



OPERATIONAL TEST
AND EVALUATION

OFFICE OF THE SECRETARY OF DEFENSE
1700 DEFENSE PENTAGON
WASHINGTON, DC 20301-1700

OCT 19 2010

MEMORANDUM FOR COMMANDER, ARMY TEST AND EVALUATION
COMMAND
COMMANDER, OPERATIONAL TEST AND EVALUATION
FORCE
COMMANDER, AIR FORCE OPERATIONAL TEST AND
EVALUATION CENTER
DIRECTOR, MARINE CORPS OPERATIONAL TEST AND
EVALUATION ACTIVITY
COMMANDER, JOINT INTEROPERABILITY TEST
COMMAND
DEPUTY UNDER SECRETARY OF THE ARMY, TEST &
EVALUATION COMMAND
DEPUTY, DEPARTMENT OF THE NAVY TEST &
EVALUATION EXECUTIVE
DIRECTOR, TEST & EVALUATION, HEADQUARTERS,
U.S. AIR FORCE
TEST AND EVALUATION EXECUTIVE, DEFENSE
INFORMATION SYSTEMS AGENCY
DOT&E STAFF

SUBJECT: Guidance on the use of Design of Experiments (DOE) in Operational Test
and Evaluation

This memorandum provides further guidance on my initiative to increase the use of scientific and statistical methods in developing rigorous, defensible test plans and in evaluating their results. As I review Test and Evaluation Master Plans (TEMPs) and Test Plans, I am looking for specific information. In general, I am looking for substance vice a 'cookbook' or template approach - each program is unique and will require thoughtful tradeoffs in how this guidance is applied.

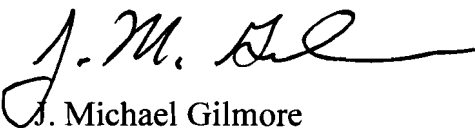
A "designed" experiment is a test or test program, planned specifically to determine the effect of a factor or several factors (also called independent variables) on one or more measured responses (also called dependent variables). The purpose is to ensure that the right type of data and enough of it are available to answer the questions of interest. Those questions, and the associated factors and levels, should be determined by subject matter experts -- including both operators and engineers -- at the outset of test planning.



Elements of experimental design that I am looking for when I approve TEMP's and Test Plans include the following:

- The goal of the experiment. This should reflect evaluation of end-to-end mission effectiveness in an operationally realistic environment.
- Quantitative mission-oriented response variables for effectiveness and suitability. (These could be Key Performance Parameters but most likely there will be others.)
- Factors that affect those measures of effectiveness and suitability. Systematically, in a rigorous and structured way, develop a test plan that provides good breadth of coverage of those factors across the applicable levels of the factors, taking into account known information in order to concentrate on the factors of most interest.
- A method for strategically varying factors across both developmental and operational testing with respect to responses of interest.
- Statistical measures of merit (power and confidence) on the relevant response variables for which it makes sense. These statistical measures are important to understand "how much testing is enough?" and can be evaluated by decision-makers on a quantitative basis so they can trade off test resources for desired confidence in results.

Design of Experiments is a structured process to identify the metrics, factors, and levels that most directly affect operational effectiveness and suitability and that should be reflected in detailed test plans. DOT&E is working with other members of the test and evaluation community to develop a two-year roadmap for implementing this scientific and rigorous approach to testing. I am looking for as much substance as possible as early as possible, but each TEMP revision can be tailored as more information becomes available. That content can either be explicitly made part of TEMP's and Test Plans, or referenced in those documents and provided separately to DOT&E for review.


J. Michael Gilmore
Director

cc:
DDT&E