FY 2011
Annual Report

In my report to you last year, I discussed four initiatives that I am undertaking as Director, Operational Test and Evaluation: field new capability rapidly; engage early to improve requirements; integrate developmental, live fire, and operational testing; and substantially improve suitability before initial operational test and evaluation (IOT&E). In this Introduction, I report on the progress made implementing these initiatives, discussing several success stories as well as areas requiring additional effort. I will first discuss key issues causing program delays in defense acquisition and the marginal cost of operational testing. I will also include a discussion of operational test and evaluation (OT&E) interest areas, as well as a summary of my monitoring and reporting activities on OT&E.

Additionally, I have included a new discussion in the Activity and Oversight chapter of this annual report containing my assessment of significant issues observed in operational testing of systems under my oversight in 2010-2011. These issues, in my view, should have been discovered and resolved prior to the commencement of operational testing. This new section also provides my identification of significant issues observed in early testing of systems during 2010-2011 that, if not corrected, could adversely affect my assessment of those systems' effectiveness, suitability, and survivability during IOT&E.

PROGRAM DELAYS

In response to continuing comments by the acquisition community that testing drives undue requirements, excessive cost, and added schedule into programs, I conducted a systematic review of recent major acquisition programs that experienced delays. I examined these programs to determine the causes and lengths of program delays, and the marginal cost of operational test and evaluation. The Under Secretary of Defense (Acquisition, Technology, and Logistics) (USD(AT&L)) had also chartered a team to assess the acquisition community's concerns regarding testing. The results of both studies indicated that testing and test requirements do not cause major program delays or drive undue costs. Dr. Carter and I signed a joint memorandum addressing these issues as well as other problems that were identified in the two studies, summarized below.

The USD(AT&L) study team found that tensions are often evident between programs and the test community and for the most part these are normal and healthy; however, there is room for improvement in these relationships and interactions. Four potential mitigations were identified:

- Stronger mechanisms for a more rapid adaptation to emerging facts
- A requirements process that produces well-defined and testable requirements
- Alignment of acquisition and test strategies (i.e., programs lack the budgetary and contract flexibility necessary to accommodate discovery)
- Open communications between programs and testers, early and often, with constructive involvement of senior leaders

Causes of program delays

My review examined 67 major programs that experienced significant delays and/or a Nunn-McCurdy breach. Thirty-six of these programs experienced a Nunn-McCurdy breach and six of these programs were ultimately canceled. (Two of the 36 Nunn-McCurdy programs experienced no delays to their schedule.) We identified five categories of problems that resulted in delays:

- Manufacturing and development (to include quality control, software development, and integration issues)
- Programmatic (scheduling or funding problems)
- Poor performance in developmental testing (DT)
- Poor performance in operational testing (OT)
- Difficulties conducting the test (such as range availability, test instrumentation problems, and other test execution problems)

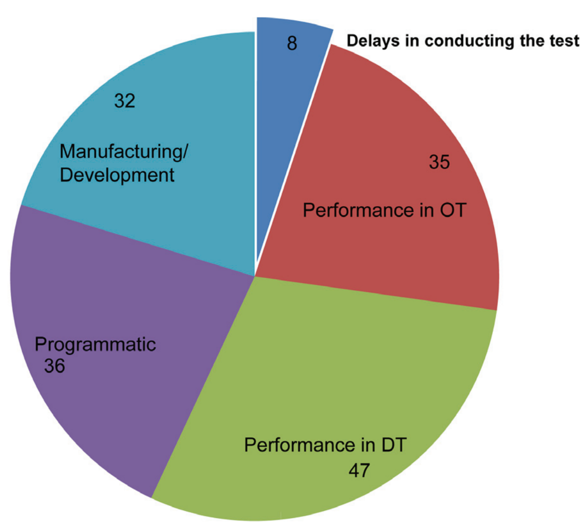

FIGURE 1.
REASONS BEHIND PROGRAM DELAYS

Of the 67 programs, we found that 56 programs (or 84 percent) had performance problems in testing (either DT, OT, or both) while only eight programs (or 12 percent) had issues conducting the tests that led to delays. Only one program had delays solely attributed to the test: the Army's Force XXI Battle Command Brigade and Below (FBCB2) operational test was delayed for one year because the test unit designated by the Army was deployed. However, the delay of the IOT&E for the FBCB2 did not affect the deployment of the satellite communications version of the system. The IOT&E was conducted later on the original terrestrial communications system, which had been previously shown to have poor performance in early operational tests. Figure 1 shows the distribution of the types of delays that occurred in the 67 programs evaluated. There were 158 instances of delays for the 67 programs in five categories (many of the programs had more than one reason for delays). Clearly, programs are most often delayed because of the results of testing, not the testing itself.

Length of delays

The length of delays for the programs examined varied from none (for two of the Nunn-McCurdy programs) to 15 years. Thirty-seven programs were delayed greater than 3 years. The delays were measured against the most recent previously published schedule; so, in a sense the total delay experienced is likely to be even longer relative to the original planned schedule. Six of the programs were eventually cancelled, and one had its Milestone B approval rescinded.

Cost of Operational Testing

The DOT&E and USD(AT&L) studies noted that the marginal cost of operational testing is a small portion of a programs' overall budget; however, the costs can be a large percentage of the budget in the year(s) in which testing occurs. Because the operational testing occurs at the end of the development process, programs typically have fewer degrees of freedom (and resources) left to resolve problems conducting such tests or correcting the problems they too often reveal. Therefore, it is important for planning for OT to commence early, so that the necessary resources can be allocated at the programs' outset.

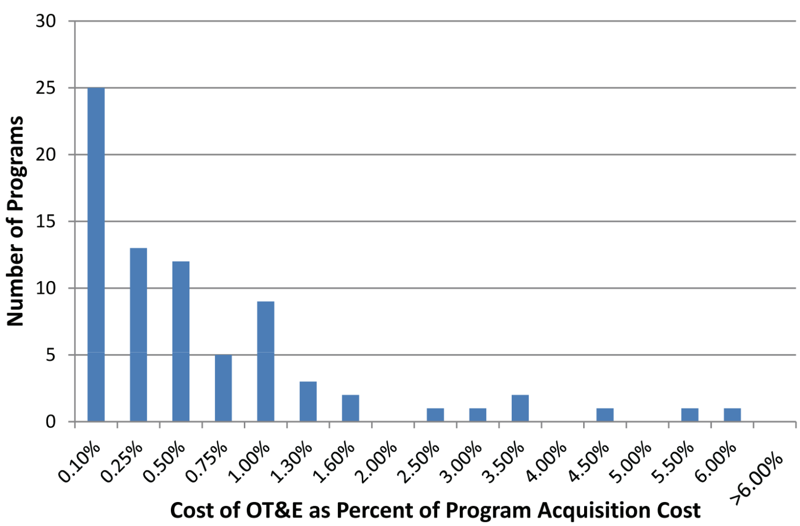

FIGURE 2.
MARGINAL COST OF OT&E RELATIVE TO PROGRAM ACQUISITION COST

We evaluated marginal cost of operational test and evaluation to programs as a percentage of total acquisition cost. A review of 78 recent programs in the Army, Air Force, and Navy showed that the average marginal cost of OT&E is approximately 0.65 percent of the total acquisition cost. Few programs that we reviewed (7 out of 78) required more than 1.5 percent of program acquisition costs for OT&E. For those programs with above average OT&E costs, a relatively low program acquisition cost was the dominant cause of larger proportional OT&E cost (e.g., Modular Aircrew Helmet with $8.4 Million acquisition costs, AIM-120C Electronic Protection Improvement Program with $87 Million acquisition costs, and the Hard Target Void Sensing Fuze with $147 Million acquisition costs). Expense of test articles and their expendability was another major driver. Figure 2 shows the distribution of the marginal cost of OT for the 78 programs we examined.

In addition to the DOT&E and USD(AT&L) studies, the Decker-Wagner report commissioned last year by the Secretary of the Army, addressed the Army's failure rate of initiating and then cancelling new development programs. The study found that between 1990 and 2010, the Army terminated 22 Major Defense Acquisition Programs (MDAPs), and that 15 of those terminations occurred since 2001. Further, excluding the Future Combat System (FCS), the Army spent more than $1 Billion per year since 1996 on programs that were eventually cancelled before completion. The study cited many reasons for the failed programs including unconstrained requirements, weak trade studies, and erosion of the requirements and acquisition workforce. However, none of the reasons cited included test and evaluation (T&E). In fact, in my opinion, earlier and more robust T&E would have revealed problems and solutions earlier when they would have been less costly to fix or allowed decision makers to cancel or restructure programs and avoid wasting billions of dollars.

PROGRESS ON DOT&E INITIATIVES

1. Field new capability rapidly.

Providing new and better equipment to our fighting forces as quickly as possible remains a top priority for the Department. Each Service operational test agency has developed methods for rapidly evaluating systems fulfilling urgent operational needs, including combining testing with the training of the first unit to be equipped and conducting quick reaction assessments. Examples of rapid acquisition programs that underwent tailored, yet rigorous live fire and operational testing this year include upgrades to the Mine Resistant Ambush Protected family of vehicles, the Stryker Double-V Hull, the MQ-9 Reaper Unmanned Aircraft System, the MQ-1C Gray Eagle Unmanned Aircraft System, the Mk 54 and Mk 48 torpedoes, and the Enhanced AN/TPQ-36 (EQ-36) Radar System.

One consequence of rapid fielding is that systems can be committed to combat operations before IOT&E and full-rate production. Under that circumstance, Congress requires DOT&E to submit Early Fielding Reports. In FY11, DOT&E delivered four such reports: the MQ-8B Vertical Take-off and Landing Unmanned Aerial Vehicle (Fire Scout), Navy Multiband Terminal, Precision Lethality Mk 82 Bomb, and the Mk 48 Torpedo. These Early Fielding Reports were also provided to the Services to support their fielding decisions and to the combatant commanders to make our joint forces aware of the capability these systems do and do not provide.

The Joint Test and Evaluation (JT&E) program, established in 1972, continues to provide rapid, non-material solutions to operational problems identified by the joint military community. DOT&E manages the JT&E program and executes it in partnership with the Combatant Commanders. Products of the program include improved tactics, techniques, and procedures (TTPs), revised operational architectures, and training packages. In addition to seven joint tests in FY11, the JT&E program conducted 14 quick reaction tests and one special project. A detailed discussion of these activities is provided in the JT&E chapter of this report.

2. Engage early to improve requirements.

The Weapon System Acquisition Reform Act of 2009 recognized that "unrealistic performance expectations" and "immature technologies" are among the root causes of trouble in defense programs. In addition, the USD(AT&L) study referenced above concluded that "the requirements process is broken;" that requirements are not well conceived; too many Key Performance Parameters complicated consideration of cost-design tradeoffs; the requirements process is slow, cumbersome, and not flexible enough to change with discovery; and finally, that it suffers from inadequate definition of anticipated operational environments and associated mission-oriented operational test requirements. The Decker-Wagner study referenced above also cited unconstrained requirements, weak trade studies, and erosion of the workforce as causes for many of the Army's failed acquisition programs. To this end, DOT&E has four dedicated staff members working within the Department's requirements-setting process, the Joint Capabilities Integration Development System (JCIDS). This year they participated on a Joint Staff task force to re-structure the JCIDS process, recommending implementation of a number of changes. One example of needed change is that the rationale for requirements be rigorously stated and justified on the basis of accomplishing missions in combat rather than on the basis of technical specifications. Throughout our participation in the task force, we have consistently emphasized the need to have measurable and testable requirements related clearly to mission accomplishment that will allow the test and evaluation community to provide timely and relevant information to decision makers. DOT&E now provides advice on the testability and relevance of proposed requirements through participation on Joint Capabilities Boards and is formally designated as an advisor to the Joint Requirements Oversight Council.

3. Integrate developmental, live fire, and operational testing.

Dedicated operational testing will always be required to provide relevant, credible evidence with inferential weight to decision makers about the operational benefits of acquiring a new weapon system. That testing must be rigorous and efficient; thus, to the extent possible, it should be integrated with developmental and live-fire testing in order to make early and full use of all the data available.

The National Research Council Panel on Statistical Methods for T&E of Defense Systems has stated many times since 1998 that current practices in the Department do not take full advantage of the benefits available from the use of state-of-the-art statistical methodology, including experimental design techniques. Thus, our testing is not as efficient as it should be. To remedy this shortcoming, my office is working with other stakeholders to develop a roadmap institutionalizing the use of scientific design and rigor in test and evaluation. The stakeholders involved include Deputy Assistant Secretary of Defense (Developmental Test and Evaluation) (DASD(DT&E)), the Service operational test agencies, and the Service T&E executives. The roadmap being developed is a multi-prong, phased effort that encompasses guidance and policy; education, training, and software tools; case studies highlighting lessons learned; and pilot projects applying scientific design tools to solve current problems in constructing tests and evaluating their results. This year we also completed a study of the military and civilian workforce demographics of the operational test agencies and the personnel within the acquisition workforce designated as T&E career field in order to learn what types of training and education are necessary for the workforce to implement these advanced experimental design and analysis techniques. The first phase of the roadmap comprises case studies drawn from each of the Service operational test agencies providing examples and best practices for experimental designs. We are developing a website that displays these case studies and lessons learned from T&E using experimental design and analysis. Moreover, we have recently completed a guide to producing a rigorous and comprehensive Test and Evaluation Master Plan (TEMP). The guide has specific instructions, examples, and guidance on topics of interest that DOT&E requires in a rigorous TEMP. This guide will be hosted on the DOT&E website with a link from the Defense Acquisition University website.

We are now beginning the next phase of the roadmap. We are working directly with each of the Service operational test agencies to identify pilot acquisition programs in which challenging problems in test design and analysis have been discovered. For these programs, we will provide statistical and experimental design expertise to develop solutions to those challenging problems. We will solicit early involvement from the developmental test and evaluation (DT&E) community so that the experimental design can encompass all phases of testing, thereby providing rigorous and efficient integrated testing.

We have established research relationships with academic institutions that have excelled in the fields of statistics, experimental design, T&E, and systems engineering. Together with the Test Resource Management Center, we are funding a three-year research consortium comprising the Air Force Institute of Technology, Naval Postgraduate School, Arizona State University, and Virginia Tech. Graduate research is being conducted on difficult statistical problems in T&E. The research also provides published work on the use of design of experiments in DoD T&E and an academic pool of analysts familiar with T&E to augment the current workforce.

A recent example of using statistical methodology and combining operational and live fire testing is with the Enhanced Combat Helmet (ECH) test program. The Marine Corps will conduct a Full-Up System Level (FUSL) test of the ECH in FY12. The purpose of this test is to provide data to support an evaluation of the vulnerability of the ECH to realistic artillery shell fragments. The test incorporates an array of ECH test articles in an operational layout to evaluate their protection in a realistic environment. The Marine Corps will use the results of this test, along with other ballistic and non-ballistic test results, to inform its procurement decisions. The ECH FUSL will provide valuable lessons for conducting future FUSL tests of combat helmets.

4. Substantially improve suitability before IOT&E.

Our fighting forces need systems that are effective when needed, not just effective when available. Weapon system reliability is a key factor in suitability; it is a primary driver of the operations and support costs of the system; and poor reliability burdens on our combat forces with unscheduled maintenance, excessive logistics burden, and down time. The timeline in Figure 3 below shows the steps the Department has taken to improve reliability, starting with a DOT&E initiative in 2006 to improve the suitability of fielded systems. Following our initiative, the Joint Staff issued a Directive making material availability a key performance parameter, the Army Acquisition Executive issued policy requiring early incorporation of Reliability, Availability, and Maintainability (RAM) planning into their regulations and required new programs to include reliability growth and testing, and we re-instated the Reliability Improvement Working Group. The Defense Science Board on Developmental Test and Evaluation in 2008 found that the use of reliability growth in development had been discontinued by the Department over 15 years ago, and that the solution to the resulting trend of producing unreliable systems would be to ensure that programs are formulated and funded to execute a viable systems engineering strategy that includes a robust reliability growth plan from inception.

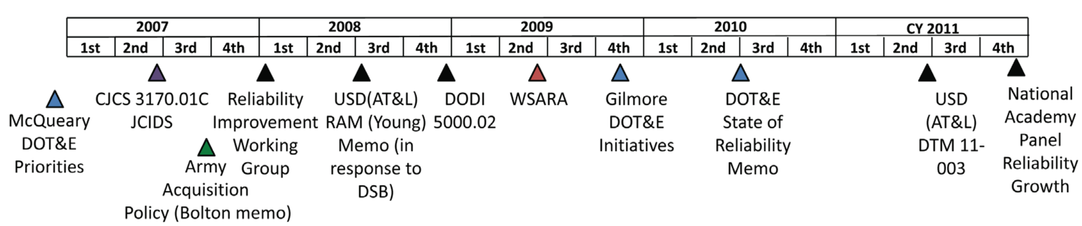

FIGURE 3.
TIMELINE OF EVENTS WITHIN DOD TO IMPROVE RELIABILITY

A system can be reliable but not suitable because of safety, human factors, training, or a combination of other factors. Conversely, a system could be unreliable but still be suitable because failures were easily repaired, there was redundancy in the system, or the reliability requirement was excessive. Figure 4 shows the cumulative scores for effectiveness, suitability, and reliability for systems on which we reported to Congress from 2006 to 2011 (a total of 52 reports). I scored each of the 52 reports to Congress as "reliable" or "not reliable" based on whether they met their reliability threshold; 36 out of 52 systems were found to be suitable while only 26 out of the 52 systems met their reliability threshold. Notably, none of these 52 systems were ultimately cancelled.

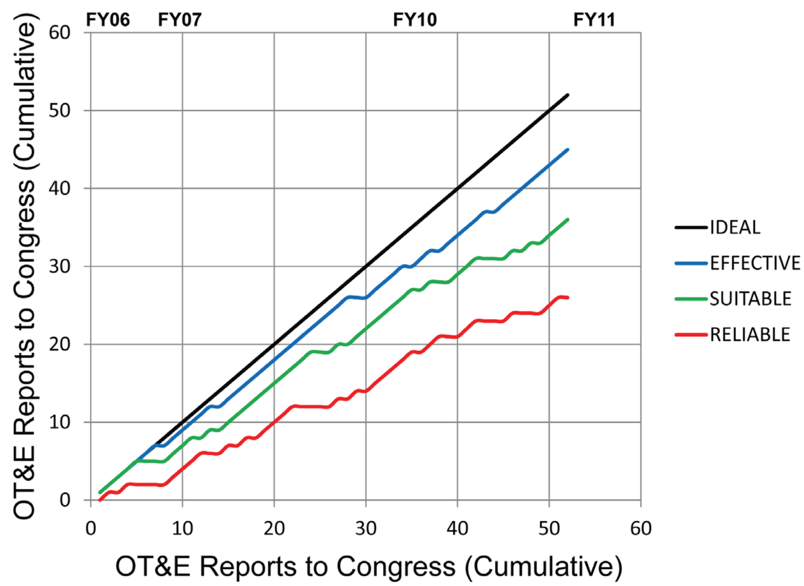

FIGURE 4.
CURRENT TRENDS IN RELIABILITY

Reliability Program Standard

In 1998, the DoD cancelled Mil-Std-785B, "Reliability Program for Systems and Equipment Development and Production." This standard was originally written in 1969 and last updated in 1980; however, industry continues to follow the -785B methodology, which, unfortunately, takes a more reactive than proactive approach to achieving reliability goals. In this standard, approximately 30 percent of the system reliability comes from the design while the remaining 70 percent is to be achieved through growth implemented during test phases. In 2008, the Defense Science Board stated that the DoD needed a standard that defense contractors can use to prepare proposals. A new voluntary reliability standard was developed by subject matter experts drawn from industry, DoD, academia, and the Services; the ANSI/GEIA-STD-0009. This standard was designated by the Deputy Assistant Secretary of Defense (Systems Engineering) (DASD(SE)) the Reliability Program Standard for Systems Design, Development, and Manufacturing to make it easy for program managers to incorporate the best practices in requests for proposals and contracts. The standard promotes four objectives:

- Understand customer/user requirements and constraints
- Design for Reliability (DfR) and re-design for reliability
- Produce reliable systems
- Monitor and assess user's experienced reliability

Thus, the standard emphasizes the need to design reliability into a system at the component level from the outset, rather than test for reliability after components have been designed and integrated to determine if retro-fixes are needed. Specific programs that have used the Design for Reliability standard include the Small Diameter Bomb II, the Stryker Nuclear, Biological, and Chemical Reconnaissance Vehicle, and the Ground Combat Vehicle. Those systems' contractors, Raytheon Missile Systems and General Dynamics Land Systems, were both active participants in the development of the ANSI/GEIA-STD-0009 (along with numerous other contractor participants).

Reliability Growth in TEMPs

We conducted a survey of 151 programs with approved Test and Evaluation Master Plans (TEMPs). Of those 151 programs, 90 percent of programs with TEMPS approved since 2008 plan to collect and report reliability data. A comparison of programs that completed a TEMP before and after June 2008 (when the Office of the Secretary of Defense (OSD) began initiatives to improve reliability) indicates improvement in several areas. Since 2008, programs are more likely to:

- Have an approved System Engineering Plan
- Incorporate reliability as an element of test strategy
- Document reliability growth strategy in the TEMP and include reliability growth curves in TEMPs
- Establish reliability-based milestone or OT entrance criteria
- Collect and report reliability data

However, as shown in Figure 4 above, no significant improvement has yet been demonstrated indicating systems are meeting their reliability thresholds; moreover:

- There is no evidence of programs using reliability metrics to ensure growth is on track.
- Systems continue to enter OT without demonstrating required reliability.
- 50 percent of programs with time scheduled to implement corrective actions met reliability thresholds compared to only 22 percent for programs without corrective action periods.

Significant Actions since 2008 Defense Science Board Study on Developmental Test and Evaluation

In response to both the Defense Science Board study (2008) and the Weapons System Acquisition Reform Act (2009), the Department took a number of actions within OSD, which in turn provided impetus for the Services to take action. In particular, within OSD,

The Systems Engineering Forum was established with DOT&E and USD(AT&L) Systems Engineering and the Service System Engineering Executives. The forum includes monthly updates from each Service on reliability improvement action items.

DOT&E has sponsored Reliability Growth Training conducted most recently by the Army Evaluation Command (AEC) and Army Materiel Systems Analysis Activity (AMSAA)). The courses offer multiple venues throughout the year and are attended by DOT&E staff, DASD(DT&E) staff, and Service personnel with responsibilities for reliability and maintainability, as well as test and evaluation.

The Reliability Senior Steering Group was established in response to the DOT&E letter to USD(AT&L) in late 2009 concerning continued poor reliability performance during initial operational testing. Senior DoD Leaders and Service Acquisition Executives compose three working groups. The primary product of this effort was the Directive Type Memorandum (DTM 11-03) on Reliability Analysis, Planning, Tracking, and Reporting, which was signed by USD(AT&L) in March 2011.

The Deputy Assistant Secretary Defense (System Engineering) now has a dedicated position for Reliability and Maintainability Engineering. The incumbent provides recommendations and advice, and chairs the Service Reliability and Maintainability Engineering Leads quarterly working group.

Specific Service Actions on Reliability Growth

Figure 5 below shows the fraction of systems meeting reliability thresholds for programs on DOT&E oversight between 2006 and 2011 (the same programs depicted in Figure 4 now broken out by Service.)

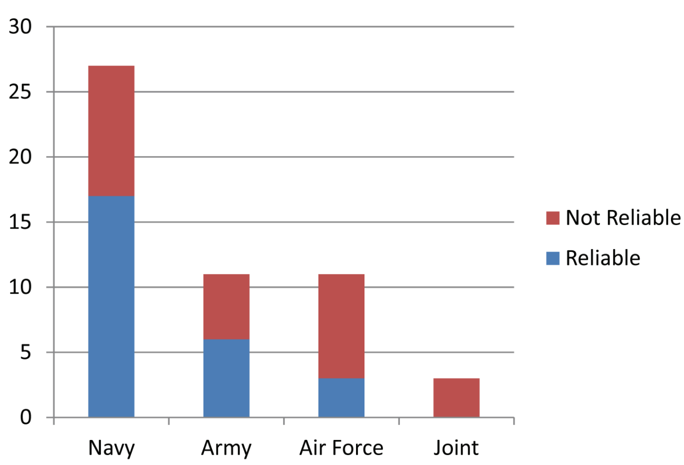

FIGURE 5.
FRACTION OF PROGRAMS MEETING RELIABILITY THRESHOLDS AT IOT&E, BY SERVICE (FROM DOT&E REPORTS TO CONGRESS 2006 – 2011)

Army. The Army Acquisition Executive issued specific policy including: that a Reliability Growth Planning Curve will be included in Engineering and Manufacturing Development (EMD) contracts; that new development programs are to execute Design for Reliability before Milestone B; and that an early reliability test threshold must be established for EMD. Additionally the Army established their Center for Reliability growth with the AEC and AMSAA, which provides training for the Army, OSD, and other Services.

As shown in Figure 5, 55 percent (6/11) of the Army programs that I reported on met their reliability thresholds. The aviation (CH-47 and UH-72) and trucks and artillery (GMLRS) performed well while networks and unmanned systems did not do well.

Navy. The Navy established a Director, Reliability and Maintainability Engineering position within the Deputy Assistant Secretary of the Navy (Research, Development, Test and Evaluation) (DASN(RDT&E)) and reliability and maintainability working groups were established at the Department of the Navy (DoN) and System Command levels. It established a network-based Integrated Reliability Software Suite for all use throughout the Service. Additionally, the Naval Air Systems Command (NAVAIR) Reliability and Engineering organization, which comprises over 200 engineers and technicians, has not been downsized during the last 15 years. The other Navy System Commands: Naval Sea Systems Command (NAVSEA), Space and Naval Warfare Systems Command (SPAWAR), and Marine Corps Systems Command (MARCORSYSCOM) are rebuilding their competencies in reliability and maintainability.

As shown in Figure 5, 63 percent (17/27) of the Navy systems that I reported on met their reliability thresholds. The majority of the reliable systems were aircraft or aircraft-related systems developed in NAVAIR, such as the H-1 upgrades to the AH-1W and UH-1N helicopters, as well as the MH-60R and MH-60S helicopters. Other reliable systems were submarines and related systems such as the USS *Virginia*, USS *Ohio*, and the TB-34 towed array. Ships and software-intensive systems were the types of systems that did not meet reliability thresholds, such as LPD-17, T-AKE *Lewis and Clark* Class of Auxiliary Dry Cargo Ships, and the APG-79 Active Electronically Scanned Array (AESA) radar and the Multi-functional Information Distribution System (MIDS) Joint Tactical Radio System (JTRS) for the F/A-18E/F Super Hornet aircraft. While the last two systems are aircraft related, their software intensive nature was problematic for reliability.

Air Force. The Air Force Material Command is sponsoring short courses in reliability at multiple venues for their acquisition and test personnel. Air Force instructions for System Engineering Plans and procedures now include the guidance for analysis and classification of potential failure modes. A Risk Identification, Integration, and "ilities" (R3I) guidebook has been published.

As shown in Figure 5, only 27 percent (3/11) of the Air Force systems that DOT&E reported on met their reliability threshold. The three systems that performed reliably were the B-2 Radar Modernization Program, Space Based Surveillance System, and the C-5 Reliability Improvement and Re-Engining Program. Other programs such as Small Diameter Bomb, Global Broadcast Service, Joint Mission Planning System, MQ-9 Reaper, Miniature Air-Launched Decoy, C-27J Joint Cargo Aircraft, and Global Hawk demonstrated poor reliability.

DOT&E Continuing Actions to Improve Suitability

- With USD(AT&L), DOT&E sponsored a National Academy of Sciences Workshop on Reliability Growth Methodology. The workshop met twice this year in March and September. Members from each of the Services presented to the panel their Service-specific actions taken to improve reliability as well as obstacles to the implementation of reliability growth methods in their systems. A report from the panel is expected in 2012.
- DOT&E continues to sponsor reliability growth training for its staff and all of DoD.
- DOT&E continues to provide in-house training to its staff to engage early in the system development and test planning process to ensure realistic reliability thresholds are established along with a test program that can support evaluating those thresholds.

Policy Supporting DOT&E Initiatives

Underlying my four initiatives is the need for rigorous, robust, and objective test and evaluation. Currently, I am actively engaged in updating the DODI 5000.02 "Operation of the Defense Acquisition System," to include requiring the use of state-of-the-art statistical methodologies, including experimental design and analysis techniques in TEMPs and test plans. The Defense Acquisition Guide, the T&E Management Guide, and T&E in Contracting Guide have been or are being updated to reflect emphasis on experimental design techniques and analysis. The DODI 5000.02 updates include changes that address rapid acquisition and the agile acquisition of information technology systems. I have provided a substantive description of how simultaneously rigorous and agile T&E of cyber systems can be conducted as part of the Department's Congressional report on Cyber Acquisitions. The updates to the Department policy and guidance draw in part on my FY11 guidance memoranda on the timeliness of OT&E plans, the use of production-representative test articles for IOT&E, and the use of design of experiments in OT&E.

OTHER INTEREST AREAS

Cyber Testing. In February 2011, the Chairman of the Joint Chiefs issued an Executive Order (EXORD) directing that all major exercises include realistic cyber-adversary elements as a training objective to ensure critical missions can be accomplished in cyber-contested environments. Although the EXORD focuses on assessments of fielded capabilities, this philosophy applies equally well to acquisition programs, and DOT&E is committed to ensuring that representative cyber environments are included in our combatant command and Service exercise assessments, as well as in the IOT&E of weapons programs. With these goals in mind, I met with the Deputy Secretary of Defense and proposed significant enhancements to Department cyber assessment capabilities. By the end of FY14, the Department should have in place the capabilities and processes to perform selected evaluations of offensive and defensive cyber-warfighting capabilities in representative cyber-threat environments. This will allow us to assess how well our fighting forces can defend against or fight through the most serious cyber attacks, as well as perform defensive and appropriate response. In order to apply these enhanced capabilities across all major exercises and acquisition programs, the Department will need to identify additional resources to expand the capacity and capabilities of the Red Teams who portray advanced cyber adversaries. This would include funding the cyber-ranges and modeling and simulation capabilities that provide operationally realistic environments for those activities inappropriate for live networks, as well as assessment teams to develop rigorous plans to ensure the cyber adversary is accurately portrayed, and assess the effects of representative cyber adversary activities.

Electronic Warfare Testing. The 2010 Tri-Service Electronic Warfare Test Capability Study, in which DOT&E participated, identified several critical AESA radar jamming capability upgrades needed for the facilities and open-air ranges currently used to evaluate U.S. weapon systems such as the F-35 and the Navy's Next Generation Jammer. These critical upgrades include:

- Next generation electronic warfare environment generator at indoor facilities and on open-air ranges to represent advanced high-fidelity threat emitter digital processing capabilities
- The capability to measure and characterize advanced U.S. jammers' multi-beam steering accuracy and power distribution at the target location at indoor facilities and on open-air ranges
- Next-generation threat surface-to-air-missile models and simulators for use in hardware-in-the-loop facilities and at open-air ranges
- A transportable urban threat representative communications environment that can be used to both stimulate U.S. communication jammers and evaluate jamming effectiveness on open-air ranges

OSD and the Navy are partially addressing the upgrades to indoor facilities, but additional investment will be needed to fund the open-air portions, as well as development of next-generation surface-to-air-missile threat simulators.

Network Integration Evaluation. The Army plans to conduct the Network Integration Evaluation (NIE) twice a year at Fort Bliss, Texas, and White Sands Missile Range, New Mexico, in order to provide a venue for operational testing of Army acquisition programs with a particular focus on the integrated testing of programs related to tactical communications networks supporting command and control. The exercises are also intended to provide an operationally realistic environment to evaluate new emerging capabilities that are not formal acquisition programs. The Army has established a leadership and governance triad comprising the Brigade Modernization Command, the Army Test and Evaluation Command, and the Training and Doctrine Command (TRADOC). A detailed assessment of the first NIE is provided in this report.

Helicopter Survivability Task Force. The Joint Aircraft Survivability Program (JASP), under DOT&E guidance, continued to work with the Office of the Assistant Secretary of Defense for Research and Engineering on the Helicopter Survivability Task Force (HSTF). This multi-disciplinary team is tasked with rapidly fielding techniques and technology to improve the survivability of helicopters in theater. JASP expertise in survivability technologies supported two specific vulnerability reduction technologies identified by the HSTF: passive fire protection for the V-22 and multi-hit transparent armor for MH-47G and UH-60 helicopters. Furthermore, the Joint Countermeasures Test and Evaluation Working Group (JCMT&E WG) that DOT&E co-chairs with the DASD(DT&E) continued to expand international cooperation in test and evaluation. Of note are the advances in common U.S./United Kingdom tactics development, improved understanding of hostile fire indication phenomenology, and Man-Portable Air Defense vulnerabilities through the use of the just concluded U.S./United Kingdom Aircraft Survival Equipment T&E Project Arrangement.

Combat Damage Assessment. I continued to support the Joint Combat Assessment Team (JCAT) in continuing its operation in Afghanistan with full-time deployments in Regional Commands – South, Southwest, and East. JCAT supported Iraq and other areas of the world remotely or by rapid deployment from Afghanistan or the Continental U.S. JCAT inspects damaged and destroyed aircraft, acquires maintenance records, and conducts interviews with aircrew and intelligence personnel to develop an accurate and comprehensive assessment of each aircraft combat damage event. They provide weapons, tactics, and logistics consultation to personnel and comprehensive briefings to commanders in charge of daily air operations. These efforts inform battlefield commanders, allowing them to adjust operational tactics, techniques, and procedures based on accurate threat assessments. Their efforts were instrumental in the investigation of the CH-47D (with 38 people onboard) that was shot down in Afghanistan on August 6, 2011.

Active Protection Systems. In response to FY08 legislation, DOT&E completed testing in August 2011 of seven foreign and domestic (two foreign, three domestic, and two combined foreign/domestic) active protection systems with the potential of protecting tactical vehicles. I will provide reports to Congress and acquisition leadership in 2QFY12. This effort will determine the capabilities of current active protection system technology and guide future acquisition decisions.

Personnel Protection Equipment. DOT&E continued oversight of personnel protection equipment testing. The Services and U.S. Special Operations Command are implementing the DoD testing protocol for hard body armor inserts published last year. The Defense Logistics Agency has incorporated the testing protocol into new contracts for sustainment stocks of hard armor inserts. The Army has incorporated the key concepts of statistical confidence and test design into its requirements for future protective systems it will develop. In partnership with the Services and the U.S. Special Operations Command, my staff developed a new combat helmet testing protocol. It ensures combat helmets provided to Service members meet ballistic protection requirements and provide uniform protection on the battlefield. I plan to work with the Services and the U.S. Special Operations Command to prepare a DoD-wide standard for testing of soft armor vests.

Joint Trauma Analysis and Prevention of Injury in Combat. In response to the DOT&E Mine Resistant Ambush Protected (MRAP) report of March 2010, former Secretary Gates tasked DOT&E to coordinate increasing the availability of data coming from the Joint Trauma Analysis and Prevention of Injury in Combat (JTAPIC) Program, as well as the Armed Forces Medical Examiner's Office. So far, DOT&E has hosted four Senior Reviews with participants from the JTAPIC Program Office and all of the JTAPIC partners, including Army intelligence, medical and materiel analysts, Navy medical researchers, and Marine Corps intelligence analysts. Additionally, the Army Surgeon General initiated the execution of two working-level Lean Six Sigma exercises with the goal of increasing the quality and volume of analytical outputs by improving internal operating processes. The improvements already made in these processes have increased the quality of the data shared among the partners, clarified the role of each partner as well as the JTAPIC Program Office, improved customer awareness of JTAPIC and its capabilities, and established common procedures that have streamlined data sharing and analytical processes among partners residing in various commands and Services.

Warrior Injury Assessment Manikin. In August 2010, I sponsored an Army-led, five-year research and development program to increase the Department's understanding of the cause and nature of injuries incurred in underbody blast combat events and develop appropriate instrumentation to assess such injuries in testing. This program, known as the Warrior Injury Assessment Manikin (WIAMan), utilizes expertise across multiple commands and disciplines within the Army to generate a medical research plan from which data will be transitioned to the materiel and T&E communities. These data will feed the design of a biofidelic prototype anthropomorphic test device (ATD) designed to evaluate occupant protection during large under-vehicle blast events, which have become the predominant threat to ground combat vehicles. Current test and evaluation techniques address occupant injuries using automotive crash test dummies and their associated injury criteria, all designed and developed for low-speed civilian car crashes. Development of a military-specific ATD for use in under-vehicle blast testing will better inform users, materiel developers, analysts, and evaluators about the levels of protection afforded by military vehicles to their occupants, and will enable more survivable vehicles to be fielded.

Stryker Double-V Hull. To support the deployment of Stryker Double-V Hull (DVH) vehicles to Afghanistan, the Army developed and began executing a robust multi-phase Stryker DVH Operational and Live Fire Test and Evaluation Program. Test and evaluation to date has confirmed that DVH systems significantly improve IED protection relative to the Stryker vehicles originally available to units in Afghanistan, meeting – and in some cases exceeding – MRAP All Terrain Vehicle (M-ATV) requirements. Stryker DVH additionally demonstrated in test that it retained operational characteristics required for operations in Afghanistan and provides increased vehicle reliability. The Stryker DVH test and evaluation program proved to be a success, and more survivable equipment is now available to the Soldier. The Army continues to conduct operational and live fire test and evaluation of multiple Stryker configurations modified with the DVH.

OT&E MISSION ACCOMPLISHMENTS, FISCAL YEAR 2011

During this fiscal year, my office monitored 311 Major Defense Acquisition Programs (MDAPs) and special interest programs. We approved 51 Test and Evaluation Master Plans, 6 Test and Evaluation Strategies, 79 Operational Test and Evaluation Plans, 6 Live Fire Test and Evaluation Strategies/Management Plans, and 4 Live Fire Test Plans.

Our reporting to both Congress and the Defense and Service Acquisition Executives has continued to increase over the past two years. This year, we delivered 31 reports, including our annual report on Ballistic Missile Defense Systems; in both FY10 and FY09 we delivered 14 reports. We also provided 10 Operational Assessments to the Acquisition Executives and 12 Major Automated Information System (MAIS) reports.

During FY11, DOT&E delivered 13 Beyond Low-Rate Initial Production Reports (BLRIPS) (three of which were combined OT&E and Live Fire Reports), 3 Follow-on Test and Evaluation Reports, 2 Live Fire Test and Evaluation reports, 8 special reports, and 4 Early Fielding Reports to the Secretary of Defense and Congress (see Table 1).

TABLE 1. DOT&E REPORTS TO CONGRESS DURING FISCAL YEAR 2011

PROGRAM	DATE
BEYOND LOW-RATE INITIAL PRODUCTION (BLRIP) REPORTS	
C-5 Reliability Enhancement and Re-Engining Program (RERP) (Combined OT&E/LFT&E)	October 2010
Suite of Integrated Radio Frequency Countermeasures (SIRFC)	October 2010
Excalibur Increment 1A-2 (Combined OT&E/LFT&E)	October 2010
TB-34 Next Generation Fat-Line Towed Array	November 2010
Warfighter Information Network – Tactical (WIN-T)	February 2011
Multi-functional Information Distribution System (MIDS) Joint Tactical Radio System (JTRS)	April 2011
Miniature Air-Launched Decoy (MALD)	April 2011
Improved (Chemical Agent) Point Detection System – Lifecycle Replacement (IPDS-LR)	April 2011
C-27J Joint Cargo Aircraft (JCA) (Combined OT&E/LFT&E)	May 2011
Low Cost Conformal Array (LCCA)	May 2011
RQ-4B Global Hawk Block 30	May 2011
Space-Based Space Surveillance (SBSS) Block 10	June 2011
Integrated Defensive Electronic Countermeasures (IDECM) Block 3 Electronic Countermeasures Suite	June 2011
EARLY FIELDING REPORTS	
Mk 48 Mod 6 Advanced Common Torpedo (ACOT) and Mk 48 Mod 7 Common Broadband Advanced Sonar System (CBASS) Torpedo with the Advanced Processor Build 4 (APB 4) Software	March 2011
Navy Multiband Terminal (NMT)	April 2011
MQ-8B Vertical Take-off and Landing Unmanned Aerial Vehicle (VTUAV)	June 2011
Precision Lethality Mark 82 (PL Mk 82) Bomb	September 2011
SPECIAL REPORTS	
M855A1 Lead-Free, 5.56 mm Cartridge	October 2010
Military Combat Helmet Standard for Ballistic Testing	December 2010
High Mobility Multi-purpose Wheeled Vehicle (HMMWV) Expanded Capacity Vehicle (ECV) Family of Vehicles (FoV)	February 2011
Ship Self-Defense Operational Mission Capability	March 2011
Special Operations Force (SOF) Mine Resistant Ambush Protected – All Terrain Vehicle (M-ATV)	May 2011
Mine Resistant Ambush Protected (MRAP) Force Protection Industries (FPI) Cougar A1 and A2 Independent Suspension Systems (ISS)	June 2011
Stryker Double-V Hull (DVH) Infantry Carrier Vehicle (ICV)	August 2011
Patriot Post-Deployment Build (PDB)-6.5 System	September 2011
LFT&E REPORTS	
Medium Tactical Vehicle Replacement (MTVR) Family of Vehicles (FoV)	July 2011
Mine Resistant Ambush Protected (MRAP) All Terrain Vehicle (M-ATV) Underbody Improvement Kit (UIK)	September 2011
FOT&E REPORTS	
MH-60R Multi-Mission Helicopter and MH-60S Combat Support Helicopter	November 2010
AN/BQQ-10 Acoustic Rapid Commercial Off-the-Shelf (A-RCI) Sonar System Advanced Processor Build 2007 (APB-07) and AN/BYG-1 Combat Control System APB-07	July 2011
Joint Biological Point Detection System (JBPDS) Phase II Whole System Live Agent	August 2011
ANNUAL REPORTS	
Ballistic Missile Defense Systems (BMDS)	February 2011

TABLE 2. DOT&E OPERATIONAL ASSESSMENT REPORTS DURING FISCAL YEAR 2011

PROGRAM	DATE
Early Infantry Brigade Combat Team (E-IBCT) Increment 1	December 2010
Mobile Landing Platform (MLP) (Early Operational Assessment)	January 2011
Shadow Tactical Unmanned Aircraft System (TUAS)	February 2011
M4E1 Joint Chemical Agent Detector (JCAD)	March 2011
Nett Warrior	April 2011
Joint Tactical Radio System (JTRS) Handheld, Manpack, and Small Form Fit (HMS) Rifleman Radio	May 2011
Ship-to-Shore Connector (Early Operational Assessment)	June 2011
Battlefield Airborne Communications Node (BACN) RQ-4B Global Hawk Block 20 Joint Urgent Operational Need (JUON)	July 2011
F-15E Radar Modernization Program (RMP)	July 2011
Miniature Air-Launched Decoy – Jammer (MALD-J)	August 2011

TABLE 3. DOT&E MAJOR AUTOMATED INFORMATION SYSTEM (MAIS) REPORTS DURING FISCAL YEAR 2011

PROGRAM	DATE
Global Combat Support System – Marine Corps (GCSS-MC)/Logistics Chain Management (LCM) Block 1, Release 1.1	October 2010
Navy Enterprise Resource Planning (Navy ERP) Release 1.1	January 2011
EProcurement Release 1.1	February 2011
Operational Utility Evaluation of the Combat Information Transport System Vulnerability Life Cycle Management System Spiral 1.5	May 2011
Operational Utility Evaluation of the Combat Information Transport System Air Force Intranet Increment I	May 2011
Global Combat Support System – Army (GCSS-A) Release 1.1	May 2011
Global Command and Control System – Maritime Increment Two Release 4.1 (GCCS-M v4.1) Force-Level	June 2011
Financial Information Resource System (FIRST) Force Structure Data Management (FSDM) Version 2.2	July 2011
Global Command and Control System – Maritime Increment 2 Version 4.1 (GCCS-M v4.l) Unit-Level (UL)	July 2011
Defense Security Assistance Management System (DSAMS) Training Module (TM) Block 4	September 2011
Global Command and Control System – Joint (GCCS-J) Joint Operation Planning and Execution System (JOPES) Version (v) 4.2.1	September 2011
Common Aviation Command and Control System (CAC2S) Increment 1, Phase 1	September 2011

CONCLUSION

We continue to make progress implementing all my initiatives and providing decision makers with analytically sound, objective information. I remain committed to assuring the Defense Department's operational and live fire tests are robust, rigorous, objective, and clearly reported. It is with pleasure that I submit this report, as required by law, summarizing the operational and live fire test and evaluation activities of the Department of Defense during Fiscal Year 2011.

J. Michael Gilmore
Director