

INTRODUCTION

The President and the Congress have given me the opportunity to serve as Director, Operational Test and Evaluation for these last two and a half years. I have been honored and humbled to serve in this capacity and I thank them. This Introduction reports on what has been accomplished during that time to further the priority goals I first identified in the FY06 report.

The DOT&E goals I will discuss are as follows:

- Improve Suitability
- Enhance operational realism in early tests, including developmental testing
- Provide timely performance information to the warfighter
- Facilitate the allocation of adequate operational testing resources
- Ensure that DOT&E personnel are well trained

One of the chief mechanisms for progress has been to review and renew existing T&E policies. Actions we took include the following: developed new policy with respect to suitability, in particular, reliability; increased manpower authorization in DOT&E to address emerging needs and increased complexity of systems; established contacts within each Combatant Command to ensure the information is available to them from our Annual Reports, our Beyond Low-Rate Initial Production Reports (BLRIPs), and our Early Fielding Reports done in accordance with Sections 231 and 139 of the FY07 National Defense Authorization Act (NDAA); and improved our Action Officer training program.

The following discussion will provide insight into the direction I have set on behalf of the DoD and for this organization.

SETTING NEW T&E POLICY

As a result of congressional direction to review existing policy in light of the many new acquisition strategies and initiatives, the DoD issued a report in July 2007 on needed changes. In December 2007, the Under Secretary of Defense for Acquisition, Technology and Logistics and DOT&E established new T&E policy. The new policy recognized that the fundamental purpose of test and evaluation is to provide knowledge to assist in managing the risks involved in developing, producing, operating, and sustaining systems and capabilities. The new policy also recognizes that T&E measures progress in both system and capability development; that T&E provides knowledge of system capabilities and limitations to both the acquisition community and the user community; and that T&E expertise must be brought to bear at the beginning of the system life cycle to provide earlier learning about the strengths and weaknesses of the system under development.

The following policies were implemented and are now in DoD Instruction 5000.02, which was signed on December 2, 2008:

- T&E expertise must be brought to bear at the beginning of the system life cycle to provide earlier learning about the strengths and weaknesses of the system under development. The goal is early identification of technical, operational, and system deficiencies, so that appropriate and timely corrective actions can be developed prior to fielding the system.
- T&E shall be conducted in an appropriate continuum of live, virtual, and constructive system and operational environments.
- Developmental and operational test activities shall be integrated and seamless throughout the Engineering and Manufacturing Development phase.
- Evaluations shall take into account all available and relevant data and information from contractor and government sources.
- Evaluations shall include a comparison with current mission capabilities using existing data, so that measurable improvements can be determined. If such evaluation is considered costly relative to the benefits gained, the program manager shall propose an alternative evaluation approach. This

INTRODUCTION

evaluation shall make a clear distinction between deficiencies uncovered during testing relative to the approved requirements and recommendations for improvement not directly linked to requirements. A DOT&E approved LFT&E strategy shall guide LFT&E activity.

- Evaluations shall be conducted in the mission context expected at time of fielding, as described in the user's capability document. The MDA shall consider any new validated threat environments that will alter operational effectiveness.
- As technology, software, and threats change, FOT&E shall be considered to assess current mission performance and inform operational users during the development of new capability requirements.

I have asked the Services to begin to collect data on current programs in order to assess if any additional policy changes are necessary.

In July 2008, the Under Secretary of Defense for Acquisition, Technology and Logistics directed the Secretaries of the Military Departments and the Directors of Defense Agencies to establish an acquisition reliability improvement policy to address the problem of inadequate system Reliability, Availability, and Maintainability (RAM). This was a major step to address one of DOT&E's top priorities to which I now turn.

GOALS IN PRIORITY ORDER

1. Improve Suitability. To address the goal of making the IOT&E a means of confirming performance, rather than revealing new failure modes, DOT&E has worked to help identify failure modes and their operational impacts early in the design and development process. During 2007, DOT&E concluded that the key issue is inadequate system reliability, which is a key component of suitability. Contributors to reliability problems include: poor definition of reliability requirements, ignoring reliability in the Request for Proposal (RFP) and in contracting, and poor tracking of reliability growth during system development. Many of these problems occur long before the IOT&E, in program formulation, and in contractor and developmental testing. Added impetus to improve suitability came from a valuable Defense Science Board (DSB) Task Force effort in 2007, the final report for which was published in June 2008. One action of particular importance, stemming from that report, was the Under Secretary's July memo, mentioned above. In particular it directed new Service and agency policy to implement RAM practices that include the following:

- Ensure effective collaboration between the requirements and acquisition communities in the establishment of RAM requirements that balance funding and schedule while ensuring system suitability and effectiveness in the anticipated operating environment.
- Ensure development contracts and acquisition plans evaluate RAM during system design.
- Evaluate the maturation of RAM through each phase of the acquisition life cycle.
- Evaluate the appropriate use of contract incentives to achieve RAM objectives.

To aid the Services and agencies in this effort, the DoD developed the following:

- RAM Cost (RAM-C) Manual to guide the development of the requirements for the established Suitability/Sustainability Key Performance Parameter and its Key System Attributes. The RAM-C Manual will provide a consistent picture of sustainment operations so both designers and testers can better perform their functions. The cost aspect of the manual is important because the DoD has made ownership cost a key system attribute. (Operation and Support Costs account for 60-70 percent of the total ownership costs.)
- Contracting language to ensure that contractors are aware of the importance the government places on reliability and total ownership costs.

INTRODUCTION

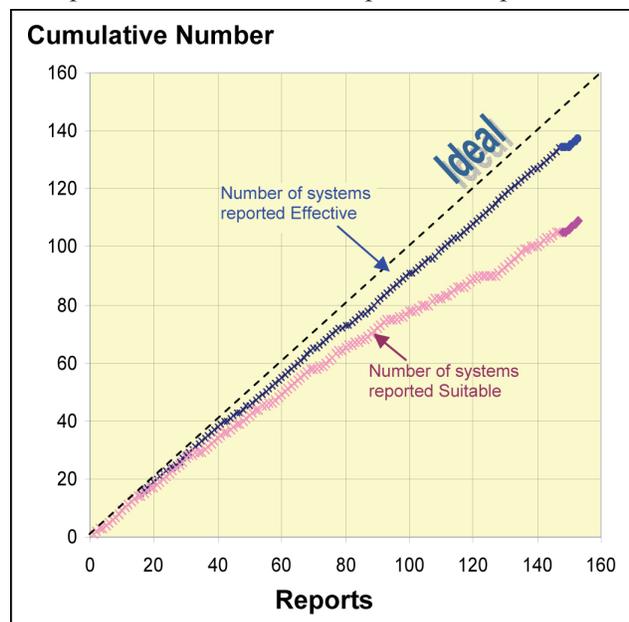
- RAM planning and evaluation tools first to assess the adequacy of the RAM program proposed and then to monitor the progress in achieving program objectives. In addition, we have sponsored the development of tools to estimate the investment in reliability that is needed and the return on investment possible in terms of the reduction of total life cycle cost. These tools include algorithms to estimate how much to spend on reliability.
- Workforce/Expertise initiatives to bring back government expertise that was lost when the importance of RAM began to be discounted. This includes refocusing the Defense Acquisition University on RAM training. For DOT&E's part in this effort, we have allocated four of the new positions we have been authorized to work with programs during the requirements definition process as part of the Joint Staff's Functional Capabilities Boards and will address RAM as part of that early influence effort. In addition, we are sponsoring training for OSD staff.

As mentioned before, a fundamental precept of the new T&E policies is that expertise must be brought to bear at the beginning of the system life cycle to provide earlier learning. Operational perspective and operational stresses can help find failure modes early in development when correction is easiest. A key to accomplish this is to make progress toward Integrated T&E, where the operational perspective is incorporated into all activity as early as possible. This is now policy, but one of the challenges remaining is to convert that policy into meaningful practical application.

In a separate action, DOT&E joined an effort to define best practices for reliability programs. Last year's report addressed how vital that effort was. Once agreed upon and codified, reliability program standards can logically appear in both RFPs and in contracts. Industry played a key partnership role in this effort. The standard, GEIA-STD-0009 has been approved, and on November 13, 2008, was American National Standards Institute certified. I see industry's increased commitment to address system reliability and suitability as evidence of growing momentum for improvement.

In summary, I remain convinced that each step in the development process can and should be used to improve suitability. While DOT&E is clearly engaged in the final operational testing of systems, we have teamed with DoD and industry partners to forge improvements in earlier steps.

As a practical matter, these steps make improvement possible, yet the results may be some time in



coming. This year, we provided eight BLRIPs. Of those, two of eight (25 percent) were not suitable for combat compared to 50 percent the year before. Some improvement might therefore be inferred, but it will be a while before a definite trend of improvement can be established. In what should become an annual reporting metric, the chart from last year's annual report has been updated with the data from FY08 (in bold) and shows improvement in the slope of the curve, which, in the ideal case would be a 45-degree slope.

2. Enhance operational realism in early tests, including developmental testing. The Defense Science Board (DSB) Task Force mentioned earlier examined the need to reinvigorate developmental test and evaluation. The final

report of the Task Force concluded that the problems in reliability can be corrected only by re-instituting a disciplined Systems Engineering process during design and development. The DSB suggested, as

INTRODUCTION

many others have, that integrating developmental and operational testing could help. Many of the DSB recommendations are now policy. Successful implementation of the policies will create more realistic and operationally-representative conditions in early testing, especially in developmental testing. Realistic stresses and loads will lead to earlier discovery of failure modes. Early operational insight and assessments can influence system design and reduce surprises in IOT&E.

As a metric of our progress toward achieving this goal, DOT&E was to provide operational insights gained prior to preliminary and critical design reviews and acquisition decision points. The chart below provides a 2008 baseline against which future progress can be measured.

METRIC	OF ALL RELEVANT PROGRAMS (FY08)	OF THOSE WITH MILESTONE THIS YEAR (FY08)	OF THOSE WITH MILESTONE NEXT YEAR (FY09)
The fraction of T&E Strategies and T&E Master Plans that test technology in relevant operational environments, including realistic threat environments, before Milestone B.	0.36	1 of 2 programs with Milestone A in FY08	Expect 4 of 4 programs with Milestone A in FY09
The fraction of programs that have a DOT&E letter report at Milestone B that assesses effectiveness, suitability, and survivability in a relevant operational environment.	0.02	0 of 7 programs with Milestone B in FY08	Expect 12 of 12 programs with Milestone B in FY09

I should emphasize that these low numbers over all programs indicate that the DoD only recently concluded that earlier OT&E involvement in the development cycle is necessary. The low percentages are metrics that are a baseline to track improvement as we move forward.

3. Provide timely performance information to the warfighters. Congress stimulated progress on this priority by requiring Early Fielding Reports when a system is committed to operations before a full-rate production decision. In FY08, DOT&E delivered three such reports in compliance with this particular part of Section 231 of the FY07 NDAA. Our goal was to provide timely and accurate assessments for fielding decisions and to make joint warfighters and commanders aware of system capabilities and limitations to performance and mission accomplishment. The DOT&E goal is that this information will be available for all systems that enter the field, fleet, or battle space.

We have established a classified website for these assessments (<http://www.dote.osd.smil.mil/assess/>) to make available DOT&E Annual Reports, BLRIP Reports, and Early Fielding Reports to the Combatant Commanders and others who have proper access.

In addition, we have established points of contact between DOT&E and each Combatant Command to ensure that joint warfighters and commanders are aware of the system capabilities and limitations, strengths and weaknesses for systems that might be deployed to them. Early fielding does not remove our responsibility to determine whether a system is effective and suitable for combat before the full-rate production decision. So DOT&E will continue to follow the Early Fielding Report with our usual BLRIP when the IOT&E is complete.

4. Facilitate the allocation of adequate operational testing resources. As I reported last year, my analysis of staffing levels indicated that DOT&E needed more resources in the form of experts. DOT&E requested, and was granted by the Deputy Secretary of Defense, an increase in staff level of 22 permanent positions. It will take time to fill these staff positions, but the process is well underway. With this increase, I believe that future Directors will be able to properly support the acquisition process and to

INTRODUCTION

respond quickly to Combatant Commanders' requests for support from our Joint Test and Evaluation Program. As noted earlier, four of the 22 positions will be focused on early involvement of T&E in the requirements and program formulation phase with an emphasis on RAM.

While DOT&E's augmentation is significant for its size, workforce augmentation remains a challenge in the Services where there are technical expertise shortfalls in the areas of Systems Engineering and testing.

During its review of test programs, my staff identifies any test-critical resource shortfalls. Test-critical resource shortfalls are those that meet the following two conditions: (1) if not available in time for IOT&E testing, would require DOT&E to declare the IOT&E inadequate, and (2) for which there is not an adequate program to develop the test capability. Only one test-critical resource shortfall has been so categorized and DOT&E has gone on record with the Navy for it: the Navy Multi-Stage Supersonic Target (MSST). The Navy response to DOT&E's memorandum of concern has been positive, leading to a contract award for development of the two-stage advanced anti-ship cruise missile target on August 22, 2008.

One other test-critical resource is worth noting because of its importance to adequate testing. This is the development of an adequate 5th Generation Fighter Target for the Air Force, Navy, and Marine Corps. Currently, the Air Force is using the QF-16 as an interim solution. A DOT&E-sponsored study is underway to determine if the QF-16 is sufficient or if an alternative, affordable solution is appropriate. The results of this study will affect both the F-35 and F-22 programs.

5. Training. To ensure that DOT&E personnel are well trained and prepared to meet the challenges presented by the evolving acquisition and testing environments, DOT&E continues to revamp its in-house training program.

Each DOT&E staff member is required to have an approved program for continued professional development, and the staff member's yearly performance appraisal will depend in part on completing that program. DOT&E now offers, as part of that professional development program, specialized training in RAM.

In another part of its professional development program, 10 DOT&E staff participated in the Deputy Secretary of Defense's Lean Six Sigma "Green Belt" training. Seven earned Green Belts.

EMERGING TEST MISSION AREAS: FORCE PROTECTION EQUIPMENT AND NET-CENTRIC AND SOFTWARE TESTING

Force Protection Testing

Based upon increased congressional interest in personnel body armor and combat helmets, the FY09 NDAA amended Title 10 Section 2366 to give the Secretary of Defense authority to designate programs for oversight pursuant to Section 2366 without restriction. The change mirrors the authority already granted the Director in Section 139 of Title 10 for operational test and evaluation oversight. In FY09, DOT&E will work with the Services to identify those programs that due to their direct contribution to warfighter lethality and survivability, particularly personal body armor and combat helmets, warrant DOT&E oversight under this new provision.

Based on previous legislation, I issued policy on force protection equipment and non-lethal weapons to the Services in 2008, establishing the framework for a collaborative and cooperative environment for the sharing of information and expertise, while meeting my statutory obligations. I believe that implementation of this policy will serve well to ensure that warfighters have the full spectrum of protection and munitions they need to have success on the battlefield of today and tomorrow.

INTRODUCTION

There were two notable examples of DOT&E involvement in force protection programs this year. DOT&E began oversight of Army testing of personnel body armor as a result of a congressional request. This request and subsequent direction by the Secretary to provide oversight was in response to the hearings held by the House Armed Services Committee (HASC) on June 6, 2007. The integrated product team formed to accomplish this task, consisting of DOT&E, Under Secretary of Defense for Acquisition, Technology and Logistics, the Army Test and Evaluation Command, and the program manager for Soldier Equipment, presented a two-phased approach to congressional staff at a November 14, 2007, meeting. Phase 1 consisted of ballistic testing in accordance with the solicitation and supported the Army's source selection process. Phase 2 includes additional ballistic testing to more rigorously characterize the ballistic performance of the plates. During 2008, Phase 1 testing was completed in accordance with test plans approved by my office and was adequate in scope and execution to support the Army's source selection process. The Army has awarded contracts for the production of enhanced small arms protective inserts (ESAPI) and XSAPI (improved ESAPI) plates to support First Article Test and Phase 2 testing. DOT&E submitted an interim report to Congress following the completion of Phase I testing. DOT&E will prepare an independent report to Congress following completion of this effort.

The second example was also a congressionally directed action, stemming from the FY08 NDAA. Congress directed the DoD to conduct a limited field user evaluation and operational assessment of qualified combat helmet pad suspension systems. After coordinating with HASC professional staff, DOT&E requested that the Army and the Marine Corps conduct independent tests. These tests were completed in the summer of 2008 and DOT&E submitted an independent report to Congress.

Net-Centric and Software Testing

As discussed last year, we have continued to work with U.S. Joint Forces Command (USJFCOM) to align joint testing and training roadmaps in the growing mission area of net-centric warfare. While our progress has been limited by major delays encountered by the largest pilot program (Net Enabled Combat Capability (NECC)), the very limited NECC testing accomplished this year underscored the need to test operationally relevant sets of capability in a live, virtual, constructive (L/V/C) continuum. Software updates to the Global Combat Support System-Joint and Defense Travel System programs also leveraged JFCOM's L/V/C capabilities, while the Air Force Operational Test and Evaluation Center conducted the Integrated Strategic Planning and Analysis Network IOT&E in concert with U.S. Strategic Command exercises.

More generally, software intensive systems such as next generation Command and Control systems and Enterprise Resource Programs consistently encounter significant problems that delay successful fielding because they fail to perform as expected in the final stages of testing.

The greatest challenge appears to be the lack of rigorous developmental testing. Too often, developmental testing resembles a feasibility demonstration with developers focusing on demonstrating that their product can work under a single set of circumstances rather than testing to ensure that the product will work under likely operational conditions. As a result, difficulties with data conversion from legacy systems, system interfaces, and the interface with the network "transport layer" are often under-emphasized.

There are three root causes of these problems. First, requirements often are not well defined or not available until the development is nearly completed. This handicaps the developer who should understand, at the beginning of development, the desired performance, the intended operating environment, and the already fielded systems with which it will have to work.

Second, development testing has not always represented a realistic environment. Some developers have assumed, because the DoD has moved to an Internet Protocol (IP), that new systems would work as if they were on the world wide web. This does not recognize the profound differences between the commercial and military situations. In the military, environment applications must span the globe using

INTRODUCTION

both satellite and terrestrial links, use extensive cryptography and, ultimately, be obliged to work with users who have comparatively limited bandwidth.

Third, developers are encouraged to focus on small modules of usable software that can be developed in short and defined periods of time (time-certain development). Taken to the extreme, time-certain development can lead to on-time delivery of software that fails to meet user needs and defers addressing the most difficult problems. Such an outcome was seen in the development of the five pilot NECC capability modules.

The path to success for these software intensive systems is remarkably similar to that of complex hardware systems: ensuring clearly articulated requirements by collaboration between the user and developer as mentioned in goal 1; a disciplined systems engineering approach, as mentioned by the DSB; and more realistic developmental testing that reflects the actual operational environment. I am pleased that Secretary Young has emphasized all three points in his reviews of NECC and the Joint Tactical Radio System.

DOT&E FISCAL YEAR 2008 OVERSIGHT AND REPORTING ACTIVITY

During this year, my office monitored 322 Major Defense Acquisition Programs (MDAPs) and special interest programs. I approved 68 Test and Evaluation Master Plans and Test and Evaluation Strategies, two LFT&E Strategies included in the Test and Evaluation Master Plans, and 86 Operational Test and Evaluation Plans for specific test events.

DOT&E delivered eight BLRIPs and one Live Fire Report to the Secretary of Defense and Congress:

Submission Date	Program Name
October 26, 2007	T-AKE <i>Lewis & Clark</i> Class of Auxiliary Dry Cargo Ships
November 1, 2007	Air Force Mission Planning System (MPS) Program Increment II (F-15)
February 1, 2008	Mk 48 Mod 7 Common Broadband Advanced Sonar System (CBASS) Phase I Torpedo
February 14, 2008	Stryker Mobile Gun System (MGS)
March 20, 2008	High Mobility Artillery Rocket System (HIMARS) with the Improved Crew Protection (ICP) Cab*
April 11, 2008	Low Band Transmitter (LBT) System
May 15, 2008	SSGN <i>Ohio</i> Class Conversion
August 22, 2008	Joint Chemical Agent Detector (JCAD)
September 15, 2008	USMC UH-1 Upgrades (UH-1Y)

(* Live Fire Testing)

DOT&E also delivered three Early Fielding Reports under the requirements of NDAA for FY07, Section 231:

Submission Date	Program Name
October 26, 2007	XM982 Excalibur Precision Engagement Projectile
April 2, 2008	SSN 774 <i>Virginia</i> Class Submarine
May 14, 2008	<i>San Antonio</i> Class Amphibious Transport Dock (LPD-17)

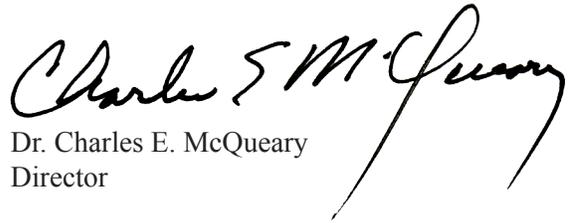
In addition to this Annual Report, we testified at four sessions of congressional meetings, provided a separate report on the Missile Defense Agency in February 2008, and responded to over 40 requests for briefings to congressional staff members.

INTRODUCTION

CONCLUSION

I am proud of the significant progress made in each of the DOT&E goals as discussed above and I greatly appreciate the support we have had from the Under Secretary of Defense for Acquisition, Technology and Logistics. I am also aware that the work of continuous process improvement is never finished. Continuity of purpose and sustained emphasis is essential to institutionalizing the lasting change needed to equip our forces with systems that work when needed.

It has been an honor and a privilege for me during these last two and a half years to have been part of an organization that is “key to weapons that work.” With that in mind, I am pleased to present the 2008 Annual Report that follows.



Dr. Charles E. McQueary
Director