INTRODUCTION

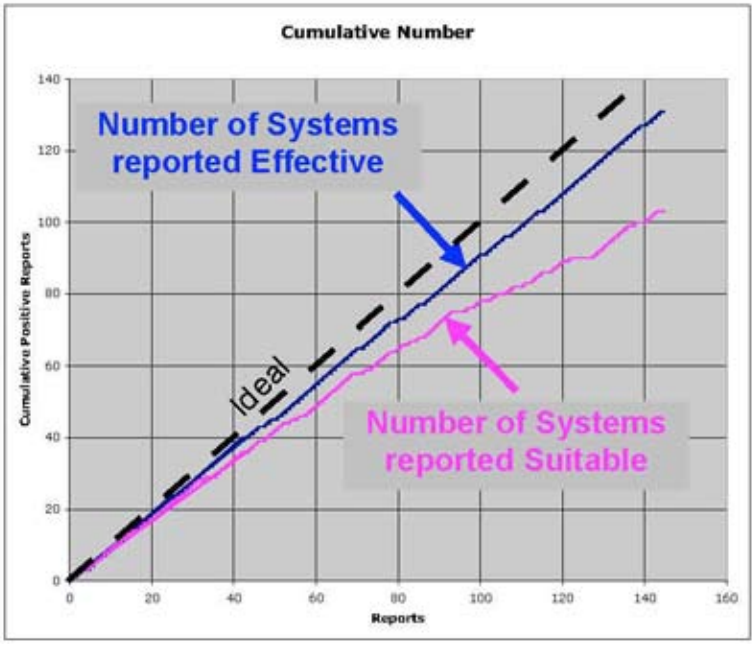
During this year, my office was privileged to monitor 288 Major Defense Acquisition Programs (MDAPs) and special interest programs. I approved 61 Test and Evaluation Master Plans and Test and Evaluation Strategies, one LFT&E Strategy, and 66 Operational Test and Evaluation Plans for specific test events. To the Secretary of Defense and the Congress, DOT&E delivered eight Beyond Low-Rate Initial Production / Live Fire Reports and three Early Fielding Reports, provided a separate report on Missile Defense, and testified at two sessions of congressional meetings.

Since my first report to you last year, I have continued to evolve and strengthen the goals and priorities I wrote about including the internal manpower study. In addition, I addressed emerging test requirements in force protection and net-centric warfare, and the report on policies and practices of test and evaluation directed in the National Defense Authorization Act (NDAA) of 2007, Section 231. The Conference Report to Accompany H.R. 1585 NDAA for Fiscal Year 2008 directs that the results of the manpower assessment be included in this Annual Report. It is included here and in Annex A.

I view these endeavors as key strategic activities and the following discussion of each of them will provide insight into the direction I have set for this organization.

GOALS IN PRIORITY ORDER

1. Improve Suitability. I continue to believe operational test and evaluation should confirm performance, rather than reveal new failure-modes. During 2007, DOT&E worked with key stakeholders, including industry, to enhance the failure-mode discovery process and eliminate surprises in operational testing. I have concluded that the key issue is inadequate system reliability, which is in turn a key component of suitability. Contributors to reliability problems include: poor definition of reliability requirements, a lack of understanding by the developer on how users will operate and maintain the system when fielded, lack of reliability incentives in contracting, and poor tracking of reliability growth during system development. While we have made progress in identifying the systemic problems causing poor suitability, program performance has not shown improvement. Of the eight Beyond Low-Rate Initial Production (BLRIP) reports published last year, most systems (88 percent) were operationally effective; but half (50 percent) were not suitable. To further put this in context, DOT&E has sent a total of 144 system reports to Congress since 1983 and we assessed 103 of the systems as suitable (72 percent). This past year's result of 50 percent reveals a continued downward trend as depicted in the chart.



The trend raises two concerns. First, system suitability – especially reliability – directly impacts our warfighter's performance. The DoD needs systems that are effective when needed, not just effective when available. Second, suitability – especially reliability – drives system life cycle costs. Put simply, poor reliability means higher sustainment cost.

As part of our efforts, we analyzed several programs that show a clear linkage between investment to improve reliability and reduced life cycle costs. Our analysis revealed reliability returns-on-investment

between a low of 2 to 1 and a high of 128 to 1. The average expected return is 15 to 1, implying a $15 savings in life cycle costs for each dollar invested in reliability. Since the programs we examined were mature, I believe that earlier reliability investment (ideally, early in the design process), could yield even larger returns with benefits to both warfighters and taxpayers.

I am convinced that each step in the development process can be used to improve suitability: the statement of requirements, the contract with industry, the design phase, early discovery of failure modes in developmental and operational test (DT/OT), and the collection of field data. While DOT&E is engaged in the operational testing of systems, we are also teaming with Departmental and industry partners to forge improvements in earlier steps.

Key stakeholders also agree with DOT&E that reliability (and its associated elements availability, and maintainability, together known as RAM) is a primary contributor to sustainment costs. In May 2007, the Joint Staff took a key first step by publishing an updated instruction about system requirements (CJCSI 3170.01F). The Joint Staff declared "Materiel Availability" a mandatory Key Performance Parameter (KPP) with two Key System Attributes (KSA): "Materiel Reliability," and "Ownership Cost." These KPP/KSAs make reliability a required metric for system evaluation, and show Departmental consensus that reliability is linked to: 1) System availability – ensuring warfighters have the system they need when they need it and, 2) Reducing total ownership cost.

Along with mandatory metrics, a clear picture must exist as to how the operational concept, the maintenance concept (how users will operate and maintain the system when fielded), and the metrics are intended to align. That picture provides the context for system design tradeoffs, and the conditions for test and evaluation. We call this picture the "Sustainment Requirements Development Report," and it is being developed to explain the rationale for a system's Materiel Availability KPP. The Joint Staff, the office of the Under Secretary of Defense for Acquisition, Technology, and Logistics, and DOT&E are collaborating on a handbook on how to develop the metrics and produce the report.

I also believe an effort to define best practices for reliability programs is vital and that these should play a larger role in both the guidance for, and the evaluation of, program proposals. Once agreed upon and codified, reliability program standards could logically appear in both Requests for Proposals (RFPs) and, as appropriate, in contracts. Industry's role is key in this area. Through an initiative with government and industry, the Government Electronics and Information Technology Association (GEIA) is working to define these best practices. A related project is ongoing work by the National Defense Industrial Association's (NDIA) Systems Engineering committee. The latter group is examining the impact of current defense acquisition policy and guidance on system suitability. The NDIA plans to deliver recommendations for acquisition policy which should help industry improve suitability. In sum, I see industry's increased commitment to address reliability and suitability as evidence of growing momentum for improvement.

As a final note, DOT&E understands the key role that operational testing will have in informing decision-makers about the new sustainment KPP. We are working with the Service Operational Test Agencies on how best to gather test data which could provide information about ownership cost. The objective is to use realistic data, from T&E, to support estimates of ownership cost.

In response to DOT&E's top priority of improving suitability, we have reached out across the DoD and to industry in a variety of initiatives. Each is important, and all – collectively – are necessary to properly set requirements, incentivize industry, oversee system design and development, and finally confirm suitability in operational T&E. The results will not be immediate, but the problem – as this year's suitability results indicate – must be addressed. If each of these initiatives is successful, over time I expect to report more operationally suitable systems.

2. Enhance operational realism in early tests, including developmental testing. During the past year, DOT&E supported a Defense Science Board (DSB) Task Force examining the need to reinvigorate

developmental test and evaluation. The interim results from that effort suggest that the frequent discovery of technical maturity problems in system operational test and evaluation, can be corrected only by re-instituting a disciplined Systems Engineering process during design and development. As I stated earlier, operational test and evaluation should be a mechanism to confirm performance, rather than one to discover new failure-modes. The DSB suggested, as many others have, that integrating developmental and operational testing could help. Among the suggestions made in the DSB interim report were the following:

- Change OSD and Service policy to mandate integrated DT/OT evaluation planning which defines testing required for all system-level evaluation.
- Enable access to all system-level test data by government DT and OT organizations as well as the prime contractor. (Separate evaluations can be accomplished by prime contractor and government test entities.)
- Give special attention to incorporating test events, where practical, designed to satisfy OT as well as DT requirements.
- Define which testing will be accomplished by the prime contractor, government DT lead, and OT as the lead agency.
- Integrate Operational Test Agencies into the deficiency report process, to include participation on Joint Reliability Maintainability Evaluation Team (JRMET) or Corrective Action Review Board throughout DT.
- Require periodic RAM assessments throughout DT to ensure early identification of problems.

Implementation of the above suggestions will create more realistic and operationally representative conditions in early testing, especially DT.

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 FY07 DOT&E delivered three such reports in compliance with this requirement.

These assessments are provided to decision-makers to help them make informed fielding decisions when systems are fielded for operational use prior to the full-rate production decision. It also helps make joint warfighters and commanders aware of system capabilities and limitations for systems that are fielded early. I believe that early fielding does not remove my responsibility to determine whether a system is effective and suitable for combat before the full-rate production decision. The Early Fielding Report will be followed by the usual Beyond Low-Rate Initial Production Report when the Initial Operational Test and Evaluation (IOT&E) is complete.

The Services have also emphasized T&E responsiveness to deployed warfighters. To support the urgent need to defeat Improvised Explosive Devices (IEDs), the Services conduct rapid testing to provide information on capabilities and limitations of systems issued directly to our warfighters to defeat IEDs. To support the Joint IED Defeat Organization, testers use flexible, streamlined, and tailored test procedures based on standard test protocols such as reusing knowledge and data, sharing data among Services and agencies, and providing concise and timely reports to support decisions. Likewise, the Air Force and Navy provide rapid evaluations of components for urgently needed capabilities such as Integrated Base Defense Security, Global Hawk, and Small Diameter Bomb employment, and Counter-Bomb detection and mitigation systems. When, in order to get equipment to those in harms way, the testing is inadequate, follow-on testing will be required.

4. Examine operational testing resources.

A. Results of Assessment of Sufficiency of Test and Evaluation Personnel. One of the key resources that DOT&E examined this year was very close to home. Title 10 specifies that "the Director shall have sufficient professional staff of military and civilian personnel to enable the Director to carry out the duties and responsibilities of the Director prescribed by law." Conference language directs a report to Congress. After careful examination I have determined that I do not have sufficient professional staff (military and

civilian) to adequately carry out the duties and responsibilities outlined for this office in 10 USC 139. Likewise, I am limited in my ability to properly support the acquisition process and to respond quickly to Combatant Commanders' requests for support from our Joint Test and Evaluation Program. The needed capabilities are inherently governmental, i.e., cannot be done by contractors. The senior leadership of the DoD is now reviewing our September request for additional staff.

I attribute the staff shortfall to the 59 percent increase – since 1994 – in programs on the DOT&E oversight list (from 184 to 288), and an increase in responsibilities associated with congressional direction to DOT&E – without a commensurate increase in staff.

The number of Action Officers on our staff has remained almost constant since 1983, despite statutory changes and acquisition initiatives which have significantly increased our workload. As a result of our analysis, we have requested a small increase in our military staff and a somewhat larger increase in our civilian staff. The military billets address the need for a current operational perspective in T&E. More detail is provided in Annex A.

B. Manpower in the Service OTAs. I also maintain a perspective of resource issues in the Service OTAs. One current interest is the OTAs' capability to report certain types of data to enable OSD to evaluate the sustainability Key Performance Parameter previously mentioned. Typical T&E events often yield maintenance and repair information and usage data that will contribute to realistic estimates of system sustainment costs. We are working with OSD offices responsible for sustainment cost estimates, and with the OTAs, to contribute relevant data. As with all new missions, there is a question of resources. Depending on results of our pilot work, I may recommend additional resources for OTAs in order to support evaluation of the new Key Performance Parameter.

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 has revamped its in-house training program. The training program has four levels.

1. Orientation, within two weeks of a new DOT&E staff member arriving, that provides basic understanding of job and duties and where to get further guidance.
2. Action Officer Course offered twice a year to give in-depth instruction on performing their responsibilities.
3. Continuing Education that presents topics intended to keep all personnel abreast of policy changes, lessons learned, new initiatives, and approaches to resolve testing challenges.
4. Professional Development designed to improve the education and leadership of assigned personnel.

As part of the level three training, I sponsored a special training this year related to the DOT&E initiative in system reliability. We arranged for an acknowledged, world-class expert in system reliability to teach best practices for assuring system reliability to select DOT&E staff, OSD Acquisition (systems engineering) staff, and analysts from the Institute for Defense Analyses. This course was designed to enable the OSD staff to interface directly with program offices, as part of the oversight mission, and coach programs on the right ways to achieve reliable systems.

As part of level four (professional development), nine DOT&E staff are participating in the Deputy Secretary of Defense's Lean Six Sigma "Green Belt" training, and my Deputies and I have either taken, or are going to take, executive training in that subject.

EMERGING TESTING MISSION AREAS

1. Force Protection Equipment. Congressional language requires the Director to provide guidance to and consult with DoD officials regarding the operational test and evaluation or survivability testing of force protection equipment, including non-lethal weapons. The language does not however, provide DOT&E the authority to oversee these programs, nor influence the scope of their test and evaluation

programs. Based upon increased congressional interest in personnel body armor and combat helmets, I believe DOT&E should have traditional oversight authority over these programs both in operational testing and survivability testing. We are working with key partners (who include: Assistant Secretary of the Army (Acquisition, Logisitics, and Technology), Army Test and Evaluation Command, Army Research Laboratory, Marine Corps Systems Command, Joint Non-Lethal Weapons Directorate, Human Effects Center of Excellence, Air Force Test and Evaluation, Joint Staff J-8, Special Operations Command, U.S. Navy (Research, Development, and Acquisition), and USD(AT&L)) to develop plans for future engagement in the force protection area.

There were two notable examples of DOT&E involvement in force protection programs this year. The first example addressed congressional concern about body armor that triggered a July 2007 decision to require DOT&E to oversee body armor testing. As this report goes to press, DOT&E continues our involvement with the Army Test and Evaluation Command while they prepare to test vendors for Army body armor. In a related supporting effort, DOT&E, the Army, and the Marine Corps have been working together for some time to select and codify a new test operations procedure.

The second example concerned a July 2007 requirement to assess the ballistic protection of the military Personnel Armor System for Ground Troops helmet in response to a Department of Justice investigation. Within 10 days of the request, DOT&E responded to the Secretary of Defense that the tested helmets met the ballistic protection requirement. Details on these matters are in the Live Fire section of this report.

2. Net-Centric Systems. Another emerging challenge is the protection of our networked information systems. The success of the United States and our coalition partners in net-centric warfare has not gone unnoticed by potential adversaries. Today, we see continual probing of our networks, as well as kinetic demonstrations of potential disruptions to our space and net-centric systems. Assessment of these systems must extend beyond preventing intrusion. Additional focus is needed in detecting intrusions, reacting to attacks, and rapidly restoring essential capabilities. We have carried out the congressional mandate to assess fielded systems, as well as our responsibilities to acquisition programs. Future efforts will require an aggressive use of live, virtual, and constructive techniques in concert with the joint training community. In FY08, we will work with our partners at U.S. Joint Forces Command to help align joint testing and training roadmaps in this growing mission area.

SECTION 231 REPORT ON T&E POLICY AND PRACTICE

Section 231 of the National Defense Authorization Act for Fiscal Year 2007, Public Law 109-364, directed a review and amendment, if appropriate, of DoD policies and practices on test and evaluation. An initial report (July 17, 2007) responding to this task identified policy initiatives under active consideration with respect to both traditional and emerging acquisition approaches:

- Testing and evaluation should concentrate on measuring improvements to mission capability and operational support based on user needs;
- Testing and evaluation programs should experiment in the sense that they should learn and understand the strengths and weaknesses of a system and its components, and the effect on operational capabilities and limitations. Decision-makers (e.g., managers, engineers, and users) can then incorporate test results into corrective actions or system enhancement initiatives;
- Developmental and operational testing should be integrated and continual to the maximum extent feasible;
- Testing and evaluation should begin early, be more operationally realistic, and continue through the entire system life cycle;
- Evaluation should be conducted in the mission context expected at time of fielding to the user and beyond and should be expressed in terms of the operational significance of the test results;
- Evaluations should include a comparison against current mission capabilities so that measurable improvements can be determined;

- Evaluations should take into account all available data and information;
- Test and evaluation should exploit the benefits of appropriate models and simulations.

Many of these items are similar to the previously discussed Defense Science Board Task Force recommendations. Changes to T&E policy were submitted to the Congress in December. The institutionalization of these policies will require changes to the Department's Acquisition Directives and Regulations.

DOT&E OVERSIGHT ACTIVITY FOR FISCAL YEAR 2007

During this year, my office monitored 288 Major Defense Acquisition Programs (MDAPs) and special interest programs. I approved 61 Test and Evaluation Master Plans and Test and Evaluation Strategies, one LFT&E Strategy, and 66 Operational Test and Evaluation Plans for specific test events.

DOT&E delivered eight Beyond Low-Rate Initial Production Reports to the Secretary of Defense and the Congress:

1. Small Diameter Bomb (SDB) - October 6, 2006
2. Global Broadcast Service (GBS) System - December 4, 2006
3. Small Unmanned Aerial System (SUAS) - December 13, 2006
4. APG-79 Active Electronically Scanned Array (AESA) Radar - April 25, 2007 (classified)
5. UH-60M Black Hawk Utility Helicopter - May 17, 2007
6. Common Submarine Radio Room (CSRR) - June 29, 2007
7. CH-47F Block II Improved Helicopter - June 29, 2007
8. UH-72A Lakota Light Utility Helicopter (LUH) - July 27, 2007

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

1. Land Warrior System - April 17, 2007
2. Stryker Mobile Gun System (MGS) - June 4, 2007 (classified)
3. Common Broadband Advanced Sonar System (CBASS) Torpedo - June 7, 2007 (classified)

Finally, in addition to the Missile Defense Agency section of this Annual Report, we provided a separate classified February 15, 2007, report and testified at two sessions of congressional meetings.

It continues to be an honor and a privilege for me to be part of an organization that is the "key to weapons that work." With that in mind, I am pleased to present the 2007 Annual Report that follows.

Dr. Charles E. McQueary
Director