STATEMENT BY
CHARLES 'PETE' ADOLPH
TO THE
SENATE COMMITTEE
ON ARMED SERVICES

3 MARCH 2009

Mr. Chairman, members of the Committee, I am Pete Adolph, the chairman of a recent Defense Science Board Task Force study of Developmental Test and Evaluation. I am pleased to present a summary of the study results. The findings and recommendations I will discuss represent a consensus of the Task Force members and do not reflect an official position of the Department of Defense.

A Defense Science Board (DSB) Task Force on Developmental Test and Evaluation (DT&E) was convened in the summer of 2007 to investigate the causal factors for the high percentage of programs entering Initial Operational Test and Evaluation (IOT&E) in recent years which have been evaluated as both not operationally effective and not operationally suitable. The following are the specific issues which the Task Force was asked to assess:

- Office of the Secretary of Defense (OSD) organization, roles, and responsibilities for Test and Evaluation (T&E) oversight. Compare organization, roles, and responsibilities in both DT&E and Operational Test and Evaluation (OT&E). Recommend changes that may contribute to improved DT&E oversight, and facilitate integrated T&E.
- Changes required to establish statutory authority for OSD DT&E oversight. Title 10 United States Code (USC) has an OT&E focus, and does not address OSD authority in oversight of DT&E. Recommend changes to Title 10 or other U.S. statutes that may improve OSD authority in DT&E oversight.
- Many IOT&E failures have been due to lack of operational suitability. Specific problems have been in the materiel readiness sustainment areas of reliability, maintainability, and availability. Recommend improvements in DT&E process to discover suitability problems earlier, and thus improve likelihood of operational suitability in IOT&E.

PROBLEM DEFINITION

In recent years, there has been a dramatic increase in the number of systems not meeting suitability requirements during IOT&E. Reliability, Availability and Maintainability (RAM) deficiencies comprise the primary shortfall areas. DoD IOT&E results from 2001 to 2006 are summarized in Figures 1 through 3. These charts graphically depict the high suitability failure rates during IOT&E resulting from RAM deficiencies.

Program	Service	ACAT	IOT&E Result		Reason
FY 2001					
F-15 TEWS	USAF	II	Effective	Not Suitable	Reliability, Maintainability, Availability
V-22 Osprey	Navy	1D	Effective	Not Suitable	Reliability, Availability, Maintainability (RAM), Human Factors, BIT
Joint Direct Attack Munitions (JDAM)	USAF	1C	Effective only with legacy fuses	Not Suitable	Integration with delivery platforms
M2A3 Bradley Fighting Vehicle	Army	1D	Effective	Suitable	
FY 2002					
Joint Primary Aircraft Training System (JPATS)	USAF	1C	Effective with deficiencies	Not Suitable	RAM, Safety, Human Factors
Cooperative Engagement Capability (CEC)	Navy	1D	Effective	Suitable	
Multiple Rocket Launcher System (MLRS)	Army	1C	Effective	Suitable	
MH-60S	Navy	1C	Effective	Not Suitable	RAM, excessive administrative and logistic repair time impacted RAM
FY 2003					
B-1B Block E Mission Upgrade Program	USAF	1D	Effective	Not Suitable	16% decrease in weapons release rate, reduction in accuracy of Mark 82 low drag weapons, 14% hit rate on moving targets
Sea wolf Nuclear Attack Submarine	Navy	1D	Effective	Suitable	Several requirement thresholds were not met but overall system effective and suitable

Figure 1. DoD IOT&E Results FY 2001-2003.

Program	Service	ACAT	IOT&E Result		Reason
FY 2004					
Evolved Sea sparrow Missile	Navy	II	Effectiveness unresolved	Suitable	Testing was not adequate to determine effectiveness.
Stryker	Army	1D	Effective	Suitable	
Advanced SEAL Delivery System (ASDS)	Navy	1D	Effective with restrictions	Not suitable	Effective for short duration missions; not effective for all missions and profiles. Not suitable due to RAM.
Tactical Tomahawk	Navy	1C	Effective	Suitable	
Stryker Mortar Carrier-B (MC-B)	Army	1D	Effective	Not Suitable	RAM and safety concerns.
FY 2005					
CH-47F Block I	Army	1C	Effective	Not Suitable	RAM; communications system less suitable than CH-47D; did not meet Information Exchange Requirements for Block I.
F/A-22	USAF	1D	Effective	Not Suitable	RAM; needed more maintenance resources and spare parts; BIT
Joint Stand-Off Weapon-C	Navy	1C	Not Effective		Not effective against moderately hardened targets; mission planning time was excessive.
Guided-MLRS	Army	1C	Effective	Suitable	
High Mobility Attack Rocket System (HMARS)	Army	1C	Effective	Suitable	
V-22 Osprey	Navy	1D	Effective	Suitable	
EA-6B (ICAP III)	Navy	II	Effective	Suitable	

Figure 2: DoD IOT&E Results FY 2004-2005.

Program	Service	ACAT	IOT&E Result		Reason
CY 2006					
Common Missile Warning System (CMWS)	Army	1C	Effective	Suitable	Effective and suitable in the OIF/OEF environment but needs further testing outside of the OIF/OEF environment.
Deployable Joint Command and Control (DJC2)	Navy	1AM	Effective	Not Suitable	Operational Test Agency, COTF, reported effective, not suitable. BLRIP not complete.
Integrated Defensive Electronic Countermeasures	Navy	II			Test suspended due to reliability problems.
Surface Electronic Warfare Improvement Program (SEWIP) Block 1A	Navy	II	Not Effective	Not Suitable	Block 1A Upgrade does not make the AN/SLQ-32 EWS operationally effective and suitable but does enhance ability to protect ships
C-130J	USAF	1C	Effective single ship; Not effective in formation	Suitable with shortfalls	Effective single ship; not effective in formation air land / air drop; not effective in non-permissive threat environment. Shortfalls in suitability due to maintainability issues
Small Diameter Bomb (SDB) Increment 1	USAF	1D	Effective with limitations	Suitable with limitations	Limited effectiveness and suitability due to bomb rack reliability and deficiencies in software used to predict optimum fuzing solutions. Oct 2006 flight operations suspended

Figure 3: DoD IOT&E Results for 2006.

Early in the DSB study, it became obvious that the high suitability failure rates were the result of systemic changes that had been made to the acquisition process; and that changes in developmental test and evaluation could not remedy poor program formulation and execution. Accordingly, the Task Force study was expanded to address the broader programmatic issues, as well as the issues previously identified.

A number of major changes in the last 15 years have had a significant impact on the acquisition process. First, Congressional direction in Fiscal Year (FY) 1996, 1997, 1998 and 1999 Defense Authorization Acts reduced the acquisition workforce (which includes developmental test and evaluation). Several changes resulted from the implementation of Acquisition Reform in the late 1990s. The use of commercial specifications and standards was encouraged, unless there was justification for the use of military specifications. Industry was encouraged to use commercial practices. Numerous military specifications and standards were eliminated in some Service acquisition organizations. The requirement for a reliability growth program during development was also deemphasized, and in most cases, eliminated. At the same time, systems became more complex, and systems-of-systems integration became more common. Finally, there was a loss of a large number of the most experienced management and technical personnel in government and industry without an adequate replacement pipeline because of the personnel cuts. The loss of personnel was compounded in many cases by the lack of up-to-date standards and handbooks, which had been allowed to atrophy, or in some cases, eliminated. It should be noted that Acquisition Reform included numerous beneficial initiatives. There have been many programs involving application of poor judgment in the last 15 years that can be attributed to acquisition/test workforce inexperience and funding reductions. It is probable that these problems would have occurred independently of most Acquisition Reform initiatives.

All Service acquisition and test organizations experienced significant personnel cuts, the magnitude varying from organization to organization. Over time, in-house DoD offices of subject matter experts (who specialized in multiple areas, such as promoting the use of proven reliability development methods) were drastically reduced, and in some cases, disestablished. A summary of reductions in developmental test personnel follows. The Army essentially eliminated their military Developmental Testing (DT) component and declared the conduct of DT by the government to be discretionary in each program. The Navy reduced their DT workforce by 10 percent but no shift of "hands-on" government DT to industry DT occurred. The trend within the Air Force gave DT conduct and control to the contractor. Air Force test personnel have been reduced by approximately 15 percent and engineering personnel supporting program offices have been reduced by as much as 60 percent in some organizations. The reduction of Acquisition Program Office and Test personnel in the Services occurred during a time when programs have become increasingly complex **(**e.g., significant increases in software lines of code, off-board sensor data integration, and systems of systems testing**).**

PRINCIPAL FINDINGS AND RECOMMENDATIONS

RELIABILITY, AVAILABILITY, AND MAINTAINABILITY (RAM)

As a result of industry recommendations in the early 1970's, the Services began a concerted effort to implement reliability growth testing as an integral part of the development process. This implementation consisted of a reliability growth process wherein a system is continually tested from the beginning of development, reliability problems are uncovered, and corrective actions are taken as soon as possible. The Services captured this practice in their reliability regulations, and the Department of Defense (DoD) issued a new military standard on reliability, which included reliability growth and development testing as a best practice task. The goal of this process from 1980 until the mid-1990's was to achieve good reliability by focusing on reliability fundamentals during design and manufacturing rather than merely setting numerical requirements and testing for compliance towards the end of development.

The general practice of reliability growth was discontinued in the mid-to-late 1990's. This discontinuance may not be a direct result of Acquisition Reform, but may be related instead to the loss of key personnel and experience, as well as short-sighted attempts to save acquisition funds at the expense of increased life cycle costs. With the current DoD policy, most development contracts do not include a robust reliability growth program. The lack of failure prevention during design, and the resulting low initial Mean Time Between Failure (MTBF) and low growth potential are the most significant reasons that systems are failing to meet their operational reliability requirements.

According to Army studies, almost 90 percent of the sustainment costs are directly correlated with the reliability of the system. Given the amount of resources consumed during sustainment, investments in reliability enhancements can provide a very large return on that investment. A case study conducted by the Logistics Management Institute (LMI), provided data that indicated an investment in total program reliability would yield a substantial reduction in support costs.

FINDINGS

- Acquisition personnel reductions combined with acquisition system changes in the last 15 years had a detrimental impact on RAM practices

  - With some exceptions, the practice of reliability growth methodologies was discontinued during System Design and Development (SDD)
  - Relevant military specifications, standards and other guidance were not used
  - Suitability criteria, including RAM, were de-emphasized
- Improved RAM decreases life cycle costs and reduces demand on the logistics system
- The Deficiency Report (DR) can be a valuable tool for early identification of RAM-related suitability problems, when used in conjunction with an adequately resourced deficiency correction system

RECOMMENDATIONS

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 RAM program, as an integral part of design and development. No amount of testing will compensate for deficiencies in RAM program formulation. To this end, the following RAM-related actions are required as a minimum:

- Identify and define RAM requirements during the Joint Capabilities Integration Development System (JCIDS), and incorporate them in the Request for Proposal (RFP) as a mandatory contractual requirement
- During source selection, evaluate the bidders' approaches to satisfying RAM requirements
- Ensure flow-down of RAM requirements to subcontractors
- Require development of leading indicators to ensure RAM requirements are met
- Make RAM, to include a robust reliability growth program, a mandatory contractual requirement and document progress as part of every major program review
- Ensure that a credible reliability assessment is conducted during the various stages of the technical review process and that reliability criteria are achievable in an operational environment
- Strengthen program manager accountability for RAM-related achievements
- Develop a military standard for RAM development and testing that can be readily referenced in future DoD contracts
- Ensure a adequate cadre of experienced RAM personnel are part of the Service acquisition and engineering office staffs

ROLES AND RESPONSIBILITIES OF GOVERNMENT TEST AND EVALUATION ORGANIZATIONS

The role of the government in the DT process has evolved over the past 50 years. With some exceptions, there has been a significant decrease in government involvement in test planning, conduct and execution, in the last 15 years.

The traditional role of the government during the DT planning phase included the identification of the test resource requirements and government test facilities, the development of the test strategy and detailed test and evaluation plans, as well as the actual conduct of T&E. When a program moved from the planning phase to the test execution phase, the government traditionally participated in test conduct and analysis; performing an evaluation of the test results for the program office. With some exceptions, this is no longer the case. Until recently, it was recognized that there should be some level of government involvement and oversight even when the contractor has the primary responsibility regarding planning and execution of the DT program.

In addition to the reduction in the number of government acquisition and test personnel, the experience level of both government and industry personnel has steadily diminished in recent years. A significant percentage of the workforce became eligible to retire since 2000, and due to prior downsizing, there has not been a steady pipeline of younger technical personnel to replace them.

FINDINGS

The changes in the last 15 years, when aggregated, have had a significant negative impact on DoD's ability to successfully execute increasingly complex acquisition programs. Major contributors include massive workforce reductions in acquisition and test personnel, a lack of up-to-date process guidance in some acquisition organizations, acquisition process changes, as well as the high retirement rate of the most experienced technical and managerial personnel in government and industry without an adequate replacement pipeline.

- Major personnel reductions have strained the pool of experienced government test personnel
- A significant amount of developmental testing is currently performed without a needed degree of government involvement or oversight and in some cases, with limited government access to contractor data

RECOMMENDATIONS

- As a minimum, government test organizations should develop and retain a cadre of experienced T&E personnel to perform the following functions:
  - Participate in the translation of operational requirements into contract specifications,
  - and in the source selection process, including RFP preparation
  - Participate in DT&E planning including Test and Evaluation Master Plan (TEMP)
  - preparation and approval
  - Participate in technical review processes
  - Participate in test conduct, data analysis, and evaluation and reporting; with emphasis on analysis and reporting
- Utilize red teams, where appropriate, to compensate for shortages in skilled, experienced T&E domain and process experts
- Develop programs to attract and retain government personnel in T&E career fields so that the government can properly perform its role as a contract administrator and as a “smart buyer”

INTEGRATED TEST AND EVALUATION

Integrated testing is not a new concept within the Department of Defense, but its importance in recent years has been highlighted, due in part to the growth of asymmetric threats and the adoption of net-centric warfare. A December 2007 OSD Test and Evaluation Policy Revisions memorandum reinforces the need for integrated testing. Implementation of integrated test concepts has been allowed to evolve on an ad-hoc basis. The time has come to pursue more consistency in integrated test planning and execution.

Collaboration between developmental and operational testers to build a robust integrated test program will increase the amount of operationally relevant data that can be used by both communities. DT and Operational Test (OT) planning is separate and this inhibits efforts by the Services to streamline test schedules, thereby increasing the acquisition timeline and program test costs.

Additionally, there is a widely held assumption by many in the OT community that only data from independent OT is acceptable for operational evaluation purposes. While not all information from DT may be useable by the Operational Test Agency (OTA) to support IOT&E, a significant amount of developmental test data can be used to partially satisfy OT requirements. More importantly, an operational perspective earlier in the developmental process has often proven to be a catalyst to early identification and correction of problems.

DoD policy should mandate integrated test planning and execution on all programs to the extent possible. To accomplish this, programs must establish a team made up of all relevant organizations (including contractors, developmental and operational test and evaluation communities) to create and manage the approach to incorporate integrated testing into the T&E Strategy and the TEMP.

FINDINGS

- Service acquisition programs are incorporating integrated testing to a limited degree through varying approaches
- Additional emphasis on integrated testing will result in greater T&E process efficiency and program cost reductions

RECOMMENDATIONS

- Implement OSD and Service policy mandating integrated DT&E/OT&E planning and execution throughout the program
  - Require sharing and access to all appropriate system-level and selected component-level test and model data by government DT and OT organizations, as well as the prime contractor, where appropriate
  - Integrate test events, where practical, to satisfy OT and DT requirements

OPERATIONAL TEST READINESS REVIEW (OTRR)

Each Service has an Operational Test Readiness Review (OTRR) process. Although it varies from Service to Service, the process generally results in in-depth reviews of readiness to undergo an IOT&E event.

FINDINGS

- A Department of Defense Instruction requires that “the Service Acquisition Executive (SAE) shall evaluate and determine materiel system readiness for IOT&E"
  - Decision authority is frequently delegated to the appropriate Program Executive Officer (PEO)
  - Materiel developer is also required to furnish DT&E report to the Under Secretary of Defense for Acquisition, Technology and Logistics (USD[AT&L]) and Director, Operational Test and Evaluation (DOT&E)

- Shortcomings in system performance, suitability, and RAM are usually identified during the OTRR
- In most cases, the operational test readiness certifying authority is well aware of the risk of not meeting OT criteria when major shortcomings exist

- Because of funding constraints, the low priority given to sustainment, as well as the urgency in recent years to get new capabilities to the Warfighter, major suitability shortcomings have rarely delayed the commencement of dedicated IOT&E

RECOMMENDATIONS

- Conduct periodic operational assessments to evaluate progress and the potential for achieving pre-determined entrance criteria for operational test events
- Conduct an independent Assessment of Operational Test Readiness (AOTR) prior to the OTRR
- Include a detailed RAM template in preparation for the OTRR
- Require the Command Acquisition Executive (CAE) to submit a report to OSD that provides the rationale for the readiness decision

OSD TEST AND EVALUATION ORGANIZATION

The Task Force was asked to assess OSD roles and responsibilities for T&E oversight. T&E has been a visible part of OSD since the early 1970's, reporting to the Research and Engineering command section when it was in charge of acquisition oversight and subsequently to the Under Secretary of Defense for Acquisition (now AT&L). The early T&E office was responsible for all T&E, ranges, resources oversight, and policy. In 1983, Congress established an independent Director, Operational Test and Evaluation (DOT&E) organization, reporting directly to the Secretary of Defense (SECDEF), responsible for operational test and evaluation policy, budget review, and assessments of operational effectiveness and suitability. The Live Fire Test (LFT) oversight function was created and added to the DT&E office responsibilities in the mid 1980's. Later, the LFT oversight function was moved to the DOT&E organization.

In 1999, the DT&E organization was disestablished. Many functions were moved to DOT&E, including test ranges and resources, and joint T&E oversight. Some of the remaining T&E personnel billets were eliminated to comply with a congressionally mandated (AT&L) acquisition staff reduction. The residual DT&E policy and oversight functions were separated and moved lower in the AT&L organization.

A 2000 DSB Task Force Study on Test and Evaluation Capabilities recommended that DoD create a test and evaluation resource enterprise within the office of the DOT&E to provide more centralized management of T&E facilities. This recommendation ultimately led to removing the test ranges and resources oversight from DOT&E, abandoning the notion of centralized management, and the establishment of the Test Resource Management Center (TRMC) in AT&L (as directed by the National Defense Authorization Act for Fiscal Year 2003).

FINDINGS

Current policy as of December 2007 mandates that developmental and operational test activities be integrated and seamless throughout the system life cycle. There must be enough experts in OSD with the ability to understand and articulate lessons learned in early testing and the ability to execute the new T&E policy. That policy is to "take into account all available and relevant data and information from contractors and government sources" in order to "maximize the efficiency of the T&E process and effectively integrate developmental and operational T&E."

- Currently there is not an OSD organization with comprehensive DT oversight responsibility, authority or staff to coordinate with the operational test office
  - The historic DT organization has been broken up and residual DT functions were moved lower in organization in 1999, and lower yet in 2002
  - Programmatic DT oversight is limited by staff size and often performed by generalists vice T&E experts
  - Recruitment of senior field test personnel is hampered by DT's organizational status
  - Existing residual organizations are fragmented and lack clout to provide DT guidance
  - System performance information and DT lessons learned across DoD has been lost
  - DT is not viewed as a key element in AT&L system acquisition oversight
  - Documentation of DT results by OSD is minimal
- Access to models, data, and analysis results is restricted by current practice in acquisition contracting, and the lack of expertise in the DT organization
- TRMC has minimal input to program-specific questions or interaction with oversight organizations on specific programs
  - Organizational separation is an impediment

RECOMMENDATIONS

- Implementation of integrated and seamless DT and OT will require, at a minimum, greater coordination and cooperation between all testing organizations
- Consolidate DT-related functions in AT&L to help reestablish a focused, integrated, and robust organization
  - Program oversight and policy, and Foreign Comparative Test (FCT)
  - Have Director, DT&E directly report to Deputy Under Secretary of Defense, Acquisition and Technology (DUSD[A&T])
  - Restore TEMP approval authority to Director, DT&E
- Integrate TRMC activities early into DT program planning
  - Make TRMC responsible for reviewing the resources portion of the TEMP
- If such an organization is established and proves itself effective, consider as part of a future consolidation moving LFT back to its original DT location (this would require congressional action and DOT&E concurrence)

Most of the organizational changes recommended above are within the purview of AT&L. The LFT change requires the concurrence of DOT&E and a legislative change to Title 10 because of the change in reporting official. All the other recommendations can be implemented within current DoD authority.