

INTRODUCTION

Since July 27th, when I was sworn in as the Director, Operational Test and Evaluation, I have reviewed the state of Operational Test and Evaluation (OT&E) and Live Fire Test and Evaluation (LFT&E) within the Department of Defense. I have met with Service Operational Test Agency (OTA) Commanders; participated in programmatic reviews; approved test plans; witnessed tests; and provided my evaluation of programs' test adequacy, operational effectiveness, operational suitability, survivability, and lethality. This introduction provides my initial thoughts on the state of OT&E and LFT&E. It also provides my thoughts on recent legislative language that gives DOT&E additional responsibilities in a changing strategic environment. Finally, I will discuss my goals and priorities to address both the issues I have observed and the recent congressional mandates.

OBSERVATIONS

Title 10, U.S. Code states that the results of IOT&E should *confirm* that the system tested is effective and suitable for combat. I strongly believe that *OT&E should be a process of confirmation and not one of discovery*. Unfortunately, my first observation is that OT&E is too often the place where performance shortcomings and new failure modes are discovered. When problems are found late in the acquisition process, the cost to fix these problems is much higher than if they were discovered earlier. In addition, the time lost when problems are found at this stage can be substantial. When our forces need a new capability, this latter penalty may be more significant than increased cost.

Second, our acquisition and test and evaluation processes must accommodate a more rapid fielding of new weapons systems or improvement to existing systems. In some cases, I have seen this done well. For example, I have seen heroic efforts to quickly test and deliver capabilities to counter the improvised explosive devices (IEDs) prevalent in Iraq and Afghanistan. In other cases, the tempo of system fielding has been too slow, or systems have been fielded (for a number of reasons) that are not effective and/or not suitable to meet the needs of our warfighters. Both the acquisition and test and evaluation communities must work together to deliver capabilities at a pace consistent with the needs of the warfighters, while maintaining a "fly before you buy" (or "fly before you field") mentality.

A third observation is that suitability needs to improve. During the past three years, 9 of our 26 (35 percent) Beyond Low-Rate Initial Production (BLRIP) reports to Congress evaluated the systems as not operationally suitable. Even those reports that assessed the system as "suitable" were often based upon follow-on testing after deficiencies were found in the initial operational testing. I am obviously not alone in recognizing this as a problem. For example, the Secretary of the Navy said in a speech in August:

...operational suitability is fundamental to any assessment of an acquisition's warfighting contribution. If a product or system cannot perform its intended function in the real world environment, it will not provide value to the warfighter. Worse, if there is an expectation of capability that is not met, this could have disastrous implications for operational plans and execution well beyond the opportunity costs commonly identified.

Both DoD and the Congress have taken initial steps to help address this issue. DoD has made materiel availability a Key Performance Parameter and issued new guidance on how to achieve reliable, maintainable, and available systems. Further, recognizing that technological maturity is a key ingredient in obtaining reliable systems, the National Defense Authorization Act of 2005 requires that:

A major defense acquisition program may not receive Milestone B approval, or Key Decision Point B approval in the case of a space program, until the milestone decision authority certifies that--

- (1) the technology in the program has been demonstrated in a relevant environment;
- (2) the program demonstrates a high likelihood of accomplishing its intended mission.

INTRODUCTION

FY07 NATIONAL DEFENSE AUTHORIZATION ACT (NDAA)

The Congress included four provisions in the FY07 National Defense Authorization Act that will affect DOT&E's responsibilities.

- **REPORTS AT EARLY FIELDING.** Whenever the Department decides to proceed to operational use (or make procurement funds available) of a major defense acquisition program before it decides to proceed beyond low-rate initial production, I am required to submit a report to the Congress and the Secretary of Defense with respect to that program as soon as practicable.
- **ADDITIONAL INFORMATION IN REPORTS.** I may now include in my reports any additional information on operational capabilities that I consider appropriate based on the testing conducted.
- **GUIDANCE ON FORCE PROTECTION EQUIPMENT.** I am to provide guidance to, and consult with, the Secretary of Defense, the Under Secretary of Defense for Acquisition, Technology, and Logistics (USD (AT&L)), and the Secretaries of the military departments with respect to OT&E and survivability testing of force protection equipment, including non-lethal weapons.
- **REVIEW OF TEST AND EVALUATION POLICY.** In conjunction with the USD (AT&L), I must review Department of Defense policies and practices on test and evaluation in order to:
 - Reaffirm the test and evaluation principles that should guide traditional acquisition programs.
 - Determine how best to apply appropriate test and evaluation principles to emerging acquisition approaches.

GOALS AND PRIORITIES

Responding to my initial observations and the provisions of the FY07 NDAA, I have formulated the following goals and priorities, toward which my staff and I will work.

Enhance operational realism in early tests, including developmental testing. With changes in the acquisition process focusing on developing and fielding systems on a shorter timeline, I see a need to incorporate operational realism into developmental testing to gain operational insights and identify failure modes as early as possible. It is worth noting that the Service OTA Commanders share this goal. Early identification and correction of problems is the only way to move in the direction of OT&E as confirmation and away from OT&E as the venue for late discovery of problems. I intend to work with the USD (AT&L), the Services, and their OTAs to explore ways of better synchronizing developmental and operational testing to enhance the discovery process during developmental testing, and eliminate surprises in operational testing.

Improve suitability. *It is far more important for a system to be effective when it is needed than when it is available.* Effectiveness and suitability are not conflicting concepts and both must be considered and assessed early in the design and developmental processes. DOT&E must focus the efforts of the program managers and the test community to identify failure modes and impacts early in these processes. As a first step in improving our understanding of this problem, DOT&E is sponsoring studies to determine the actual costs of "unsuitability" and to determine the optimum place in the development cycle to insert resources to enhance suitability. We must also ensure our efforts to improve suitability encompass all aspects including reliability, availability, maintainability, human-machine interfaces, safety, transportability, training, etc

Provide timely performance information to the warfighters. Given the current strategic and operational environments, the acquisition process is changing, and the test community must adapt to those changes. While DOT&E must continue to support full-rate *production* decisions, we must also be able to provide decision makers with assessments that help them make informed *fielding* decisions when systems are being considered for operational use prior to the full-rate production decision. We must also assist the

INTRODUCTION

OTAs in ensuring our joint warfighters and commanders are aware of system capabilities and limitations when systems are fielded early.

Warfighters need to know about the capabilities and limitations of the system with respect to the spectrum of missions, threats, and possible scenarios. They need to know the tactics, techniques, and procedures that are most effective in getting the mission accomplished, especially if there will be rapid fielding. This requires a mission-focused, realistic, operational test. Even when full-scale operational testing is not feasible prior to early fielding, useful information on capabilities, limitations, techniques, and procedures can be developed during developmental testing, albeit with less confidence, if some degree of realism is added to gain operational insights. All this is to suggest that the best way to provide timely, useful information is a continual test and evaluation process that is mission-focused. The key will be early engagement with programs, more careful following of all testing (contractor and early developmental tests), and greater use of modern, sophisticated techniques for collecting, sharing, and evaluating information. Finally, we must recognize the inherent limitations in the evaluation that can be done prior to early fielding and ensure warfighters understand the implications and risks associated with employing such systems in combat.

Support the testing of force protection equipment. As mentioned above, FY07 changes to the U.S. Code now require my guidance and consultation with respect to operational and survivability test and evaluation of force protection equipment, including non-lethal weapons. Members of my staff have already been involved with the testing of force protection equipment, but it has been on an *ad hoc* basis. Specifically, this year they provided guidance and expertise in the joint testing and evaluation of helmet suspension systems. Similarly, my staff recently provided expertise, data sharing, and guidance in development of common test standards for personnel body armor, armor protection of wheeled vehicles, and support for testing against IEDs. The specifics of that support can be found in the LFT&E section of this annual report.

In order to standardize our involvement with force protection programs, I propose to develop a department policy on testing of force protection equipment. My staff will work with the Service OTAs to share expertise, make available technical advice, and provide support to expedite operational and survivability test and evaluation of those systems. I will then be in a position to provide informed counsel with respect to operational and survivability test and evaluation to the warfighting, acquisition, and fielding decision-makers.

I will need to broaden the expertise of my staff to be able to look at the different measures against which these systems, especially the non-lethal systems, will be evaluated. For example, typical programs evaluate system lethality and measure it against a threshold. Non-lethal weapons, on the other hand, must not only achieve an effectiveness threshold, but must also not exceed a safety threshold to ensure they are “non-lethal” in the expected method of employment.

Examine the allocation of operational testing resources. Everyone recognizes that there must be adequate resources dedicated to OT&E to ensure test adequacy and determine operational effectiveness and suitability. We must also recognize that the above initiatives are not “free goods” and that our greater and earlier involvement in testing means we must assess the consequences on our current approach to OT&E and LFT&E. My office has a government staff of about 45 individuals and we currently oversee 293 of the approximately 1,400 DoD acquisition programs. We must determine whether we have the right criteria for determining what programs are placed on oversight and/or whether we have adequate resources to perform our mission. Further, the additional responsibilities and new paradigms may require additional or different training and expertise.

INTRODUCTION

Assist in the review of test and evaluation policy. Within the next seven months, the FY07 legislation requires that DOT&E and USD (AT&L) review test and evaluation policy to reaffirm the test and evaluation principles that guide traditional acquisition programs and determine how best to apply appropriate these principles to emerging acquisition approaches. The policy question is significant, and has concerned the Department for some time. In fact, DOT&E and USD (AT&L) asked the National Academies, through the National Research Council, to study test and evaluation in the new acquisition environment. Their report, *Testing of Defense Systems in an Evolutionary Acquisition Environment*, was published in 2006. It provides a useful starting point from which to address the policy question.

Consistent with many of the thoughts discussed above, the study concluded that testing should be a continuum in which the primary goals should be to experiment, learn about the strengths and weaknesses of newly added capabilities, and facilitate the ability of program managers to use the results to improve overall system performance. To do this, early testing should emphasize the detection of design inadequacies and failure modes. Early detection of potential operational failure modes and the limits of performance will require DoD testing to develop alternative strategies for testing.

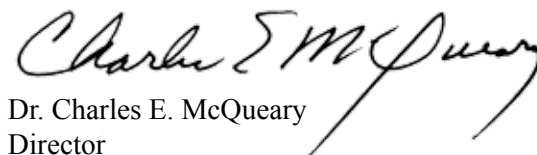
The study also recommended improving developmental testing; requiring contractors to share all relevant data on system performance; formally reviewing technology maturity before using it in a program; and increasing the expertise in areas such as combining information, software engineering, and physics-based and operational-level modeling. The final recommendation was to review proposed changes with a recognition that the current acquisition system already has a counterproductive incentive system and that the "...flexibilities inherent in the evolutionary acquisition process present greater opportunities for these counterproductive incentives to be expressed."

CONCLUSION

At DOT&E, there is a tradition of asking two questions: "Does it work?" and "How do we know?" Those questions mean demonstrating operational effectiveness, suitability, survivability, and lethality in full end-to-end tests with realistic missions, threats, tactics, and operations. To document and sustain that tradition, in FY06 we:

- Monitored 293 Major Defense Acquisition Programs (MDAPs) and special interest programs
- Approved 57 Test and Evaluation Master Plans / Test and Evaluation Strategies
- Approved 7 LFT&E Strategies and Test Plans
- Approved 53 Operational Test and Evaluation Plans for specific test events
- Delivered six Beyond Low-Rate Initial Production Reports (including one in October 2006):
 - EA-6B Improved Capability Three (ICAP III) Weapon System
 - Joint Biological Agent Identification and Diagnostic System
 - MH-60R Multi-Mission Helicopter (combined OT&E / LFT&E report)
 - Common Missile Warning System (classified)
 - Surface Electronic Warfare Improvement Program
 - Small Diameter Bomb Increment One (combined OT&E / LFT&E report)
- Delivered additional reports to Congress on Missile Defense

It is an honor and a privilege for me to be part of an organization that is a "key to weapons that work." With that in mind, I am pleased to present the 2006 Annual Report that follows.



Dr. Charles E. McQueary
Director