Problem Discovery Affecting OT&E

Adequate developmental and operational testing are essential for determining whether systems provide an effective, suitable, and survivable warfighting capability to our Soldiers, Sailors, Airmen, and Marines. Developmental testing, in particular, serves as an early means to identify problems in the performance of weapon systems. The later a performance problem is discovered in a program's development timeline, the more costly and more difficult it is to correct it. Provided it is done adequately and rigorously, developmental testing also serves to determine if a program is ready for operational testing. Furthermore, discovery in operational testing has the potential to delay fielding while problems are corrected, or in the worst case, reveal a fatal flaw; neither of which is desirable.

Background

In 2010, Congress expressed concern that significant problems with acquisition programs are being discovered during operational testing that: (1) should have been discovered in development testing and (2) should have been corrected prior to operational testing. In response to this congressional request, I added this section to my annual report as a means to survey, across all DOT&E oversight programs, the extent of problem discovery occurring late in program development. Unfortunately, each year, operational testing continues to reveal performance problems for a significant number of programs that should have been discovered in developmental testing.

Evaluation of Problem Discovery

My evaluation of this issue falls into several cases, which are illustrated in Figure 1:

- ***Case 1.*** In the worst case (illustrated in red), problems were discovered solely in operational testing. The implication is that developmental testing (DT) was not conducted or was not adequate to uncover the problem prior to operational testing (OT). These cases illustrate that when decision makers focus too much on budget and schedule and not enough on the outcomes of testing (and the need to conduct adequate developmental testing), there is an increased likelihood of observing problems in operational testing.
- ***Case 2.*** A second case (illustrated in orange) includes those programs where problems were observed in operational testing that were also observed in developmental testing prior to the operational test period. Here, the implication is that the program chose to proceed to operational testing and accept the risk of potentially experiencing a poor operational testing outcome. Unfortunately, the problems were observed again and had an adverse effect on the determination of operational effectiveness, suitability, and/or survivability: a situation that is entirely avoidable.
- ***Cases 3 and 4.*** Two additional cases, illustrated at the bottom of Figure 1, show the desired paradigm: early testing is conducted; problems with system performance are uncovered and recognized for their potential effect on the upcoming determination of effectiveness, suitability, and survivability; and the program has the opportunity to resolve problems before entering operational testing.
  - In Case 3, programs made the decision to correct the problem(s) identified in early testing, which is laudable in light of the fact that it delayed the program and its entry into operational testing.
  - In Case 4, early testing uncovered problems, and the program has an opportunity to correct the problems. For this case, I recommend the program take action to address the issue before proceeding to the IOT&E/FOT&E period. It is noteworthy that many of the problems identified early were discovered during an operational assessment or limited user test; this reveals the value of conducting such early operationally realistic test events. I have expanded this section of the report over previous years, with specific details provided to enable programs to take action.

My discussion below identifies programs applicable to each of these cases and includes the reasons (if known) specific to each program.

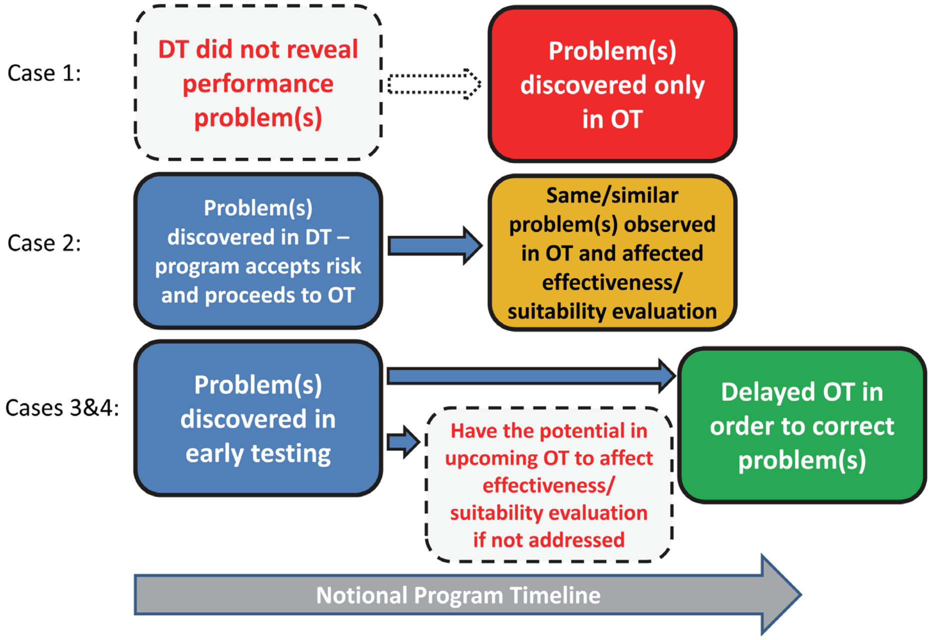

Case 1:
DT did not reveal performance problem(s)
Problem(s) discovered only in OT
Case 2:
Problem(s) discovered in DT – program accepts risk and proceeds to OT
Same/similar problem(s) observed in OT and affected effectiveness/ suitability evaluation
Cases 3&4:
Problem(s) discovered in early testing
Have the potential in upcoming OT to affect effectiveness/ suitability evaluation if not addressed
Delayed OT in order to correct problem(s)
Notional Program Timeline

Figure 1. Illustration of Problem Discovery Cases Observed in Oversight Programs

Conclusions

Some of the cases discussed below reveal that problem discovery only could have occurred in operational testing because that is when the operational implications of a performance deficiency become clear. This again reflects the value of operational testing – without such testing, the problems would have been discovered by the Services during operational use, and in the worst case, during actual conflict. There will always be a need for operational testing; nonetheless, in most of the cases below, the discovery of problems in operational testing was entirely avoidable.

Several solutions exist to curb the trends observed here:

- Programs should generate and execute schedules that allow adequate time for thorough developmental testing, and time to troubleshoot and resolve deficiencies. The results of testing should be used to guide program development decisions, including the need to extend developmental testing (and potentially delay operational testing until problems are corrected), and to ensure the system will meet its intended operational use.
- Programs should conduct developmental testing with a focus on the mission. In some cases, this will require developmental testing to go beyond specification compliance testing to demonstrate the desired system performance in an operational context.
- Services should develop concepts of operations and concepts of employment earlier so that developers can better understand how the system will be used in the field and can inform both system design and developmental test design.
- The requirements and acquisition communities need to work closely to develop requirement documents that ensure specification requirements are written to incentivize contractors and program managers to focus on demonstrating mission capabilities. These requirements should also clearly define performance expectations across the conditions the system is intended to be used, not just for a narrowly defined set of conditions.
- Often, effectiveness shortfalls and/or suitability shortfalls found in operational testing are discovered because operational use profiles (how the Soldier uses the equipment) reveal failure modes (reliability) or performance shortfalls that are unique to the operational test environment; such shortfalls would not have been revealed under the more structured, controlled, and benign conditions common to development testing. Development testing is often limited to verifying narrowly-defined requirements regardless of the operational relevance of those specifications. When the user takes the system to more operationally realistic conditions (more difficult threats; more difficult, but still relevant, operational environments), these performance failures are discovered.

The Deputy Assistant Secretary of Defense (DASD) Developmental Test and Evaluation (DT&E) is implementing initiatives consistent with these solutions that will be discussed in that office's upcoming report.

If requirements are set in a manner to ensure high performance under benign conditions, then developmental testing will likely only examine performance in those specified conditions. Therefore, well-defined requirements, especially the contractual specifications that are derived from the system's concept of employment, can help drive the developmental testing to examine performance under the conditions expected in the field. Furthermore, the early test events should also provide information to the requirements and resource sponsors for the system to ensure that the documented requirements are still relevant and feasible. Operational testing, by definition, must examine performance across the expected operational envelope.

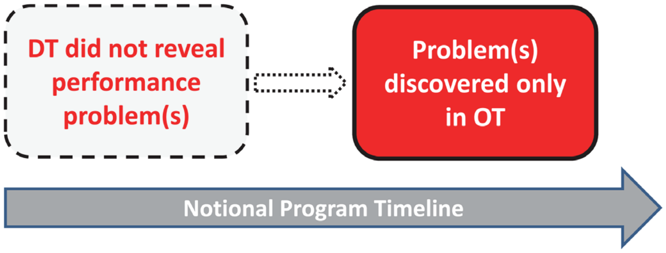
Summary

In 2013, 44 programs had significant problem discovery affecting OT&E. Of these, 12 are considered to be Case 1, meaning problems were discovered solely in operational testing (IOT&E or FOT&E). Ten programs fall into the Case 2 category, where problems that were identified in developmental testing were re-identified in operational testing. Six programs are considered to be Case 3, where problems were discovered in early testing and the program delayed operational testing to correct the problem. For these cases, I consider the developmental test and evaluation process to have been successful and the program to have responded appropriately. The remaining 16 programs fall under Case 4, where early testing has identified problems that need to be corrected. The value of this early identification of programs cannot be overstated. The benefit is lost, however, if these deficiencies are not corrected prior to IOT&E.

I have also included an assessment of cybersecurity vulnerabilities discovered during operational testing. I categorize these discoveries under Case 1, as they should have been discovered earlier in the systems' development. Operational testing of 33 programs in FY12 and FY13 revealed over 400 cybersecurity vulnerabilities, about 90 percent of which could have been found and corrected earlier in the systems' development.

I also provide updates to the problem discovery cases listed in my FY12 Annual Report. Last year, I documented 23 systems with significant discovery during testing: 6 of those systems had discovery in early testing, of which 5 implemented fixes that were verified by successful OT&E, are currently in OT&E, or are planning OT&E. Of the 17 programs that discovered significant issues during their IOT&E in 2011-2012, 10 have implemented fixes that were either verified in successful OT&E or are planning additional operational test periods; 2 of the remaining 7 programs were cancelled. Thus, while significant issues are being discovered late in the acquisition cycle, most programs are addressing the discoveries and verifying fixes in follow-on operational testing.

CASE 1:
PROBLEMS DISCOVERED IN 2013 DURING OPERATIONAL TESTING THAT SHOULD HAVE BEEN DISCOVERED DURING DEVELOPMENTAL TESTING


DT did not reveal performance problem(s)
Problem(s) discovered only in OT
Notional Program Timeline

IOT&Es IN FY13 WITH DISCOVERY	OTs (OTHER THAN IOT&E) IN FY13 WITH DISCOVERY
AIM-9X Air-to-Air Missile Upgrade	Acoustic Rapid Commercial Off-the-Shelf Insertion (A-RCI) AN/BQQ-10 (V) Submarine Sonar System
AIM-120 Advanced Medium-Range Air-to-Air Missile (AMRAAM)	Defense Enterprise Accounting and Management System (DEAMS)
Joint Battle Command – Platform (JBC-P)	DoD Automated Biometric Identification System (ABIS)
Miniature Air Launched Decoy (MALD) and MALD-Jammer (MALD-J)	Mk 54 Lightweight Torpedo
Multi-Static Active Coherent (MAC) System	Public Key Infrastructure (PKI) Increment 2
Surveillance Towed Array Sensor System (SURTASS) and Compact Low Frequency Active (CLFA)	Warfighter Information Network – Tactical (WIN-T)
All Programs Tested in FY12-13: Discovery of Cybersecurity Vulnerabilities	

Acoustic Rapid Commercial Off-the-Shelf Insertion (A-RCI) AN/BQQ-10 (V) Submarine Sonar System

A-RCI is composed of the computer processors and displays that process the data collected from submarines' acoustic arrays. It encompasses the primary components of U.S. submarines' combat systems and enables submarines to conduct all missions. The active operating mode of the Low Cost Conformal Array (LCCA), the mode in which the sonar pings and listens for the echoes, was unable to be evaluated due to a flaw in system software. Due to coding problems, the sonar was incapable of functioning in high reverberation environments, making detection of ships nearly impossible.

Early testing did not catch the problem because the software issue was not apparent in the more benign environmental conditions of the early developmental testing. The problem was discovered just hours before the commencement of the operational test of the system. Because of the late discovery, operational testing of the remaining components of the sonar system proceeded without examining the active operating mode capability.

Subsequent to the operational test, the Navy developed a software update to correct this issue and verified proper functionality with in-lab testing, including playback and analysis of recorded at-sea data. Operational testing of the active operating mode of the LCCA with this software update is still required and has not yet been conducted.

AIM-9X Air-to-Air Missile Upgrade

AIM-9X is the latest generation short-range, heat-seeking, air-to-air missile. IOT&E of the AIM-9X Block II missile was paused in April 2013 after multiple flight test failures. Two hardware reliability failures were traced to poor manufacturing. Additionally, IOT&E revealed problems with missile guidance. Missiles made porpoise-like maneuvers that contributed to misses when combined with inertial measurement units that showed errors occurring after launch shock. This launch shock problem occurred once during developmental testing, but the missile guided successfully to target. Currently the Program Office is pursuing root cause investigation with poor inertial measurement hardware units and guidance, navigation, and control (GNC) software as possible causes.

AIM-120 Advanced Medium-Range Air-to-Air Missile (AMRAAM)

The AIM-120 AMRAAM is a radar-guided air to-air missile with capability in both the beyond-visual-range and within-visual-range arenas. A single launch aircraft can engage multiple targets with multiple missiles simultaneously when using AMRAAM. Problems affecting missile performance and suitability were discovered in IOT&E in FY12, and the IOT&E was suspended until the problems were resolved. Specific details are classified. IOT&E resumed in May 2013, but the program continues to experience delays, and IOT&E is not projected to be complete until FY14.

Defense Enterprise Accounting and Management System (DEAMS)

DEAMS replaces legacy systems using an enterprise architecture with commercial off-the-shelf (COTS)-based financial accounting software (such as general ledger, accounts payable, accounts receivable, financial reporting, and billing). An initial operational assessment (OA-1) occurred in 2012, commensurate with the initial limited deployment of the system. The Air Force Operational Test and Evaluation Center began a second operational assessment (OA-2) of DEAMS Release 2.2 in August 2013, with the intent to determine if the issues discovered during OA-1 were remedied, and that processes and procedures had been put in place to allow for continued operational use.

Although the OA was not a formal IOT&E, it was conducted on a live and fielded system; many of the problems discovered could have been found earlier had adequate developmental testing been conducted. Results of OA-1 and initial deployment indicated numerous software defects (over 200) and showed that there was essentially no method or process for adequate configuration control. Furthermore, the live system was used to troubleshoot and fix severe deficiencies instead of employing a robust developmental regression testing process. A degree of regression testing automation is being employed that should reduce developmental test time and allow for greater depth of testing in future code development.

DoD Automated Biometric Identification System (ABIS)

The DoD ABIS is the result of a Joint Urgent Operational Need request and consists of information technology components and biometric examiner experts that receive, process, and store biometrics from collection assets across the world, match new biometrics against previously stored assets, and update stored records with new biometrics and contextual data to positively identify and verify actual or potential adversaries. While operational as ABIS 1.0, the system has not had any formal OT&E in its over 10-year existence, with only limited testing done by the Program Management Office and users to support new software releases, specifically ABIS 1.2.

Since 2010, there have been four failed attempts to deploy the ABIS upgrade, with the latest failed attempt in August 2013. The upgrade disabled critical interfaces with ABIS customers, preventing high-priority customers from receiving timely, accurate match results while maintaining compliance with established sharing agreements. The Director, Defense Forensics and Biometrics Agency recommended that the legacy ABIS 1.0 be restored after customers reported significant operational impacts to missions. Issues discovered during these deployment attempts should have been found beforehand through developmental test and evaluation.

Joint Battle Command – Platform (JBC-P)

JBC-P is a multi-Service situational awareness and mission command tool that automatically propagates the position of friendly forces, allows friendly forces to manually place allied and threat elements, and allows units to send preformatted and free-text messages across echelons from individual vehicles to Corps headquarters.

The JBC-P system exhibited problems in operational testing that were not identified in developmental testing, including spontaneous computer reboots, software unpredictability, and message management problems (duplicate entries and message format changes during transmission). Reliability failure modes were observed in the IOT&E that had not been observed in previous developmental testing, which indicates that the system's software development was immature.

Miniature Air Launched Decoy (MALD) and MALD-Jammer (MALD-J)

MALD is a small, low-cost, expendable, air-launched vehicle that replicates how fighter, attack, and bomber aircraft appear to enemy radar operators. The Air Force designed the MALD-J as an expendable, close-in jammer to degrade and deny an early warning or acquisition radar's ability to establish a track on strike aircraft while maintaining the ability to fulfill the MALD decoy mission. MALD-J IOT&E was conducted throughout FY13. The MALD and its follow-on MALD-J variant have been extensively tested over a number of years. However, the MALD-J variant poses significant potential for self-interference and is particularly reliant on accurate navigation to remain effective.

All MALD-J vehicles launched during developmental testing performed within the navigational accuracy requirements. During IOT&E at an open-air flight test range (a more challenging operationally representative environment), several MALD-J vehicles experienced unexpected navigational accuracy issues. There were several different causes of the navigational errors, all classified, but all arose from technical performance issues that should have been uncovered during developmental testing.

Mk 54 Lightweight Torpedo

The Mk 54 Lightweight Torpedo is the primary Anti-Submarine Warfare (ASW) weapon used by U.S. surface ships, fixed-wing aircraft, and helicopters. In May 2013, for one phase of operational testing of the Mk 54 torpedo with Block Upgrade software, the Navy planned to launch the weapons from MH-60R helicopters against a stationary submarine surrogate target off the coast of California. The plans called for the use of specific torpedo tactical presets that had been optimized for this scenario. This preset had not been examined in developmental testing.

Discussions between fleet aviation personnel, Navy testers, and torpedo developers revealed that the MH-60R could not execute the desired presets and that published tactical guidance and documentation were inaccurate. This incident led to a broader Navy investigation that identified gaps in communication and coordination between the undersea warfare community, which manages the torpedo programs, and the Naval aviation community, which is responsible for airborne fire control systems and tactical development.

Multi-Static Active Coherent (MAC) System

The MAC system is an active sonar system composed of two types of buoys (source and receiver) and an acoustic processing software suite. It is employed by the Navy's maritime patrol aircraft (P-3Cs and eventually P-8As) to search for and locate threat submarines in a variety of ocean conditions. During operational testing of the MAC sonobuoys system, P-3C maritime patrol aircraft deployed and monitored large fields of these sonar sensors in order to search for target submarines. As per approved test plans, the Navy conducted the tests at various sites in order to evaluate MAC detection capability in a variety of acoustic environments. Relevant conditions include sound speed profile, ambient noise, bathymetric profile, and bottom composition.

Testing revealed that the presentation of a valid target to the operator can vary significantly between environments and likely target types, making operator training and recognition of target-specific characteristics critical to performance. These differences were not identified in developmental testing, since all developmental testing was restricted to an environment where these effects could not have been studied. Data from a May 2013 test had to be invalidated because of the discovery of the phenomenon during the operational testing. Based on the data collected in operational testing, the Navy revised the employment concept and conducted additional training for the crews, and then repeated the operational test in October 2013.

Public Key Infrastructure (PKI) Increment 2

PKI Increment 2 provides authenticated identity management via password-protected Secret Internet Protocol Routing Network (SIPRNet) tokens to enable DoD members and others to access the SIPRNet securely, and encrypt and digitally sign e-mail. The Joint Interoperability Test Command conducted a combined FOT&E I and II of the PKI Increment 2 from January 8 through February 1, 2013, to verify correction of system deficiencies discovered during the IOT&E in 2011 for Spirals 1 and 2, and to evaluate preliminary Spiral 3 enhancements, respectively. The FOT&Es were originally scheduled to be completed in FY12, but were postponed due to system development delays. Furthermore, a stop-test in December 2012 resulted from systemic configuration management problems and lack of coordinated test-preparation. Delays in delivering the Integrated Logistics System (ILS) capability for token ordering and shipping contributed to delays in the delivery of several key Spiral 3 capabilities, including an Alternate Token Capability to support system administrator roles on the SIPRNet.

The FOT&E identified problems with blacklisting and token reuse in the token management system, and the operational testing exposed usability and auditing problems in ILS; none of these areas were adequately examined during developmental testing. The ILS was not effective for tracking tokens returned for reuse, was cumbersome to use, and did not provide the necessary functions to replace existing spreadsheet tracking mechanisms. More operationally relevant use cases should have been executed during developmental testing to avoid discovering these problems in the operational test. System user involvement in developmental testing likely would have identified ILS inadequacies early in the system design and development.

Surveillance Towed Array Sensor System (SURTASS) and Compact Low Frequency Active (CLFA)

SURTASS/CLFA is a low frequency, passive and active acoustic surveillance system installed on tactical auxiliary general ocean surveillance ships as a component of the Integrated Undersea Surveillance System. The Navy conducted the first phase of IOT&E in the Western Pacific in September 2012 to evaluate the ability of SURTASS/CLFA to detect submarine targets at long ranges as part of a large area search. The test revealed that the system is prone to detecting surface ships and presenting them as valid submarine targets, creating a false alarm problem. Although similar results were seen in developmental testing, the significance of the problem was only made clear when the system was put in an operationally realistic war time scenario.

Warfighter Information Network – Tactical (WIN-T)

WIN-T is a three-tiered communications architecture (space, terrestrial, and airborne) serving as the Army's high-speed and high-capacity tactical communications network. Testing of the WIN-T vehicle kits, specifically the Soldier Network Extension and the Point of Presence, during the WIN-T IOT&E in May 2012 and the WIN-T FOT&E in May 2013 showed that the systems were too complex for Soldier operation and troubleshooting. Additionally, mission command applications were sluggish. These key problems were not identified in the Risk Reduction Events (conducted at contractor facilities using engineers as operators) held prior to the operational tests.

Discovery of Cybersecurity Vulnerabilities

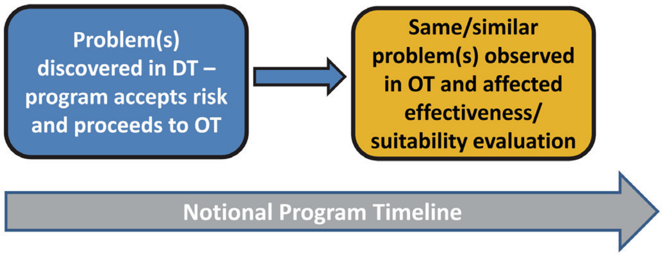
Where appropriate, programs that conducted operational testing in FY13 included a cybersecurity assessment – suitably scoped for the system under test – as part of the operational test program. DOT&E assessed 33 of these programs from FY12 and FY13 whose operational tests included cybersecurity assessments.

Over 400 Information Assurance (cybersecurity) vulnerabilities were uncovered during the vulnerability assessment and/or the penetration testing that occurred during the operational test period. Of those, approximately half were serious (Category 1) vulnerabilities that could allow debilitating compromise to a system, and approximately three-quarters of the systems reviewed had one or more serious vulnerabilities. The three most common Category 1 vulnerabilities were: (1) out-of-date/unpatched software, (2) configurations that included known code vulnerabilities, and (3) the use of default passwords in fielded systems. All of the problem discoveries could have and should have been identified prior to operational testing.

An assessment of the problems found reveals that only about 11 percent of those 400 vulnerabilities required an operational environment/operational test to uncover; 89 percent of the 400 vulnerabilities found in FY12 and FY13 could have been found in developmental testing. The review did not demonstrate whether these vulnerabilities were discovered in developmental testing but not remediated (Case 2 below), or if they were uniquely discovered in operational testing due to an inadequate developmental test process. However, the fact that so many vulnerabilities are being found late in a program's acquisition cycle is one of the main reasons why DOT&E and USD(AT&L) are collaborating on a revised cybersecurity policy. There is general agreement that systems must be assessed for cybersecurity earlier in a system's development. Testing over the past several years has indicated the need to move the discovery and resolution of system vulnerabilities earlier in program development, and the revised cybersecurity T&E process addresses this need.

CASE 2:
PROBLEMS IDENTIFIED IN DT&E THAT WERE RE-IDENTIFIED IN OT&E

Beginning this year I am reporting findings for oversight programs for which problems were identified in DT&E and then were re-identified in OT&E (10 programs). This is illustrated as the second type of undesirable problem discovery, since it could have been avoided.


Problem(s) discovered in DT – program accepts risk and proceeds to OT
Same/similar problem(s) observed in OT and affected effectiveness/ suitability evaluation
Notional Program Timeline

PROBLEMS IDENTIFIED IN DT&E THAT WERE RE-IDENTIFIED IN OT&E	
AIM-120 Advanced Medium-Range Air-to-Air Missile (AMRAAM)	Global Command and Control System – Joint (GCCS -J)
Cooperative Engagement Capability (CEC)	H-1 Upgrades – U.S. Marine Corps Upgrade to AH-1Z Attack Helicopter and UH-1Y Utility Helicopter
E-2D Advanced Hawkeye	Handheld, Manpack, and Small Form Fit (HMS) Manpack Radio
F-15E Radar Modernization Program (RMP)	Mission Planning System (MPS)/Joint Mission Planning System – Air Force (JMPS-AF)
Global Broadcast System (GBS)	P-8A Poseidon Multi-Mission Maritime Aircraft (MMA)

AIM-120 Advanced Medium-Range Air-to-Air Missile (AMRAAM)

AIM-120 AMRAAM is a radar-guided air to-air missile with capability in both the beyond-visual-range and within-visual-range arenas. IOT&E began in 2012. Problems that had been identified in DT&E reoccurred, which caused a pause in the IOT&E until May 2013. Specific details are classified.

Cooperative Engagement Capability (CEC)

The CEC is a system of hardware and software that allows the sharing of radar and weapons systems data on air targets among U.S. Navy ships, U.S. Navy aircraft, and some U.S. Marine Corps units. Developmental testing of the USG-3B CEC variant installed on the E-2D Advanced Hawkeye, conducted in FY12, revealed problems with the system's determination of relative sensor alignment, problems related to the system's capability to maintain a consistent air contacts picture on other CEC platforms (such as CEC-equipped ships and E-2Ds), and reliability problems. These problems were re-discovered during FOT&Es conducted in FY13.

E-2D Advanced Hawkeye

The E-2D Advanced Hawkeye is a carrier-based Airborne Early Warning and Command and Control aircraft. The Navy conducted the E-2D IOT&E from February to September 2012. Four major deficiencies, found during developmental testing, were also observed during the IOT&E:

- Accuracy issues found in developmental testing still existed in IOT&E.
- Because CEC software deficiencies that caused the CEC system to create multiple tracks for the same contact were still occurring at the start of the E-2D IOT&E, CEC testing was decoupled from the E-2D IOT&E. The multiple track problem remained during the CEC FOT&E that occurred immediately after the E-2D IOT&E.
- Radar track re-labeling was observed in developmental testing, but the full magnitude of the problem only manifested itself under the conditions of IOT&E.
- Poor radar reliability and availability were seen in developmental testing and persisted into IOT&E.

F-15E Radar Modernization Program (RMP)

The F-15E is a twin engine, tandem seat, fixed-wing, all weather, multi-role fighter aircraft. The RMP replaces the F-15E legacy APG-70 mechanically scanned radar with an active electronically scanned array system designated the APG 82(V)1, and is designed to retain functionality of the legacy radar system while providing expanded mission employment capabilities. F-15E RMP developmental flight testing began in January 2011. IOT&E started in April 2013 and completed in September 2013. The program experienced software maturation challenges during developmental test. Radar software maturity anomalies resulted in multiple unplanned software releases requiring additional regression testing to mature the radar functionality.

The program originally intended that later operational flight program releases would focus on software stability/Mean Time between Software Anomaly (MTBSA) fixes without additional functionality and performance changes. Due to challenges in maturing performance and functionality, the program exhausted its developmental schedule and funding before achieving the user's MTBSA requirement. Preliminary results from operational testing show software stability performance did not meet the 30-hour MTBSA goal, as predicted in the FY12 Annual Report.

Global Broadcast System (GBS)

The GBS is a one-way satellite communications system that works in a manner similar to satellite television. The Defense Enterprise Computing Center (DECC) upgrade consolidates several Navy ground sites into a single facility that creates broadcasts and provides technical support to users. The Air Force conducted a Force Development Evaluation of the GBS DECC upgrade from July through September 2013 at the Oklahoma City, Oklahoma DECC site; Mechanicsburg, Pennsylvania, DECC site; and Schriever AFB, Colorado.

Problems were discovered in developmental testing when users attempted to reauthorize receive suites to participate in the network. The program took corrective actions, but because of cost and schedule constraints, chose not to conduct additional developmental testing to verify these corrective actions were sufficient to provide system restoral capability. During operational testing, the same problems were seen.

The inexperience of personnel, poor operating procedures, and technical shortcomings were noted in previous developmental testing. Operational testing found similar deficiencies. Training and documentation for the GBS Operations Center personnel were not suitable for troubleshooting GBS user problems. Operations Center personnel needed to call contractor support to resolve more than half of the technical help desk tickets submitted during the operational test. Also, while transitioning from the main site at Oklahoma City to the backup site at Mechanicsburg, the absence of automated processes for reauthorizing users contributed to the extended time it took to restore service to all GBS users. The program knowingly entered operational testing with these immature procedures in place.

Global Command and Control System – Joint (GCCS-J)

GCCS-J is a command and control system utilizing communications, computers, and intelligence capabilities. The system consists of hardware, software (commercial and government off-the-shelf), procedures, standards, and interfaces that provide an integrated near real-time picture of the battlespace necessary to conduct joint and multi-national operations. Operational testing of GCCS-J version 4.3 Global was originally planned for May 2013; however, because of system immaturity, the program decided to conduct additional developmental testing to allow more time to find and fix deficiencies. Operational testing was conducted in August 2013, and while not adequate, was sufficient to determine that the system is not effective and not suitable.

While laudable that the program delayed operational testing to conduct additional developmental testing, several significant deficiencies were identified again during the second developmental test period, and the program did not again delay entry into operational testing, where the deficiencies were found again. Deficiencies included:

- Target lists that have been created and locked in GCCS-J 4.3 cannot be opened as read only using legacy versions of GCCS-J.
- The fielded version of the Generic Area Limitation Environment used to process electronic intelligence data could not pass processed data to the GCCS-J Common Operational Picture.
- Target lists take too long to replicate between GCCS-J 4.3 and legacy versions of GCCS-J. This issue was also seen during developmental testing, and must be retested using an operationally relevant test server.
- When large target lists are being synchronized across multiple versions of GCCS-J, the list is marked "validated" or "approved" before the synchronization process has completed. This will require a change to the synchronization process, followed by retesting using an operationally relevant test server.
- The process of upgrading the target folders in the new database structure resulted in incorrect security classification markings being used. At a minimum, the target folder should reflect the highest classification level of any information contained in the target folder.

H-1 Upgrades – U.S. Marine Corps Upgrade to AH-1Z Attack Helicopter and UH-1Y Utility Helicopter

This program upgrades the AH-1W attack helicopter to AH-1Z and the UH-1N utility helicopter to the UH-1Y. In 2010, the Navy began full-rate production and fielding of the AH-1Z aircraft following successful completion of Phase III IOT&E. Since 2010, the Navy has continued to develop software to correct previously noted deficiencies and provide new capabilities. By 2012, Software Configuration Set (SCS) version 6.0 had become mature enough to warrant FOT&E before fielding the new version. The Navy requested that Commander, Operational Test and Evaluation Force conduct FOT&E (OT-IIIB) of the new version of software.

Effectiveness, suitability, and survivability of H-1 Upgrades aircraft with SCS 6.0 are degraded by occasional software blanking of the electronic warfare display. If SCS 6.0 detects any failure (actual or false) in the aircraft survivability equipment (APR-39 and AAR-47), SCS 6.0 causes the electronic warfare display to go blank. Manual deployment of chaff and flares remains possible. Although detected during developmental testing, the operational implications of this loss of electronic warfare situational awareness were not apparent until operational testing.

Handheld, Manpack, and Small Form Fit (HMS) Manpack Radio

The HMS program evolved from the Joint Tactical Radio System program and provides software-programmable digital radios to support tactical communications requirements. The Manpack radio is a two-channel radio with military GPS. The Deputy Assistant Secretary of Defense for Developmental Test and Evaluation (DASD(DT&E)) stated in 2012 that the Manpack radio was not sufficiently mature to enter Multi-Service Operational Test and Evaluation (MOT&E). Waveform performance, particularly for the Single Channel Ground and Airborne Radio System (SINCGARS) was poor, and reliability was very low. However, the Army proceeded to conduct the MOT&E.

DOT&E assessed the Manpack as not operationally effective and not operationally suitable, primarily because of SINCGARS performance and low reliability. The Army has not conducted operational testing since the May 2012 MOT&E to demonstrate improvements to Manpack. There have been multiple low-rate initial production procurements totaling 5,326 radios, and the Army has fielded the system to the 101st Airborne Division.

Mission Planning System (MPS)/Joint Mission Planning System – Air Force (JMPS-AF)

MPS is a package of common and platform-unique mission planning applications. The IOT&E for the JMPS Mission Planning Environment version 1.3 for the E-8 Joint Surveillance Target Attack Radar System began in 2011. During this initial phase, incorrect magnetic variation computations and unreliability of the process to transfer mission planning data to the aircraft were uncovered; these problems had also been observed in developmental testing prior to IOT&E. The operational test was paused and restarted more than a year later to ensure that these deficiencies had been corrected.

The program went back into testing in 1QFY13, demonstrating that these two deficiencies were corrected. Other problems observed during developmental testing and found again during the first phase of the IOT&E include:

- The system's inability to automatically calculate flight plans with orbits based on user inputs
- Problems calculating take-off and landing data
- Failures in the implementation of vector vertical obstruction data

P-8A Poseidon Multi-Mission Maritime Aircraft (MMA)

The P-8A Poseidon MMA is a fixed-wing aircraft that will replace the P-3C Orion; its primary mission is to detect, identify, track, and destroy submarine targets (ASW), but it also is intended to conduct Anti-Surface Ship Warfare and Intelligence, Surveillance, and Reconnaissance (ISR). The Navy conducted IOT&E of the P-8A Increment 1 from September 2012 through March 2013. Nearly all of the major deficiencies that were identified during the developmental test period were re-discovered during the IOT&E; many of these deficiencies led to DOT&E determining that P-8A is not effective for the ISR mission and is unable to execute the full range of ASW Concept of Operations at its Initial Operational Capability (IOC).

Prior to IOT&E, DOT&E sent two memoranda to the Navy emphasizing the potential operational impact of critical performance deficiencies identified during developmental testing.

- Synthetic Aperture Radar imagery collection capabilities were severely limited due to radar stability problems, target cueing errors, and image quality problems, which severely degraded ISR mission performance.
- Communication and data transfer system interoperability problems limited receipt of tactical intelligence updates and transmission of P-8A imagery intelligence products to operational users.
- Electronic Support Measures deficiencies limited threat detection and localization, seriously degrading capabilities and aircraft survivability across all major missions.
- Developmental testing identified significant maritime surface target tracking errors while operating in the radar track-while-scan mode. Operational testing confirmed and further quantified these errors, which degrade operator capabilities to maintain an accurate surface operational picture while executing mission operations.

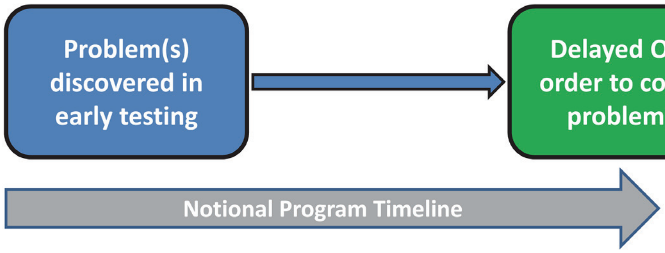
Detailed DOT&E analysis of developmental test results indicated that the P-8 radar was not meeting detection requirements for some types of critical surface targets. Operational testing confirmed these results and characterized the operational impact of the performance limitations on the ASW mission. Additional details are classified and can be found in DOT&E's October 2013 IOT&E report.

Although the P-8A Increment 1 system provides an effective small area, cued ASW search, localization, and attack mission capability, similar to the legacy P-3C system, the Navy's decision to cancel plans to integrate the Improved Extended Echo Ranging capability into P-8A ensured that the aircraft would have no wide-area ASW search capability at IOC. Additionally, fundamental limitations with the P-8A's current sensor technology restrict search capabilities against more stressing adversary targets, making the P-8A not effective at ASW in some mission scenarios. The Navy intends to use the Multi-static Active Coherent (MAC) sonobuoy system to address these shortfalls, and will test the capability in the P-8A Increment 2 program.

The Navy plans to conduct additional developmental testing after the IOT&E to verify the correction of some of the system deficiencies identified during IOT&E.

CASE 3: PROBLEMS DISCOVERED IN EARLY TESTING AND THE PROGRAM WAS DELAYED TO CORRECT THE PROBLEM

These cases could be considered instances in which the developmental test and evaluation process was successful and the program responded appropriately. Early testing can be both early developmental testing as well as operational assessments conducted prior to Milestone C. The latter have proven to be essential for identifying problems early.


Problem(s) discovered in early testing
Delayed OT in order to correct problem(s)
Notional Program Timeline

PROBLEMS IDENTIFIED IN DT&E THAT DELAYED OT&E	
Air Operations Center – Weapons System (AOC-WS)	Ground/Air Task Oriented Radar (G/ATOR)
Battle Control System – Fixed (BCS-F)	Small Tactical Unmanned Aerial System (STUAS) Tier II
F/A-18E/F Super Hornet and EA-18G Growler	Vertical Take-Off and Landing Unmanned Aerial Vehicle (VTUAV) (Fire Scout)

Air Operations Center – Weapons System (AOC-WS)

The AOC-WS is the senior command and control element of the U.S. Air Force's Theater Air Control System and provides operational-level command and control of air, space, and cyberspace operations, as well as joint and combined air, space, and cyberspace operations. The Air Force originally planned to conduct both developmental and operational testing of AOC-WS 10.1 Recurring Event (RE)12 in December 2012. The AOC-WS 10.1 RE12 test article and associated documentation that entered operational testing in August 2013 was the direct output of a thorough developmental test-fix-test cycle. Extended developmental test and evaluation efforts ensured that this test article successfully passed operational test Phase II without any significant deficiencies.

The RE12 test article in December 2012 was built on top of a flawed RE11 test baseline. The developmental test process recommended a clean rebuild of the RE11 baseline, followed by a rebuild of the RE12 test article. This was consistent with the plan for fielding to operational sites. Developmental testing in December 2012 identified 2 known significant deficiencies that had not been fixed and 10 new significant deficiencies. The developmental test-fix-test cycle continued until all significant deficiencies were verified fixed.

Battle Control System – Fixed (BCS-F)

The BCS-F is a tactical air battle management command and control system that provides the two continental U.S. North American Aerospace Defense Command (NORAD) air defense sectors, as well as the Hawaii and Alaska Regional Air Operation Centers, with COTS hardware using an open-architecture software configuration. The system operates within the NORAD air defense architecture and is employed by the U.S. and Canada. During developmental testing, several problems were found with the hardware and software configurations of the servers, firewalls, intrusion detection systems, and system guards that generated vulnerabilities in the system's defenses.

The start of IOT&E was delayed while the contractor and Program Office corrected the deficiencies and tested the corrections to ensure the deficiencies were fixed. A key problem underlying many of the deficiencies was that the documentation was insufficient, which contributed to problems with software installation and configuration.

F/A-18E/F Super Hornet and EA-18G Growler

The Super Hornet is the Navy's premier strike-fighter aircraft that replaces earlier F/A-18 variants in carrier air wings. The F/A-18E/F software is being incrementally upgraded. The most recent software version is known as Software Configuration Set (SCS) H8E. Phase 1 of operational testing for SCS H8E took place from June 2012 to May 2013 after a delay of six months, because the Navy discovered problems during developmental testing in 6 of the 14 new SCS H8E capabilities. Ultimately these problematic capabilities were deferred to a later operational test and SCS H8E (Phase 1) proceeded with the remaining planned capabilities.

Several of these deferrals resulted from the Navy's difficulty in integrating electronics support on the Super Hornet while others would have allowed the aircraft to detect the position of an emitter using onboard sensors only, integrate the latest version of a self-protection jammer, and navigate through civilian airspace using GPS navigation instead of the traditional Tactical Air Navigation (TACAN) system.

Ground/Air Task Oriented Radar (G/ATOR)

G/ATOR is a three-dimensional short- to medium-range tactical radar designed to detect, identify, and track low-level cruise missiles, manned aircraft, and unmanned aerial vehicles as well as rockets, mortars, and artillery fire. The Marine Corps' G/ATOR program conducted three developmental test periods beginning in July 2012 and continuing until April 2013. An operational assessment was to be conducted in April 2013, but because reliability problems primarily related to software deficiencies were identified during the preceding developmental test periods, the operational assessment was postponed and a Field Users Evaluation was conducted instead.

G/ATOR reliability-related software deficiencies have continued and have kept the radar from meeting its Mean Time Between Operational Mission Failure (MTBOMF) requirements. After allowing additional time for the software to further mature prior to the program's Milestone C decision (scheduled for 1QFY14), the program added a fourth developmental test period to assess improvement. While laudable, the program's reliability growth plan has not been fully defined; it remains unclear if G/ATOR will meet key reliability metrics by the start of IOT&E (scheduled for 3QFY17).

Small Tactical Unmanned Aerial System (STUAS) Tier II

The STUAS consists of five RQ-21A unmanned air vehicles, surface components, and assorted government-provided equipment; it is intended to provide units ashore with a dedicated persistent battlefield intelligence, surveillance, and reconnaissance capability. During integrated testing, developmental testers identified an issue with the STUAS sensor payload. Frequently during flight, the imagery provided by the payload would freeze, flicker, and drift, or the operators would lose payload control. The remedial action was to conduct a "soft" reset similar to rebooting a computer. If the soft reset (or multiple soft resets) did not restore payload functionality, the operator would conduct a "hard reset," which consisted of powering off and then powering on the payload. Developmental testers did not see the 1 to 4 minutes required to restore functionality as a detriment to system effectiveness.

During the operational assessment in support of Milestone C, the frequency of payload resets, along with the time required to restore functionality, caused operators to lose track of targets or interrupted ongoing missions; this caused operational testers to conclude that the payload reset issue had the potential to render the system not effective during IOT&E. Detailed analyses identified issues with the payload to air vehicle interface (electrical and software).

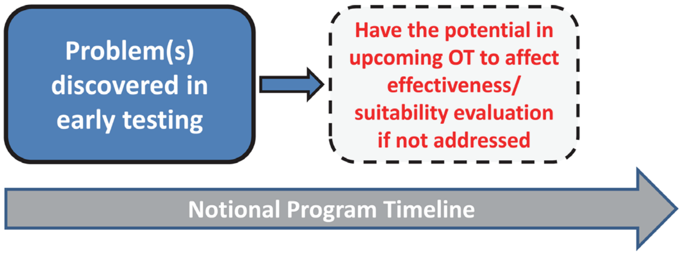
After Milestone C, the Program Office inserted an additional integrated test period before IOT&E and implemented modifications to the air vehicle, which contributed to a three-month delay in the IOT&E. The last integrated test period demonstrated that the payload reset problem has been corrected and that changes to the recovery procedures have resulted in less damage on recovery. As a result, these two are not expected to be issues for the IOT&E.

Vertical Take-Off and Landing Unmanned Aerial Vehicle (VTUAV) (Fire Scout)

The Fire Scout is a helicopter-based tactical unmanned aerial system comprised of up to three MQ-8 air vehicles with payloads, a shipboard integrated Ground Control Station with associated Tactical Common Data Link, and the UAV Common Automatic Recovery System. In 2009, the Navy produced a draft VTUAV Developmental Test to Operational Test Transition Report, which assessed the system's readiness to enter IOT&E using the MQ-8B air vehicle. The draft report stated: "The VTUAV system is not recommended to proceed to IOT&E based on the high risk of an OPEVAL [operational evaluation] determination of not operationally suitable." Because of this draft recommendation, VTUAV did not enter IOT&E as scheduled in early 2010. Since that time, the Navy decided not to proceed with full-rate production of the MQ-8B, and will delay the VTUAV IOT&E until the MQ-8C replaces the MQ-8B at some future date.

CASE 4:
PROBLEMS DISCOVERED DURING EARLY TESTING, THAT IF NOT CORRECTED, COULD ADVERSELY AFFECT MY ASSESSMENT OF OPERATIONAL EFFECTIVENESS, SUITABILITY, AND SURVIVABILITY DURING INITIAL OPERATIONAL TEST AND EVALUATION

I include this section of the report to identify early in a program's development problems that need to be corrected to improve the potential for a successful IOT&E. The list includes programs that conducted either early developmental testing or an operational assessment that was conducted prior to Milestone C. The latter have proven to be essential for identifying problems early and clearly continue to reveal their value to the acquisition process. Most of these entries identify problem discoveries in early testing that need to be corrected soon, as their IOT&E or FOT&E periods are approaching within the next two or three years.


Problem(s) discovered in early testing

Have the potential in upcoming OT to affect effectiveness/ suitability evaluation if not addressed

Notional Program Timeline

DISCOVERIES IN EARLY TESTING IN FY13 THAT SHOULD BE CORRECTED PRIOR TO IOT&E	
CVN-78 *Gerald R Ford* Class Nuclear Aircraft Carrier	LHA-6 Amphibious Assault Ship
Defense Enterprise Accounting and Management System (DEAMS)	Littoral Combat Ship (LCS) (Includes Seaframes and Mine-Countermeasures Mission Package with the Remote Minehunting System (RMS) and Airborne Mine Neutralization System (AMNS))
DoD Automated Biometric Identification System (ABIS)	M109 Family of Vehicles (FoV) Paladin Integrated Management (PIM)
Handheld, Manpack, and Small Form Fit (HMS) Manpack Radio	Next Generation Diagnostic System (NGDS)
Handheld, Manpack, and Small Form Fit (HMS) Rifleman Radio and Nett Warrior	Public Key Infrastructure (PKI) Increment 2
Integrated Defensive Electronic Countermeasures (IDECM)	Q-53 Counterfire Target Acquisition Radar System
Integrated Electronic Health Record (iEHR)	RQ-4B Global Hawk High-Altitude Long-Endurance Unmanned Aerial System (UAS)
Joint Warning and Reporting Network (JWARN)	Surface Ship Torpedo Defense (SSTD) System: Torpedo Warning System (TWS) and Countermeasure Anti-torpedo Torpedo (CAT)

CVN-78 *Gerald R Ford* Class Nuclear Aircraft Carrier

The CVN-78 *Gerald R. Ford* class of aircraft carriers is the first new aircraft carrier design in more than 30 years and will replace the CVN-68 *Nimitz* class. Compared to the *Nimitz* class, CVN-78 has design features intended to enhance its ability to launch, recover, and service aircraft, such as a slightly larger flight deck, dedicated weapons handling areas, and increased aircraft refueling stations. In FY13, the Navy completed an operational assessment for CVN-78 that examined design documentation and data from developmental testing.

The CVN-78 test schedule is aggressive, leaving little time to fix problems discovered in developmental testing before IOT&E begins. Based on past comments that CVN-78 had inadequate developmental testing, the Program Office has been working to incorporate additional developmental test events into the test program. Nonetheless, major developmental test events are still scheduled to occur after IOT&E begins. DOT&E concludes this aggressive schedule increases the likelihood that problems will be discovered during CVN-78's IOT&E, which could inhibit the successful completion of testing.

There are concerns with the reliability of key systems that support sortie generation on CVN-78. These systems include the new catapults, arresting gear, dual-band radar, and weapons elevators. These systems are critical to CVN-78 operations and will be tested for the first time in their shipboard configurations after they have been installed in CVN-78. To date, the Navy has conducted limited reliability testing of these systems. They have either poor or unknown reliability. Poor reliability of these critical systems could cause a cascading series of delays during flight operations that would affect CVN-78's ability to generate sorties, make the ship more vulnerable to attack, or create limitations during routine operations. DOT&E assesses the poor or unknown reliability of these critical subsystems will be the most significant risk to CVN-78's successful completion of IOT&E. If reliability problems with these systems drive CVN-78's sortie generation rate well below *Nimitz* performance, the result could be significant to strategic planners.

Due to known problems with aircraft carrier combat systems, there is a high risk that CVN-78 will not achieve its self-defense requirements. Although the CVN-78 design incorporates several combat system improvements over the *Nimitz* class, these improvements are unlikely to address all of the known shortfalls

CVN-78 cannot support multiple Common Data Links (CDLs) and this fact limits the carrier's ability to communicate with current and future systems, including MH-60 helicopters, P-3 and P-8 aircraft, unmanned aerial vehicles, and other assets. DOT&E concludes the lack of CDL coverage on CVN-78 will limit its operational effectiveness and pose a risk to successful completion of IOT&E.

Two common problems with the first ship of a new class is that training and documentation for new systems are provided too late to train the crew before the start of IOT&E; current CVN-78 plans indicate that these problems will affect CVN-78's IOT&E as well. The CVN-78 Master Integrated Schedule for Logistics shows the production status of required technical documentation. Based on that schedule, Integrated Logistics Support documentation for training, operation, and maintenance of many unique CVN-78 systems are likely to be delivered late.

Defense Enterprise Accounting and Management System (DEAMS)

DEAMS replaces legacy systems using an enterprise architecture with commercial off-the-shelf (COTS)-based financial accounting software (such as general ledger, accounts payable, accounts receivable, financial reporting, and billing). The Air Force began a second operational assessment (OA-2) of DEAMS Release 2.2 in August 2013. The intent of OA-2, to be completed in February 2014, is to determine if the issues discovered during a previous operational assessment (OA-1) in 2012 were remedied, and that processes and procedures have been put in place to allow for continued operational use. The DOT&E assessment from OA-1 cast doubts on the ability of the system to support financial management for the Air Force. In contrast, the current system has the potential to be both operationally effective and suitable. The problems below, some of which were mentioned in Case 1 above, have the potential to affect a future determination of effectiveness and suitability if not addressed.

- Feedback from new users at McConnell AFB, where DEAMS was deployed in October 2012, indicated that the training they had received was inadequate. They noted that it focused on navigating DEAMS but did not provide them with a real understanding of the system and its application to their day-to-day work process. McConnell users also stated that they need more on-site technical support during DEAMS implementation.
- Effective workarounds for existing software defects have been well documented at the Defense Finance and Accounting Service in Limestone, Maine, but workarounds have not been documented within the Air Force.
- Although configuration management has improved, a large number of defects remain open and several currently required capabilities and enhancements are still being developed and are not planned for implementation until 2014.
- The percent of subsidiary accounts reconciled to general ledger accounts does not meet the 95 percent threshold requirement. This could significantly affect the ability of DEAMS to attain an unqualified financial audit by FY17 as required.

DoD Automated Biometric Identification System (ABIS)

The DoD ABIS is the result of a Joint Urgent Operational Need request and consists of information technology components and biometric examiner experts that receive, process, and store biometrics from collection assets across the globe, match new biometrics against previously stored assets, and update stored records with new biometrics and contextual data to positively identify and verify actual or potential adversaries. While operational as ABIS 1.0, the system has not had any formal operational testing in its over 10-year existence, and the follow-on release, ABIS 1.2 has failed to demonstrate adequate maturity during four unsuccessful demonstrations since 2010.

Several ABIS 1.2 deficiencies have been identified during early testing including lack of approved requirements, lack of a baseline system against which to make comparisons, lack of configuration management plans and processes to support tracking of fixes and new requests, and lack of a standards conformance program to enable interoperability certification.

Unless all of the above prerequisites to a successful IOT&E are addressed, DoD ABIS 1.2 will likely be found not operationally effective nor operationally suitable in the IOT&E scheduled for 3QFY14.

Handheld, Manpack, and Small Form Fit (HMS) Manpack Radio

The HMS program evolved from the Joint Tactical Radio System program and provides software-programmable digital radios to support tactical communications requirements. The Manpack radio is a two-channel radio with military GPS. In September 2012, the Army conducted a Government Development Test (GDT) 3 to demonstrate improvements in deficiencies identified in the 2012 MOT&E. During GDT 3, the Manpack radio demonstrated improved waveform performance but poor reliability. If reliability is not improved, it could adversely affect the performance during the next operational test.

Additionally, a number of key Manpack required capabilities, such as the ability to pass data and voice between different radio networks, have not yet been fully tested. The Army plans to test these requirements during GDT 4 in January 2014. Conducting operational testing without proving these capabilities in a developmental test will increase the likelihood of Manpack demonstrating poor performance during operational testing.

Handheld, Manpack, and Small Form Fit (HMS) Rifleman Radio and Nett Warrior

Nett Warrior is an integrated, dismounted Soldier situational awareness system for use by leaders during combat operations. The Rifleman Radio, AN/PRC-154A, is a component of the Nett Warrior system. Nett Warrior is designed to facilitate command, control, and sharing of battlefield information and to integrate each leader into the digitized battlefield. The Army intends to use Nett Warrior to provide mission command and position location information down to the team leader level. In the Nett Warrior Limited User Test during Network Integration Evaluation 13.2, the AN/PRC-154A classified radio did not support the mission of the test unit.

The radio provided inconsistent digital communications, and the majority of the unit leaders indicated that voice quality was degraded beyond 500 meters. The radio experienced delays in re-joining the network, and experienced problems with battery over-heating and rapid battery depletion. If the problems with the radio are not fixed, the effectiveness of the Nett Warrior to provide situational awareness will be severely limited, and future operational effectiveness and operational suitability assessments of the radio will be adversely affected.

Integrated Defensive Electronic Countermeasures (IDECM)

The IDECM system is a radio frequency, self-protection electronic countermeasure suite on F/A-18 aircraft. The system is comprised of onboard and off-board components. The onboard components receive and process radar signals and can employ onboard and/or off-board jamming components in response to an identified threat. IDECM Block 4 program completed an operational assessment in FY13. The operational assessment was originally planned to consist of flight testing and three laboratory tests with hardware-in-the-loop. One of those laboratory tests was postponed because the system was insufficiently mature, and a second was rescheduled because of a higher priority program. Partially because the system was immature at the time of the test, and partially by design, very little developmental flight testing had occurred prior to the operational assessment.

As a result of poor record-keeping, some aspects of suitability could not be assessed for the analysis of the operational assessment; however, sufficient information was available to determine that reliability was extremely low. The primary contributors to these failures were system instability and resets. While the Navy in general was aware of the problems – its system anomaly database had over 100 open anomalies at the time of the operational assessment – the Service had focused on tracking each mode of failure rather than their frequency. If reliability does not significantly improve prior to accomplishing FOT&E, it is likely the system will be assessed as both not effective and not suitable because IDECM's poor reliability will preclude effective use in combat.

In addition to these documented shortfalls, the Navy must collect complete and comprehensive suitability data to enable the assessment of availability, maintainability, and built-in test. The Navy needs to improve interoperability between IDECM Block 4 and the radar warning receiver and fire control radar.

Since the operational assessment, the prime contractor has released several updates to the system software and further laboratory and flight testing have been accomplished in preparation for the FOT&E, currently scheduled for early CY14. It is not yet clear whether these efforts have been sufficient to address all the shortfalls noted above.

Integrated Electronic Health Record (iEHR)

The DoD and Veterans Affairs (VA) will use the iEHR program to implement an electronic health record that both organizations can use to meet the healthcare needs of their beneficiaries and the clinicians providing healthcare. Increment 1 will provide a Single Sign-on (SSO) capability for multiple applications via the users' Common Access Card, and a Context Management (CM) capability to allow fast user switching between applications while keeping the patient and associated clinical information in context. The Interagency Program Office designed and developed SSO-CM using the capabilities of COTS products. The U.S. Army Medical Department Board planned to conduct an SSO-CM operational assessment in November 2012, but testing was delayed due to system defects and site configuration problems.

- Four developmental test events identified a total of 32 defects: 14 in the initial test, 7 in the first System Integration Test (SIT-1), 7 in SIT-2, and 4 in SIT-3. At the end of SIT-3, 13 defects remained open. At the completion of SIT-3, the program manager further delayed the operational assessment.
- DOT&E rejected the operational assessment plan because it did not demonstrate that the SSO-CM systems would work with, and not interfere with, the Interagency Program Office's primary deliverables, which are the DoD and VA iEHR accelerators.
- The Program Executive Officer for the DoD Healthcare Management Systems should work with DOT&E to develop an adequate plan for an operational assessment of the SSO-CM functionality and the impact on Health Data Sharing and Interoperability.

Joint Warning and Reporting Network (JWARN)

JWARN is a chemical, biological, radiological, and nuclear (CBRN) warning and reporting software application intended to provide men and women in combat with an integrated analysis and response capability to minimize the effects of hostile CBRN attacks. The Army Test and Evaluation Command conducted the JWARN Increment 1 Modernization operational assessment in a laboratory setting at the Central Technical Support Facility (CTSF) at Fort Hood, Texas, from July 25–31, 2013. During the operational assessment, the immaturity of Army Command Web and network instability diminished the capability of JWARN web application operators to provide timely warnings to units at risk. Since there is no other developmental test venue for the Army network other than the CTSF, these problems could not be predicted or knowable by the program manager prior to the operational assessment. The Army should schedule a developmental test event in the CTSF with a goal of achieving a stable network prior to operational testing.

Configuration problems with the command and control infrastructure virtual machine software, which supports lower-level tactical messaging, prevented Variable Message Format warning messages from being exchanged between battalions using JWARN and company units using Joint Battle Command – Platform (JBC-P) in both unicast and multicast modes. This limitation precluded an end-to-end evaluation of battalion-to-company or company-to-battalion hazard warning using JWARN.

LHA-6 Amphibious Assault Ship

LHA-6 is a large-deck amphibious ship designed to support a notional mix of fixed- and rotary-wing aircraft. Completed testing of the Ship Self-Defense System (SSDS) Mk 2-based combat system on the CVN-68 class carrier indicates that it is not likely that LHA-6's nearly equivalent SSDS Mk 2-based combat system will meet the ship's Probability of Raid Annihilation requirement against all classes of anti-ship cruise missiles (ASCMs). Additionally, LFT&E analysis completed to date identified potential problems in susceptibility and vulnerability that would likely result in the LHA-6 being unable to maintain or recover mission capability following a hit by some threat weapons.

Littoral Combat Ship (LCS)
(Includes Seaframes and Mine-Countermeasures Mission Package with the Remote Minehunting System (RMS) and Airborne Mine Neutralization System (AMNS))

The LCS is the Navy's newly-designed surface ship intended to accommodate a variety of individual warfare systems (mission modules) assembled and integrated into interchangeable mission packages. Testing conducted in FY13 and analysis of data from FY12 testing continued to identify deficiencies in the LCS seaframes and essential mission systems:

- Analysis of equipment casualty reports filed by LCS 1, LCS 2, and LCS 3 showed that the reliability of both seaframe variants has been degraded by frequent critical system failures during early operations and testing. Failures of the LCS 1 seaframe's diesel-powered generators, air compressors, and propulsion drive train components have degraded the seaframe's reliability during developmental testing and early operations. The operational reliability of the LCS 2 variant's seaframe has been degraded by equipment failures, including problems with operator consoles, power generation equipment, components of the Total Ship Computing Environment and the ship's internal networks, propulsion drive train components, communications systems, and mission package support systems.
- The Remote Multi-Mission Vehicle (RMMV), which is a component of the Mine Countermeasures (MCM) mission package, has a history of poor reliability that if not corrected would affect the assessment of LCS's operational suitability in conducting MCM operations. Following a second phase of vehicle improvements and reliability growth testing, the Navy reported that RMMV reliability was meeting Navy requirements. However, DOT&E's review showed that the Navy's assessment excluded some critical failures and was based on failure definitions and scoring criteria that were inconsistent with those used during the program's Nunn-McCurdy review; the estimates also do not reflect the expected reliability in more operationally realistic mission scenarios where vehicle usage is more stressed. An upcoming shore-based operational assessment will provide another opportunity to evaluate the system's reliability.
- The MCM mission package performance during developmental testing has been degraded by immature mission systems, low sensor detection performance in some operational conditions, high false alarm rates, unproven tactics, and low operator proficiency.
- The Navy completed developmental testing to assess Multi-Vehicle Communications System (MVCS) upgrades and improvements to the launch, handling, and recovery systems for the RMMV. Following testing, the Navy reported that additional efforts are required to retire risks associated with RMMV launch and recovery. Sailors also reported that communications between an RMMV equipped with MVCS upgrades and LCS 2 were unreliable throughout the test.
- DOT&E completed analysis of data from an FY12 shore-based operational assessment of the MH-60S helicopter and Airborne Laser Mine Detection System (ALMDS) and found that ALMDS detection depth does not meet the Navy's requirement. This deficiency will make it necessary to extend the detection envelope of the AN/AQS-20A Sonar Mine Detecting Set to restore the desired overlap with the demonstrated ALMDS envelope. The Navy conducted additional developmental testing of the AN/AQS-20A using a surface craft to tow the sensor and expert operators to evaluate the AN/AQS-20A capability to detect and classify near-surface mines during post-mission analysis. While this has the potential to ameliorate the deficiency, the Navy has not yet completed an operational test of this capability with the RMMV, controlled by fleet operators, towing the sensor and fleet personnel performing the post-mission analysis of the sonar data. Additional testing will be required in other environments as well to fully characterize the capability.

- The Navy completed developmental testing to evaluate the performance of the Airborne Mine Neutralization System (AMNS) when it is operated in high current and reported problems with compass corrections and fiber-optic communications losses. These failures have the potential of making AMNS not effective since even minor currents are expected in many operational environments. Additional testing is needed to determine the maximum current in which the system is still operable, and determine the operational impact of the performance deficiency.
- The Navy's Quick Reaction Assessment uncovered classified deficiencies in LCS 1's capability to protect the security of information.

M109 Family of Vehicles (FoV) Paladin Integrated Management (PIM)

The PIM program is a sustainability and survivability upgrade of the currently fielded Paladin M109A6 self-propelled howitzer and companion M992A2 resupply vehicle. The Army conducted the PIM Limited User Test (LUT) in November 2012 to support the program's Milestone C decision.

The PIM LUT Pilot Test and collective live firing events revealed issues with the M82 primer when firing M232A1 Modular Artillery Charge System (MACS) Charge 5 propellant. The M82 primer deforms and jams in the cannon firing mechanism due to higher breech pressures when firing MACS Charge 5 propellant. This problem had been observed in developmental testing, but the scope of the problem and operational implications were not widely understood until the LUT Pilot Test. There were no plans to address the issue. Problems encountered during training and the pilot test prompted replacement of MACS 5 with another propellant during the LUT.

The Program Executive Officer, Ground Combat Systems and Program Executive Officer, Ammunition established a special research team to identify solution options involving modification of the propellant, redesign of the breech and firing mechanisms, development of alternative ignition systems, and/or restriction of the use of MACS propellant to no more than four increments. If the issue is not resolved before the FY16 IOT&E, it is unlikely the test unit will be responsive when firing missions with MACS 5 propellant.

Next Generation Diagnostic System (NGDS)

The NGDS is a U.S. Food and Drug Administration (FDA)-cleared reusable, portable biological pathogen diagnostic and identification system capable of rapidly analyzing clinical and environmental samples. The U.S. Army Medical Department Board conducted an early operational assessment of three candidate NGDS systems in 3QFY13. The three candidates were commercial off-the-shelf medical diagnostic devices.

One of the vendor systems encountered major reliability problems during testing, resulting in systems having to be replaced. Other vendor systems experienced minor hardware problems, such as loose wiring connections, that could also affect suitability. One vendor system used complex operating procedures that at times proved difficult for operators to follow correctly and often resulted in invalid results. Ensuring protocols are clear and operators are appropriately trained to operate the system will be key as the program moves to MOT&E.

Public Key Infrastructure (PKI) Increment 2

PKI Increment 2 provides authenticated identity management via password-protected SIPRNet tokens to enable DoD members and others to access the SIPRNet securely and to encrypt and digitally sign e-mail. The program continues to add capability through spiral development, and these spirals will undergo testing in the future. Limited and poorly designed developmental testing was directly attributable to the problems observed in previous operational testing. While the Program Management Office has made some initial attempts to correct the configuration management issues, adequate Configuration Control Board structure and overall repeatable processes for defect identification and resolution still do not exist.

Unless the program can fix the configuration management processes for prioritizing needed capabilities and improve configuration control processes for ensuring deployments can be sustained without impacting availability and reliability, DOT&E may once again assess the PKI as not operationally effective and suitable for current and future Spirals.

Q-53 Counterfire Target Acquisition Radar System

The Q-53 radar is designed to detect, classify, and locate projectiles fired from mortar, artillery, and rocket systems using a 90-degree or continuous 360-degree sector search. Early developmental testing indicates the Q-53's probability of detection and location accuracy against volley-fired weapons is worse than the performance demonstrated against single-fired weapons. Volley-fire is the technique of firing multiple weapons from the same location at a single target. Although the Army has not identified a volley-fire requirement for the Q-53 radar, volley-fire is a standard threat technique and will be used as a threat tactic in the FY14 Q-53 IOT&E.

Developmental testing was conducted under conditions that do not match all expected threat employment profiles; therefore, IOT&E results have the potential of being different than observed in developmental testing. If corrections are not made and the IOT&E results reveal the same performance deficiencies observed in developmental testing, then DOT&E's assessment of operational effectiveness could be affected.

RQ-4B Global Hawk High-Altitude Long-Endurance Unmanned Aerial System (UAS)

The RQ-4 Global Hawk is a remotely-piloted, high-altitude, long-endurance airborne Intelligence, Surveillance, and Reconnaissance system that includes the Global Hawk unmanned air vehicle, various intelligence and communications relay mission payloads, and supporting command and control ground stations. In March 2013, the Air Force conducted an Operational Utility Evaluation (OUE) of the RQ-4B Global Hawk UAS. The OUE discovered previously unidentified shortfalls in synthetic aperture radar stationary target imagery capabilities. These capabilities do not currently meet established operational requirement thresholds for image resolution. Multi-Platform Radar Technology Insertion Program (MP-RTIP) operator displays and control interfaces are also immature, which significantly increases operator workload during target-intense operations.

During OUE missions, frequent MP-RTIP sensor faults required sensor operators to halt intelligence collection operations to reset or restart the system. Resulting sensor downtime reduced on-station intelligence collection time by 23 percent. Additionally, contactor maintenance and supply support was required to compensate for immature system-level reliability, maintenance training, documentation, and logistics support systems.

Surface Ship Torpedo Defense (SSTD) System: Torpedo Warning System (TWS) and Countermeasure Anti-torpedo Torpedo (CAT)

The SSTD is a system-of-systems that includes two new sub-programs: the TWS program (an Acquisition Category III program) and CAT (not an acquisition program until FY16). TWS is being built as an early warning system to alert on and localize incoming threat torpedoes. While TWS was designed to employ both active and passive sonar to detect incoming threat torpedoes, hardware reliability failures forced the Navy to delay development of the active component. During early testing from March through August 2013, using the purely passive detection approach, the Navy observed that TWS was subject to false alarms and poor detection performance.

The Navy temporarily addressed this problem by assigning a civilian contractor acoustics specialist to monitor and report indications of threat detections using displays not normally available to the ship's crew. Contractors provided this service during the November 2013 Quick Reaction Assessment aboard USS *George H. W. Bush* (CVN-77), and are expected to deploy with the ship in FY14.

PROGRESS UPDATES ON DISCOVERIES REPORTED IN THE FY12 DOT&E ANNUAL REPORT

FY12 Discoveries in IOT&E that should have been Resolved prior to Operational Test

In FY12, I identified 17 systems that had significant issues in IOT&E that should have been discovered and resolved prior to commencement of operational testing. Two of the 17 programs were cancelled: Mine Resistant Ambush Protected (MRAP) Dash Ambulance and MRAP Caiman Multi-Terrain Vehicle (CMTV). For the ALR-69A Radar Warning Receiver, the Program Office has implemented a fix for the program, but operational testing will not be completed until a future aircraft program integrates the system. The Standard Missile-6 (SM-6) Program Office is studying potential fixes. The following updates the status of the remaining 13 systems.

Fixes Implemented and Demonstrated in FOT&E

- Key Management Infrastructure (KMI) Increment 2
- Mission Planning System (MPS)/Joint Mission Planning System – Air Force (JMPS-AF)

Fixes Implemented but New Issues Discovered

- Distributed Common Ground System – Army (DCGS-A)

Fixes Implemented; Currently in OT or Planning Additional OT

- AGM-88E Advanced Anti-Radiation Guided Missile (AARGM) Program
- Battle Control System – Fixed (BCS-F)
- E-2D Advanced Hawkeye
- E-3 Airborne Warning and Control System (AWACS) Block 40/45 Upgrade
- Handheld, Manpack, and Small Form Fit (HMS) Manpack Radio
- Handheld, Manpack, and Small Form Fit (HMS) Rifleman Radio
- Miniature Air Launched Decoy (MALD) and MALD – Jammer (MALD-J)
- MV-22
- SSN 774 *Virginia* Class Submarine
- Warfighter Information Network-Tactical (WIN-T)

No Fixes Planned

- None

FY12 Discoveries in Early Testing that should be Corrected prior to IOT&E

In FY12, I identified six systems that had significant issues in early testing that should be corrected before IOT&E. The following provides an update on the progress those systems made in implementing fixes to those problems. Five of the six programs have or are implementing corrective actions that will be tested and assessed in either LFT&E or OT&E.

Fixes Implemented and Demonstrated in OT or LFT&E

- Bradley Engineering Change Proposal (ECP)

Fixes Implemented, but Effect is Unknown; Currently in OT or Planning OT

- F-15E Radar Modernization Program (RMP)
- Joint Standoff Weapon (JSOW) C-1
- Littoral Combat Ship (LCS)
- Patriot Advanced Capability-3 (PAC-3)

Some Fixes Implemented; Testing Constrained Pending Future Acquisition Decisions

- None

No Fixes Planned or Plans not Determined

- Multi-Static Active Coherent (MAC) System