that advances in technology are taking over much of the manual control tasks of the past. The human role is shifting to monitoring these highly complex automated systems. Thus, the emphasis is now on such tasks as monitoring the operator's evaluation of expected state versus perceived state and on assessment of cognitive function. For many of these situations there will be more than one way to approach measurement of a task. For example, many different measure-ment instruments have been produced to measure situation awareness by various research programs.

Complex decisionmaking usually cannot be summarized by a single score or measure. Multiple measures (instruments) are required to examine the various facets of the decision problem, such as, process and outcome measures and team and individual measures. Some lessons learned include the need to pilot a new measurement instruments prior to its first use to uncover problems encountered in using them. Problems may include (a) lack of clarity regarding use of the form, (b) data that do not fit into prescribed categories, (c) more time is required than was allotted to complete the measure-ment instrument, (d) artificial constraints were imposed, and (e) the level of detail requested was not appropriate (too detailed or not detailed enough). It can sometimes be difficult to predict all possible responses that will be made by the experimental participants. This reinforces the need to pilot test a measurement instrument to obtain some idea of the range of responses the experimental conditions will elicit. Let the research question or the operational issue drive measure development. Sometimes a simple, direct approach is more effective than a

complex approach. When the goal of the experiment is to extend a theory, a more complex approach to measurement will often be necessary. When the goal is an applied objective, with more concrete questions to be answered, a more straightforward measurement instrument may be appropriate.