

INTRODUCTION

I was confirmed by the Senate on September 21, 2009, as the Director, Operational Test and Evaluation, and sworn in on September 23. It is a privilege to serve in this position. I will work to assure that all systems undergo rigorous operational test and evaluation to determine whether they are operationally effective, suitable, and survivable. I will also assure that both civilian and military decision makers know the test results so that they can make informed decisions about acquiring those systems and how to employ them.

With pleasure 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 2009.

Because I was confirmed late in the 2009 Fiscal Year, most of the content in the main body of this report is based on what occurred before my tenure began. This Introduction, in contrast, provides my views regarding how I will execute the duties of the office I now hold. For example, I will institute changes in test and evaluation to better support rapid acquisition of improved capabilities for our nation's deployed forces. I will also make certain that ongoing initiatives are aligned fully with the important changes brought about by the Weapon System Acquisition Reform Act of 2009.

ACQUISITION REFORM ACT OF 2009

Fielding systems quickly and successfully depends critically on starting programs right and having sufficient, competent oversight. These are central tenets of the Weapons System Acquisition Reform Act of 2009. Implementing the letter and intent of the Act is an important task. The law affects the requirements process; requests for proposals; development planning – especially with respect to reliability growth; the workforce; and contractual support with respect to conflict of interest.

The Act recognizes that “unrealistic performance expectations” and “immature technologies” are among the root causes of trouble in defense programs. I believe the test and evaluation community can, during the requirements-setting process, identify such potential problems early in the life of programs. Last year, DOT&E added four staff members to work within the Department's requirements-setting process – currently the Joint Capabilities Integration Development System (JCIDS) – to assure that requirements for major acquisition programs are feasible, testable, and relevant. DOT&E participation in requirements-setting is discussed further in the Initiatives section of this Introduction under the topic “Engage early in the requirements process.”

The Weapons System Acquisition Reform Act of 2009 provides for a Director of Systems Engineering and a Director of Developmental Test and Evaluation (DT&E). I plan to work closely with them both to assure that all test and evaluation activities of the Department of Defense are fully integrated and to reinvigorate robust systems engineering and development planning within the Department. Of particular importance is the Act's emphasis on reliability, availability, and maintainability in major defense acquisition programs. The Act calls on the new offices to report on whether the Services have plans for adequate numbers of trained personnel to improve reliability, availability, maintainability, and sustainability as an integral part of rigorous systems engineering and developmental testing. DOT&E continues to support training events in reliability growth and is requiring reliability growth to be addressed specifically in future test and evaluation plans. Such emphasis has been, and will continue to be, a priority for DOT&E. Later in this Introduction I review the progress the Department has made this year toward improving reliability.

The Act requires the Secretary of Defense to revise the Defense Supplement to the Federal Acquisition Regulation to provide uniform guidance and tighten existing requirements to guard against organizational conflicts of interest by contractors in major defense acquisition programs. This will affect how we obtain contract assistance, and in response DOT&E will increase its use of Federally Funded Research and Development Centers and bring jobs into the government.

INTRODUCTION

NEW INITIATIVES

I reviewed with the senior leadership of DOT&E the state of OT&E in light of the urgent needs of our deployed forces, the new legislation, and the existing priorities under which DOT&E has operated.

I will direct the energies of DOT&E into the following four initiatives, which subsume the office’s previous 2009 priorities, address the Acquisition Reform Act of 2009, and incorporate the intent of the Secretary of Defense. The initiatives I will undertake are the following:

1. Field new capability rapidly,
2. Engage early to improve requirements,
3. Integrate developmental, live fire, and operational testing, and
4. Substantially improve suitability before IOT&E.

The relationship between the office’s previous priorities and the 2010 Initiatives is illustrated in Table 1 below. In the following sections, I will examine the 2010 initiatives and the office’s performance with respect to the priorities that guided DOT&E actions during FY09.

TABLE 1. RELATIONSHIP BETWEEN DOT&E’S 2009 PRIORITIES AND 2010 INITIATIVES

2009 Priorities 2010 Initiatives	1. Improve Suitability	2. Instill Operational Realism in Testing	3. Provide Timely and Accurate Information	4. Engage Early	5. Institutionalize Continuous Process Improvement
1. Field rapidly		✓	✓	✓	✓
2. Engage early in requirements		✓	✓	✓	✓
3. Integrate testing	✓	✓	✓	✓	✓
4. Substantially improve suitability	✓	✓		✓	✓

1. Field new capability rapidly

Secretary of Defense Gates has made clear that his top priority is to get the capabilities needed by our fighting forces to them as quickly as possible. The test and evaluation community has played a key role in fielding new capabilities rapidly—a role that I want to further strengthen and make even more helpful. Examples include the Mine Resistant Ambush Protected Vehicle (MRAP), MQ-9 Reaper, and the A/AO-10 C. In these cases, actions taken by Service Operational Test Agencies saved weeks to months in the time-to-field. Many adopted the approach of combining testing with the training of the first unit to be equipped, which shortened the timeline, provided real-time rigorous and objective feedback on system performance, and assured that the tactics, techniques, and procedures (TTPs) our forces need to employ new equipment were ready as the equipment was deployed.

Probably the best example of successful rapid acquisition is the MRAP Combat Vehicle. According to Brigadier General Michael M. Brogan, USMC, Commander, Marine Corps Systems Command, in testimony before the House Armed Services Committee on October 8:

The entire program was accomplished within the existing acquisition regulation. All of the actions normally required of an acquisition category 1-D program have been done by MRAP. They weren’t all done in a normal sequence, and many of them were tailored. But they have all been accomplished. The key was to view those regulations as permissive, not prohibitive, to see opportunities and not challenges, to look for possibilities and not obstacles and always the focus was on the 19-year-old lance corporal that we are charged to support.

INTRODUCTION

At the same hearing, General Brogan also said that the involvement of DOT&E was a key factor in the success of MRAP – important vulnerabilities were discovered through testing, and design changes were accomplished in near real-time; testing also played a key role in developing TTPs. MRAP is now regarded as a model for rapid acquisition.

To extend DOT&E's efforts to support rapid fielding as far as possible, I have begun a systematic review of programs to assess whether there are remaining candidates for early fielding or accelerated testing. If testing has already confirmed that the system would be effective and suitable in the current theaters of operation, those findings will be identified to fielding authorities. If only a small amount of testing remains in order to make the determination, we will examine the possibility of accelerating testing. We will assess risk and assure that accelerated testing reveals full capabilities and limitations. In addition to programs themselves, I am reviewing T&E procedures to see if they can be streamlined to better support rapid fielding. I am also reviewing the mechanisms we have to provide feedback to Program Offices to assure that when testing indicates equipment has problems, we get the fix into theater quickly.

Developing TTPs is critical to assuring that our forces can make full use of new capabilities as soon as they are fielded. The Joint Test and Evaluation (JT&E) program has been very successful assisting Combatant Commanders (COCOMS) with Quick Reaction Tests that evaluate TTPs. The Quick Reaction Tests provide quick-turn, evaluated solutions, in this case within 10 months. We will continue to stress the availability of that resource to the Combatant Commanders and seek ways to conduct those tests more quickly. The JT&E Program, established in 1972, expanded its reach to the combatant commands with the addition of five new members on its Senior Advisory Council this year. The council now has representatives from Joint Staff, the Services, and seven of the 10 combatant commands. Central Command and Northern Command have been the most active in using the JT&E Program as a means of solving issues as evidenced by their sponsoring seven projects each.

The JT&E projects address a wide range of issues. For example, the Joint Sniper Defeat project developed TTPs for employing new technology to detect the direction of sniper fire and target a specific area when friendly forces are under sniper attack. The Joint Command and Control for Net-Enabled Weapons project developed the concept of operations and procedures for post-launch redirection of weapons like the Tomahawk cruise missile. The procedures allow a change of targets after a missile launch so that if a more valuable target emerges during fly-out it can be attacked.

One consequence of efforts to rapidly field new capability is that systems are committed to combat operations before full-rate production. Under that circumstance, Congress has required DOT&E to submit Early Fielding Reports. In FY09, DOT&E delivered two such reports in compliance with Title 10, Section 2399 of U.S. Code. Copies of these and all our reports were provided to the Combatant Commanders to support their fielding decisions and to make joint warfighters and commanders aware of systems' capabilities and limitations with respect to performance and mission accomplishment. DOT&E has established points of contact with each Combatant Commander to assure that they are aware of the capabilities and limitations – both the strengths and weaknesses – of systems that might be deployed to their theaters. In addition, DOT&E uses a classified website to make available DOT&E Annual Reports, Beyond Low-Rate Initial Production (BLRIP) Reports, and Early Fielding Reports to the Combatant Commanders and others who need them.

2. Engage early to improve requirements

The Department's experience indicates that unless programs start with clear, sensible, and rationalized requirements, the program and its testing suffer tremendously and to the detriment of our fighting forces. The DOT&E experience has been that no amount of testing can compensate or correct for unjustified or unrealistic performance expectations.

Program requirements are often identified but not supported by a rigorous analytic rationale. Such a rationale is essential for performing proper engineering trade-offs and making test decisions during design

INTRODUCTION

and development. In other cases, requirements are inconsistent with program funding and schedules or with Combatant Commanders' expectations. In the case of the Joint High Speed Vessel, for example, initial concepts of operations stated that the Combatant Commanders would use the vessel to conduct missions, such as support of Special Operations forces and providing joint command and control, that were inconsistent with the program's funding and threshold requirements. That funding and those requirements specified a commercial ferry to be used in benign environments. DOT&E's reporting on the results of an early operational assessment for the JHSV highlighted these issues for action by the Services and Combatant Commanders.

To engage early in the development of requirements, the test community must become involved in what is currently called the Joint Capabilities Integration and Development System (JCIDS). Participation in JCIDS fulfills a long-standing recommendation of the National Academies.

DOT&E staff members who assess programs are taking the following actions to assure that systems have adequate requirements and are tested in realistic operational environments:

- First, staff are reviewing requirements as they are developed within JCIDS to assure they are unambiguous, testable, operationally relevant, and technically realistic.
- Second, staff are reviewing the Test and Evaluation Strategy (TES) and Test and Evaluation Master Plan (TEMP) for each project and working with developmental testers to assure that testing in operational environments is initiated during development and continues with increasing stress of the system through operational testing.
- Third, staff are identifying operational concerns to Program Offices at the earliest possible time so that they can be resolved in a timely manner.

It is important to identify early in a program's life whether their requirements may necessitate the development of new test resources such as threats or targets. In 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, would require DOT&E to declare the IOT&E inadequate, and (2) there is not an adequate program to develop the lacking test capability. Only one test-critical resource shortfall (aerial target drones) has been so categorized this year.

3. Integrate developmental, live fire, and operational testing

DoD Instruction (DoDI) 5000.02 currently requires "integrated testing" but continues to treat developmental and operational testing as entirely separate. For example, the instruction states:

The Program Manager shall design DT&E objectives appropriate to each phase and milestone of an acquisition program. ... The O(perational) T(est) A(gency) and the PM shall collaboratively design OT&E objectives appropriate to each phase and milestone of a program, and these objectives shall be included in the Test and Evaluation Master Plan.

There will always be a need for dedicated operational testing to confirm systems work in combat. Nonetheless the separateness of developmental testing from operational testing has caused problems in the development process that have been documented by the Defense Science Board and the National Academies. Most notably the lack of operational realism in early testing hides failure modes and limitations that then become evident only at the end of a program when fixing the problems is expensive, time-consuming, and, often, simply not possible. The solution is to introduce greater realism into testing earlier in order to understand those failure modes. I will move the department forward to integrate developmental, live fire, and operational test and evaluation.

A key means to achieve integrated testing, endorsed by DOT&E and the Operational Test Agency Commanders in April 2009, is Design of Experiments (DOE). DOE comprises the early use of rigorous scientific and statistical methods to plan and execute tests, and evaluate their results. Properly used, DOE will result in more effective and efficient T&E. The DT&E and OT&E offices are working together with the Operational Test Agencies and Developmental Test Centers to develop ways to apply DOE

INTRODUCTION

across the whole development and operational test cycle for a program, not just for individual test events. One important advantage of DOE is that it allows a rigorous and objective statement to be made of the confidence levels we have in the results of the testing. The Weapons System Acquisition Reform Act of 2009 makes specific mention that, for cost estimates, the confidence level used in establishing an estimate must be disclosed along with the rationale for selecting such confidence level, and, if such confidence level is less than 80 percent, the justification for selecting a confidence level of less than 80 percent. The evaluation of performance revealed through testing should be stated with similar rigor whenever possible. I intend to move T&E forward to use DOE in all test programs and thus provide that rigor.

Developing a workforce of persons skilled in all aspects of DOE can take many years, and we will work to establish necessary training capabilities. But in the near term, we will continue to emphasize the process as outlined in the DOT&E / Operational Test Agency Commanders Design of Experiments agreement, i.e., begin in early concept exploration to identify driving factors and conditions and continue to explore them throughout the product life cycle. This process aligns with accepted system engineering best practices for the development, production, and fielding of reliable systems.

Getting early operational realism into developmental testing can occur only if the resources needed to do so are identified and allocated. This particularly relates to developmental testing conducted before IOT&E. Currently, DOT&E staff members are becoming more engaged in the planning of early testing to assure that performance requirements will be tested in relevant environments for operational testing. As a metric of our progress toward achieving this goal, the percent of programs with a realistic test environment documented in the TEMP at Milestone B is 86 percent, and at Milestone C is 94 percent. Further, only 7 percent had resource gaps that DOT&E had to identify at Milestone A, and 13 percent had gaps at Milestone B. The challenge will be to identify and use the needed test resources in the early stages of development to find problems and failure modes at a time when they are easier to fix.

4. Substantially improve suitability before IOT&E

Suitability, and specifically reliability, is the principal area in which systems are found to be deficient during operational testing. The Defense Science Board Task Force on Developmental Test and Evaluation (DT&E), which was chartered by the USD(AT&L) and DOT&E to examine the reasons behind high suitability failure rates, found the following:

...the single most important step necessary to correct high suitability failure rates is to ensure programs are formulated to execute a viable systems engineering strategy from the beginning, including a robust reliability, availability, and maintainability (RAM) program, as an integral part of design and development. No amount of testing will compensate for deficiencies in RAM program formulation.

The new Weapons System Acquisition Reform Act of 2009 and DoDI 5000.02 require a reliability growth program.

Reliability is also the main driver of life-cycle costs and warfighter confidence in systems, maintenance force size, spare parts needs, and, ultimately, mission success. Increased reliability and how to establish a good reliability growth program have been a chief policy initiative of DOT&E for a number of years. We have made some progress in this area through implementation of formal reliability policies by the military services, incorporation of formal reliability growth planning within development programs, and by conducting reliability testing throughout programs' development.

In December 2008 the Department reissued DoDI 5000.02 with new guidance addressing reliability. The Instruction required the following:

P[rogram] M[anager]s for all programs shall formulate a viable Reliability, Availability, and Maintainability (RAM) strategy that includes a reliability growth program as an integral part of design and development. RAM shall be integrated within the Systems Engineering processes, documented in the program's Systems Engineering Plan (SEP)

INTRODUCTION

and Life-Cycle Sustainment Plan (LCSP), and assessed during technical reviews, test and evaluation (T&E), and Program Support Reviews (PSRs).

For this policy guidance to be effective, the Services must incorporate formal requirements for early RAM planning into their regulations, and assure development programs for individual systems include reliability growth and reliability testing; ultimately, the systems have to prove themselves in operational testing. Incorporation of RAM planning into Service regulation has been uneven. The Air Force, instead of following the DoDI 5000.02, changed its regulation to read:

The PM shall implement a reliability growth program if the initial mandatory sustainment KPPs and supporting materiel reliability KSA are not met.

This regulation achieves the exact opposite of the guidance in DoDI 5000.02. It guarantees that reliability problems will be found too late to be corrected cost-effectively. Clearly more work needs to be done to implement the DoD Instruction.

A second way of measuring progress is to consider actual program planning. Currently, 44 percent of programs on oversight and reviewed this year have a reliability plan, and 45 percent of programs are tracking reliability. Of the programs on DOT&E's current oversight list that have completed IOT&E, 66 percent met their reliability requirements. While these numbers represent an improvement from 2008 (see Figure 1), there is substantial room for continued improvement.

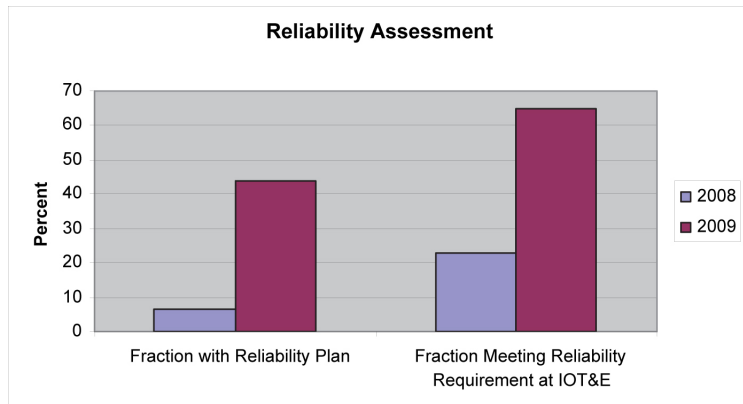


FIGURE 1. PROGRAM RELIABILITY PLANNING

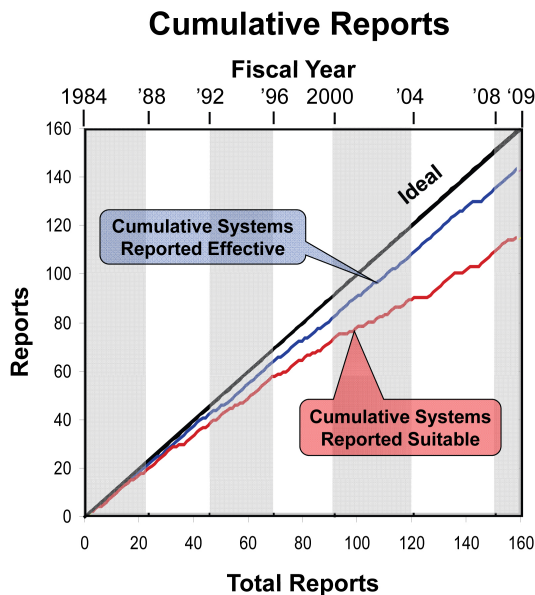


FIGURE 2. BEYOND LOW-RATE INITIAL PRODUCTION REPORT FINDINGS

Yet another way to monitor progress is to examine the results of testing as reflected in the reports we send to the Secretary and the Congress. This final measure responds slowly to the efforts we are making because programs take a long time to get to the final operational test, and improved processes at the end of a program have a difficult time compensating for problems that occurred before our efforts began. This fiscal year, we provided eight Beyond Low-Rate Initial Production reports for programs on oversight. Of those, two were not suitable for combat compared to two of nine the year before. The chart from last year's annual report has been updated in Figure 2 with the data from FY09 and shows no improvement in suitability. Over the 25 years of DOT&E's existence, about 75 percent of defense systems are found to be suitable in operational testing. As noted in the discussion of Figure 1, the current measure is worse than this.

Positive steps the Department took this past year to improve suitability include the following:

INTRODUCTION

- In June 2009, the Department published the Department of Defense Reliability, Availability, Maintainability, and Cost Rationale Report Manual (RAM-C) on realistic reliability, availability, and maintainability requirements and estimates describing methods for developing their life-cycle cost.
- The Department designated as a DoD Standard the ANSI/GEIA Standard 0009, Reliability Program Standard for Systems Design, Development, and Manufacturing to make it easy for program managers to incorporate the best reliability practices in requests for proposals (RFPs) and contracts. This is very important, because if the RFP does not ask for a reliability growth program, the contractor will not bid it; and, if reliability growth is not included in the subsequent contract with the winning bidder, they will not provide it. Designation as a DoD Standard allows (but does not require) program managers to incorporate compliance with the standard in contracts.

Actions taken specifically by DOT&E to improve suitability include the following:

- DOTE continues to support a training course for all of DoD in reliability growth engineering and testing.
- DOT&E continues to revamp its in-house training program, training staff to engage early in the development process by addressing requirements, operationally realistic test environments, and integrated testing.
- DOT&E now offers, as part of its professional development program, special training in RAM and DOE.
- DOT&E participates in the Program Support Reviews conducted by the System Engineering office of the USD(AT&L).

These initiatives will improve the reliability of our systems and should cause more systems to be evaluated as “suitable” during IOT&E. We have refined this priority into the initiative to “Significantly Improve Suitability before IOT&E.” It continues to be at the center of our attention as an organization. Going forward, DOT&E will work to assure that programs incorporate reliability growth planning, testing, and data collection at their inception, and practice reliability growth throughout their duration.

AREAS OF PARTICULAR CONCERN

Body Armor

During the last year, there was concern expressed by the Government Accountability Office (GAO) about the Army’s testing of body armor. GAO observed both Preliminary Design Model testing of new plate designs and then, further testing between November and December 2008, called First Article Testing, on those designs. GAO was concerned about the degree to which the Army followed its established testing protocols during these tests and whether the body armor purchased based on the tests would provide the needed protection to our Soldiers. The report noted however, “GAO did not provide an expert ballistics evaluation of the results of testing.”

Protecting our Soldiers is critical and I have engaged the National Academies and its experts to review the Army’s testing of body armor and make recommendations for improvement or correction regarding any and all of the issues raised in GAO’s report. The Army has embraced the need for this independent review by the National Academies.

Missile Defense

DOT&E has begun a study of the Department’s new four-phased, adaptive approach for missile defense in Europe. The goal of our study is to determine how the Missile Defense Agency’s plan for testing should be changed to incorporate realistic operational assessment of the capabilities provided under the phased adaptive approach. We will examine what can be tested, when it can be tested, and what rigor, objectivity, and confidence we can have in the test and evaluation results.

INTRODUCTION

OT&E MISSION ACCOMPLISHMENTS, FISCAL YEAR 2009

During this fiscal year, my office monitored 322 Major Defense Acquisition Programs (MDAPs) and special interest programs. We approved 50 Test and Evaluation Master Plans and Test and Evaluation Strategies, two LFT&E Strategies included in the Test and Evaluation Master Plans, and 70 Operational Test and Evaluation Plans for specific test events.

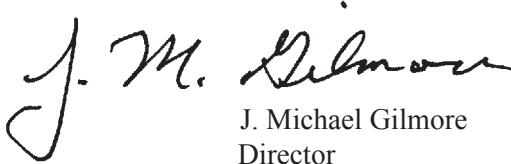
During FY09, DOT&E delivered eight BLRIPs (three of which were combined OT&E and Live Fire Reports) one report solely on live fire, and four Early Fielding reports to the Secretary of Defense and Congress (see Table 2).

TABLE 2. DOT&E REPORTING DURING FISCAL YEAR 2009

Program	Report Type	Date
Battlespace Command and Control Center (BC3) Air Force Central Command (AFCENT) Increment 1 Testing	OT&E Early Fielding Report	October 2008
MH-60S Block 3A Armed Helicopter Weapon System (AHWS)	Combined OT&E / LFT&E BLRIP Report	October 2008
Surface Electronic Warfare Improvement Program (SEWIP) – Block 1B2	OT&E BLRIP Report	October 2008
Logistics Vehicle System Replacement (LVSR)	LFT&E Report	December 2008
Guided Multiple Launch Rocket System (GMLRS) - Unitary (classified Annex)	Combined OT&E / LFT&E BLRIP Report	December 2008
MQ-9 Unmanned Aircraft System (UAS)	OT&E BLRIP Report	March 2009
Joint Biological Point Detection System (JBPDs)	OT&E BLRIP Report	June 2009
Air Force Mission Planning System (MPS) Increment III (F-16)	OT&E BLRIP Report	July 2009
Battlespace Command and Control Center (BC3) Air Force Central Command (AFCENT) Increment 2 Testing	OT&E Early Fielding Report	September 2009
MC-12W Liberty Project Aircraft (LPA)	OT&E Early Fielding Report	September 2009
Extended Range Multi-Purpose (ERMP) Unmanned Aircraft System Quick Reaction Capability	OT&E Early Fielding Report	September 2009
EA-18G Airborne Electronic Attack (AEA) Aircraft (classified Live Fire Report)	Combined OT&E / LFT&E BLRIP Report	September 2009
B-2 Radar Modernization Program (RMP) Mode Set One (MS 1)	OT&E BLRIP Report	September 2009

CONCLUSION

I am proud of the work DOT&E has done during this past year and I am honored to have been given the responsibility to lead this outstanding organization. I will build on DOT&E's success by helping to field new capabilities rapidly, engaging early in the requirements process, integrating developmental and operational testing, and substantially improving suitability at IOT&E. I am committed to assuring the Defense Department's operational testing and live fire tests are rigorous, objective, and clearly reported.


 J. Michael Gilmore
 Director