Problem Discovery Affecting OT&E

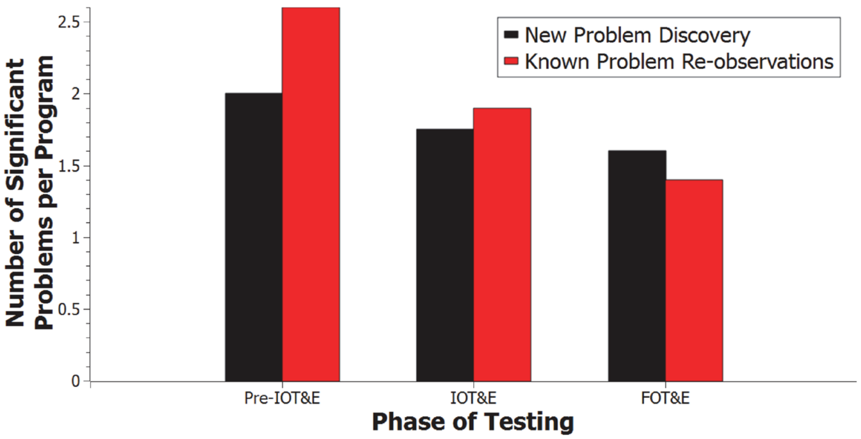
Background

In 2011, Congress expressed concern that acquisition programs are discovering problems during operational testing (OT) that: (1) should have been discovered in developmental testing (DT), and (2) should have been corrected prior to OT. In response to this congressional concern, I added this section to my Annual Report, as a means to survey across all DOT&E oversight programs covered in this report, the extent of late problem discovery and to identify known problems that jeopardize a system's successful performance in upcoming OT.

This is the fourth time this section has been included in my Annual Report, and this iteration presents a more in-depth review of the programs included in this report. Last year, this section consisted of short case studies that discussed problems that were identified during OT or DT. This year's section includes data that break down into several relevant categories the effectiveness, suitability, and cybersecurity problems that were either observed during OT or that jeopardize a system's successful performance in an upcoming OT event (i.e., if known problems are not fixed, a finding of not effective, not suitable, and/or not survivable could occur).[1] The results presented in this section continue to show that OT is necessary, and that we continue to find significant and substantial problems during OT that were either not previously observed or could not be observed in DT. Also, as documented in this section, OT continues to identify problems that were previously discovered but not fixed.

Overview of Problem Discovery in OT

Figure 1 below shows a breakdown of the number of significant problems (per program and by the phase of testing) and where the problems were newly discovered or already known. As expected, the rate of new problem discovery in early OT that occurs prior to Initial Operational Test and Evaluation (IOT&E) (pre-OT conducted to inform acquisition and/or early fielding decisions) is higher than the rate of problem discovery in both IOT&E and Follow-on Operational Test and Evaluation (FOT&E). This is a desirable trend because the earlier a problem is discovered, the easier it is to fix it, and it is consistent with DOT&E's initiative for early involvement in test programs. The ratio for new problem discovery (black bars) is the highest (two "significant problems" per program) for early OT. Significant problems are those that would have a negative impact on DOT&E's assessment of effectiveness, suitability, or cybersecurity.



		Pre-IOT&E	IOT&E	FOT&E
New Problem Discovery	Total Problems	10	14	8
	Total Programs	5	8	5
	Ratio	**2**	**1.75**	**1.6**
Known Problem Re-observations	Total Problems	13	19	14
	Total Programs	5	10	10
	Ratio	**2.6**	**1.9**	**1.4**

FIGURE 1. PROBLEM DISCOVERY RATIOS IN PRE-IOT&E, IOT&E, AND FOT&E

For re-observations of known problems, the rate is also higher in early OT, and is higher overall than for new problem discovery (red bars). This result indicates that while early OT is effective in demonstrating the operational impact of known problems prior to IOT&E, OT is observing more known problems in all phases of testing compared to new problem discovery.[2] In cases where known issues prior to OT are significant (indicating a lack of system maturity), DOT&E has suggested not doing the OT because the resources expended conducting the test would not be worth the little or irrelevant information gained from an OT at that time. This year, DOT&E suggested foregoing planned OT events for the F-35 Joint Strike Fighter (JSF), Air and Missile Defense Radar (AMDR), and Remote Minehunting System (RMS) because of several known performance issues.

[1] Cybersecurity problems are evaluated through OT and are considered in DOT&E's survivability assessments. Survivability problems discovered through Live Fire Test and Evaluation are not included in this discussion of OT.

[2] For pre-IOT&E testing, observing known problems is not a major issue because the program still has time to correct them prior to IOT&E; this fact underscores the importance of conducting an operational assessment prior to the Milestone C or Low-Rate Initial Production decision.

Other trends are:

- About one-third of the programs that underwent OT during FY14 did so successfully; that is, they did not uncover problems significant enough to negatively affect my assessment of operational effectiveness, suitability, or cybersecurity.
- For new problem discovery, about half of the effectiveness problems found in OT were not discoverable in DT because the operationally realistic conditions required in OT were needed to discover the problem (i.e., testing under realistic combat conditions by typical military users).
- More than two-thirds of the programs that commenced IOT&E or FOT&E in FY14 with known suitability problems implemented (and in many cases tested) fixes to these problems prior to the OT. This is an area where DOT&E's initiatives on reliability growth are having a positive effect.
- Looking to the future and consistent with the first bullet above, about one-third of the programs with upcoming OT events in the next three years have not yet exhibited any effectiveness, suitability, or cybersecurity problems significant enough to jeopardize successful performance in OT. (However, we know that about half of the new problems observed during OT cannot be observed in early testing because of the need for operationally realistic environments.)
- Thirty percent of the programs undergoing OT in FY14 only re-observed previous known problems during OT; no additional significant problems were found.
- Pre-IOT&E test events are more likely to be delayed to allow time to correct problems compared to delaying either IOT&Es or FOT&Es.
- A majority of programs (10/13) that observed problems during IOT&E re-observed at least one problem that was known prior to the IOT&E.
- Five of the nine programs that re-observed known effectiveness issues during an IOT&E or FOT&E in FY14 did not identify fixes to address these problems prior to operational test.

EVALUATION OF PROBLEM DISCOVERY

Programs with an FY14 OT

I surveyed 81 programs that either underwent OT in FY14 or will undergo OT within the next three years (some programs fall into both categories), and are reported on in this Annual Report.[3] The results presented in this section, including those in Figure 1 above, focus on these programs. I classified the programs that underwent OT into one of three main categories: (1) successful performance in OT; (2) new performance problems discovered; and (3) known performance problems re-observed. The more detailed review conducted this year also allowed me to categorize individual problems and Program Office responses to these problems, whereas last year's report only categorized problems at the program level. Otherwise, the categories used in this year's report are similar to those used in previous years. These categories are described in Table 1. For problems that were discovered during OT, I assess whether these problems affected effectiveness, suitability, or cybersecurity and whether they reasonably could have been discovered prior to the OT event.

TABLE 1. PROBLEM DISCOVERY CATEGORIES FOR PROGRAMS WITH AN OT EVENT OR A DOT&E REPORT IN THE LAST YEAR

Category		Description
Successful OT	Successful OT without delays	No significant problems were discovered during OT that would negatively affect DOT&E's assessment of operational effectiveness, suitability, and/or cybersecurity.
	Successful, but delayed, OT	Problem(s) were discovered that delayed entry into OT so they could be addressed, thus contributing to a successful OT outcome.
New problem discovery		Problem(s) significant enough to negatively affect DOT&E's assessment of operational effectiveness, suitability, and/or cybersecurity were discovered for the first time in OT.
Known problem re-observations		Problem(s) were observed in OT that were known prior to entering OT and significant enough to negatively affect DOT&E's assessment of operational effectiveness, suitability, and/or cybersecurity.

Programs with an upcoming OT

For programs that are scheduled to undergo an OT event within the next three years, I identified those that have not uncovered any problems that jeopardize a system's successful performance in upcoming OT events, and those with problems significant enough that, if uncorrected, would negatively affect my assessment of operational effectiveness, suitability, and/or cybersecurity. I classify these programs into one of three categories: (1) no problems for upcoming OT; (2) problems delayed upcoming OT; and (3) problems have not delayed upcoming OT. They are described in more detail in Table 2.

[3] The original congressional request specified programs scheduled to commence operational testing within the next two years. I expanded that window to three years to include programs that delayed their entry into OT so they could fix known problems before commencing OT.

TABLE 2. PROBLEM DISCOVERY IN CATEGORIES FOR PROGRAMS IN THIS ANNUAL REPORT SCHEDULED TO UNDERGO OT WITHIN THE NEXT 3 YEARS

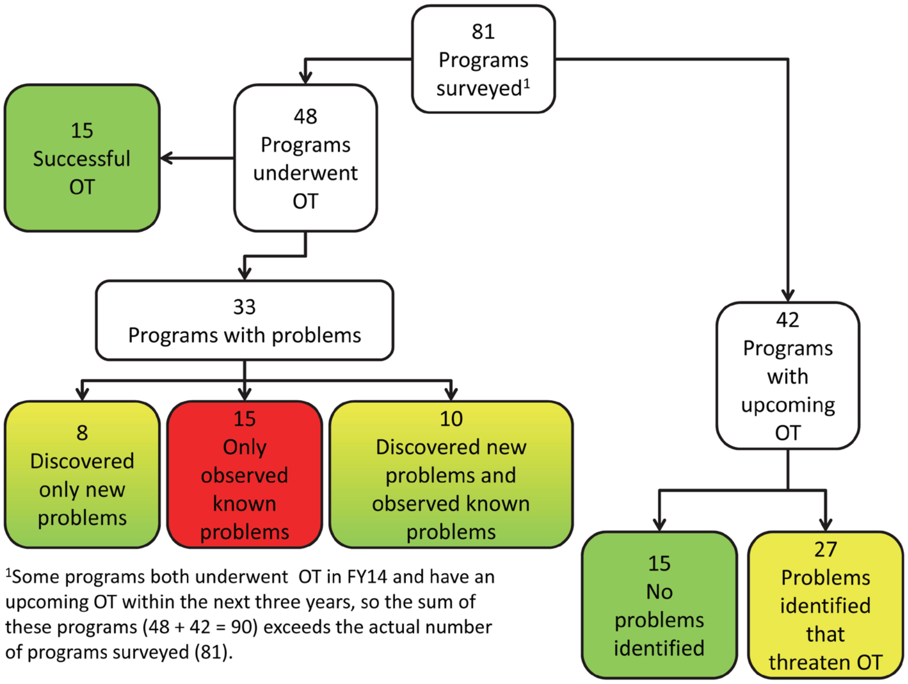
Category	Description
No problems for upcoming OT	The program has not exhibited any problems that would negatively affect DOT&E's assessment of operational effectiveness, suitability, and/or cybersecurity for the upcoming OT.[1]
Problems delayed upcoming OT	Problems exist that, if not corrected, would negatively affect DOT&E's assessment of operational effectiveness, suitability, and/or cybersecurity for the upcoming OT. OT has been delayed so these problems can be addressed prior to commencing OT.
Problems have not delayed upcoming OT	Problems exist that, if not corrected, would negatively affect DOT&E's assessment of operational effectiveness, suitability, and/or cybersecurity for the upcoming OT, but OT has not (yet) been delayed to address these problems.

1. Such programs could be examples where the program development process, including DT and OT, appears to be moving along well. Alternatively, the testing to date might not have been sufficiently stressing to surface any problems.

SUMMARY OF RESULTS

For problem discovery in FY14, I found a mixture of positive trends and areas that may need improvement. The results are shown in Figure 2. Blocks that are colored green signify positive trends, while the block in red signifies areas that need improvement. The yellow block represents an outcome that is in-between or neutral. The two blocks with a yellow/green color gradient are a combination of mixed results. The outcomes shown in Figure 2 are discussed in more detail in the sections that follow.

About one-third of the programs (15 of 48) that underwent OT during FY14 did so successfully; that is, they did not uncover problems significant enough to negatively affect DOT&E's assessment of operational effectiveness, suitability, and cybersecurity.[4] About two-thirds of the programs (33 of 48) that underwent OT during FY14 encountered problems that negatively affected DOT&E's assessment of their operational effectiveness, suitability, and/or cybersecurity. These problems were either new problems discovered in the OT event, or were re-observations of known problems. Of these 33 programs, 8 programs discovered only new problems, 15 only re-observed known problems, and 10 both discovered new problems and re-observed known problems.


81 Programs surveyed[1]
48 Programs underwent OT
15 Successful OT
33 Programs with problems
8 Discovered only new problems
15 Only observed known problems
10 Discovered new problems and observed known problems
42 Programs with upcoming OT
15 No problems identified
27 Problems identified that threaten OT

[1]Some programs both underwent OT in FY14 and have an upcoming OT within the next three years, so the sum of these programs (48 + 42 = 90) exceeds the actual number of programs surveyed (81).

FIGURE 2. PROBLEM DISCOVERY IN OT RESULTS MEASURED BY PROGRAM COUNTS

For programs with upcoming OT events in the next three years, I determined that slightly more than one-third (15 of 42) of the programs currently exhibit no effectiveness, suitability, or cybersecurity problems significant enough to jeopardize their successful performance in upcoming OT, which is to say that no problems have yet been found that, if not corrected, would negatively affect my assessment of operational effectiveness, suitability, or cybersecurity. Of the remaining two-thirds of programs (27), I identified 23 that have effectiveness, suitability, and/or cybersecurity problems that, if not corrected, could negatively affect my operational assessments. The remaining four programs have items that potentially jeopardize successful performance in OT, but these relate more to schedule or process as opposed to effectiveness, suitability, or cybersecurity. Examples include test schedules in Test and Evaluation Master Plans that are not executable; reliance on other programs that are facing development challenges; and failed, cancelled, or delayed DTs that jeopardize successful performance in OT.

[4] Note that even in these cases, OT provides recommendations or potential improvements to improve system performance for the warfighter.

PROGRAMS THAT CONDUCTED OT IN FY14

For programs that conducted an OT event in FY14, my analysis consists of the following:

- Program-Level – Analysis of the number of programs that fall into each of the categories listed in Table 1, broken down by the types of problems found by each program (effectiveness, suitability, or cybersecurity), and across the three phases of OT (pre-IOT&E, IOT&E, or FOT&E)
- Problem-Level – For programs that experienced significant problems during their OT, analysis of the number of problems found by all the programs, broken down by the types of problems and across the three phases of OT
- Responses to Known Problems – For programs that re-observed known problems, analysis of the number of programs that identified, implemented, and in many cases tested fixes to these problems prior to the OT

Program-Level

Analysis of the number of programs in which problems have been observed is necessary to assess the scale of problem discovery in OT. Note that some programs, as shown in Figure 2, observed both new and known problems, so these can contribute to program counts for both types of problems. The types of problems observed can be related to effectiveness, suitability, or cybersecurity; programs can observe multiple types of problems during their OT.

TABLE 3. FY14 OT RESULTS BASED ON NUMBER OF PROGRAMS

Category		Number of Programs[1]	Number of Programs by Type of Problem[2]		Number of Programs by Phase of Testing		
					Pre-IOT&E	IOT&E	FOT&E
Successful OT	Successful OT without delays	11			1	6	4
	Successful, but delayed, OT	4			1	2	1
New problem discovery		18	Effectiveness	11	4	4	4
			Suitability	9	2	4	2
			Cybersecurity	5	0	4	1
				Total	**5**	**8**	**5**
Known problem re-observations		25	Effectiveness	13	4	4	5
			Suitability	14	3	7	4
			Cybersecurity	4	1	1	2
				Total	**5**	**10**	**10**

1. Forty-eight programs underwent an OT in FY14. Fifteen had successful OTs (11 + 4), and 33 uncovered problems. The number of programs that experienced new problem discovery or re-observed known problems adds up to more than 33 in the table because some programs experienced both new problem discovery and known problem re-observations, thus contributing to both counts.
2. The count of programs that discovered/observed problems during testing exceeds the totals in the "Numbers of programs" column because some programs discovered multiple types of problems.

Of the 15 programs that successfully completed OT in FY14, 4 programs delayed their OT in order to fix performance problems prior to the OT. Table 3 shows these broken down by program discovery category, type of problem observed, and phase of testing. A list of the programs that fall under these categories in Table 3 is included at the end of this section.

More than one-third of the programs (18 of 48) undergoing OT in FY14 discovered new problems. When aggregated, these programs were divided nearly evenly between observing effectiveness (11) and suitability (9) problems, along with a few cybersecurity (5) problems. When broken down by the phase of testing, the ratio of programs discovering new suitability problems to the discovery of new effectiveness problems is lowest in pre-IOT&E testing. This suggests that suitability issues manifest themselves later in the testing lifecycle, but the sample size is too small to definitively state whether this is a trend, or simply random chance.

The proportion of programs re-observing known problems in OT remains high.[5] Table 3 shows that 25 of the 48 programs that underwent OT in FY14 encountered known, significant problems. Fifteen of the 25 programs (see Figure 2) encountered only known problems during their OT, while the other 10 also discovered new problems.

Problem-Level

Analysis based on the number of problems observed during OT can help characterize the completeness of testing prior to OT. Such results are shown in Table 4. Table 4 is similar to Table 3, except it indicates the number of problems encountered during OT instead of the number of programs encountering problems.[6] Some of the new problems observed were not discoverable prior to commencing OT while others were. For new problem discovery, about half of the effectiveness problems found (8 of 17) were not discoverable prior to the OT. Such problems generally require the operationally realistic (or "test-as-you-fight") environment that is the hallmark of OT in order to be discovered. For new suitability problems discovered, this drops to one-third (3 of 9).

[5] This result is shown in Figure 2, but is not directly observable in Table 3 because some programs observed both new problems and known problems.

[6] The problems referred to here are the number of "significant" problems, not all problems. Recall that significant problems are those that would negatively affect my assessment of operational effectiveness, suitability, or cybersecurity.

The proportion of the number of known problems encountered in OT to new problem discovery may be an area where improvement is possible. Table 4 shows that 42 known problems were re-observed in OT, compared to 27 new problems being discovered (of which 11 were not discoverable in pre-OT testing).

TABLE 4. FY14 OT RESULTS BASED ON NUMBER OF PROBLEMS

Category	Number of Problems by Type[1]		Number of Problems by Phase of Testing[1]		
			Pre-IOT&E	IOT&E	FOT&E
New problem discovery	Effectiveness	18 (8)	8 (4)	5 (2)	5 (2)
	Suitability	9 (3)	2 (1)	5 (1)	2(1)
	Total	**27 (11)**	**10 (5)**	**10 (3)**	**7 (3)**
Known problem re-observations	Effectiveness	19	7	7	5
	Suitability	23	5	11	7
	Total	**42**	**12**	**18**	**12**

1. Numbers in parentheses are the number of problems that were not discoverable prior to OT. For example, in IOT&E, five new effectiveness problems were identified in FY14 across all programs undergoing IOT&E. Of these, two were not discoverable prior to IOT&E.

Responses to Known Problems

As shown in Table 5, about half of the programs that commenced IOT&E or FOT&E in FY14 with one or more known effectiveness problems did not identify fixes to address these problems prior to OT (both 3 of 5 in IOT&E and 2 of 5 in FOT&E).[7] The situation is better, however, when the program commenced OT with known suitability problems. In this case, 5 of 7 programs in IOT&E, and 3 of 4 in FOT&E implemented (and in many cases tested) fixes to these problems prior to OT. Note, however, that by breaking down the results thus far, the number of programs in each category (fix not identified or fix implemented (and in many cases tested)) is small. The sample size is too small to definitively state whether this is a trend or simply random.

TABLE 5. ACTIONS TAKEN TO MITIGATE KNOWN PROBLEMS PRIOR TO ENTERING OT

Category	Number of Programs	Number of Programs by Type of Problem		How was the problem addressed prior to OT?			
				IOT&E		FOT&E	
				Fix Not Identified[1]	Fix Implemented (tested)[2]	Fix Not Identified[1]	Fix Implemented (tested)[2]
Known problem re-observations (IOT&E or FOT&E)	20	Effectiveness	9	3	2 (1)	1	3(1)
		Suitability	11	2	5 (3)	1	3 (3)
		Cybersecurity	3	1	1 (1)	1	1 (0)

1. Number of programs that had at least one problem for which no fix was identified.
2. Number of programs that had at least one problem for which a fix was implemented (tested).

Specific Programs that had OT in FY14

Successful OT

Fifteen of the 48 programs that underwent OT in FY14 experienced successful performance in OT. The majority of these (11 of 15) that successfully completed an OT event did so without having to delay OT. These programs are listed below in Table 6 and are examples of a successful development process, including DT and OT. Additional details on any of these programs can be found in the program-specific entries in the main body of this report.

TABLE 6. PROGRAMS THAT HAD SUCCESSFUL OT IN FY14

Successful OT (No Delays)	Successful OT (with Delays)
AH-64E	Cobra King (formerly Cobra Judy Replacement)
AIM-120 Advanced Medium-Range Air-to-Air Missile (AMRAAM)[1]	Excalibur Increment 1B M982E1
F/A-18E/F Super Hornet and EA-18G Growler	Global Command and Control System – Joint (GCCS-J)[2]
Global Command and Control System – Maritime (GCCS-M)	MQ-9 Reaper Armed Unmanned Aircraft System (UAS)
Integrated Personnel and Pay System – Army (IPPS-A)	
Joint Tactical Network (JTN)	
MH-60R Multi-Mission Helicopter	
MH-60S Multi-Mission Combat Support Helicopter	
Nett Warrior	
RQ-7BV2 Shadow Tactical Unmanned Aircraft System (TUAS)	
Department of Defense (DOD) Teleport	

1. This was the Electronic Protection Improvement Program (EPIP), a software upgrade to a previously-fielded missile.
2. Emerging results from the OT have not been completely analyzed.

[7] Programs that had known problems prior to commencing pre-IOT&E testing are not counted here because in most cases there may be sufficient time prior to starting IOT&E to address these problems. Furthermore, not all required system capabilities might be present for pre-IOT&E events.

New problem discovery

Of the 48 programs that underwent OT in FY14, 18 discovered problems that had not been seen before. There are a variety of reasons why some problems are not observed prior to OT. In some cases, problems can be uncovered only by testing under the operationally realistic conditions that characterize formal OT. The sooner these problems are discovered, the better. Hence, finding these problems in the pre-IOT&E phase of OT, such as an operational assessment prior to the Milestone C or Low-Rate Initial Production decision, is highly desirable. Other problems are discovered for the first time that could have been found and addressed during testing prior to OT, such as dedicated Developmental Test and Evaluation (DT&E). Table 7 gives a list of the programs that had new problem discovery and indicates whether these problems were discoverable earlier. Note that many programs experience both new problem discovery in OT and re-observation of known problems. These programs are highlighted in grey in Table 7. Additional details on any of these programs can be found in the program-specific entries in the main body of this report.

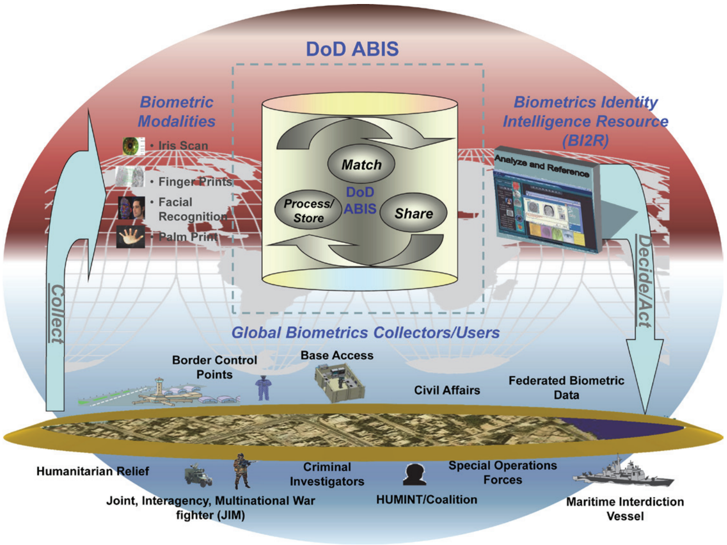
TABLE 7. PROGRAMS THAT HAD NEW PROBLEM DISCOVERY IN OT IN FY14

Program	OT Event Type	Service	Discoverable Prior to OT			Not Discoverable Prior to OT		
			Effectiveness	Suitability	Cybersecurity	Effectiveness	Suitability	Cybersecurity
AN/PRC-117G	Pre-IOT&E	Army	×					
DOD Automated Biometric Identification System (ABIS)	Pre-IOT&E	Army	×			×		
Defense Medical Information Exchange (DMIX)	Pre-IOT&E	Joint	×	×				
Joint Light Tactical Vehicle (JLTV) Family of Vehicles (FoV)	Pre-IOT&E	Joint					×	
Surface Ship Torpedo Defense (SSTD) System: Torpedo Warning System (TWS) and Countermeasure Anti-Torpedo (CAT)	Pre-IOT&E	Navy	×					
Ballistic Missile Defense System (BMDS)	IOT&E	MDA	×					
Distributed Common Ground System – Marine Corps (DCGS-MC)	IOT&E	Marine Corps		×				×
F-15E Radar Modernization Program (RMP)	IOT&E	Air Force		×				
Joint Battle Command – Platform (JBC-P)	IOT&E	Joint			×			
Joint High Speed Vessel (JHSV)	IOT&E	Navy			×	×	×	
Miniature Air-Launched Decoy (MALD) and MALD – Jammer (MALD-J)	IOT&E	Air Force	×					
Q-53 Counterfire Target Acquisition Radar System	IOT&E	Army	×		×			
RQ-21A Blackjack (formerly Small Tactical Unmanned Aerial System (STUAS))	IOT&E	Navy		×				
AIM-120D Advanced Medium-Range Air-to-Air Missile (AMRAAM)	FOT&E	Air Force				×	×	
Air Operations Center – Weapon System (AOC-WS)	FOT&E	Air Force				×	×	
Mark XIIA Mode 5 Identification Friend or Foe (IFF)	FOT&E	Navy	×					
MK 54 Lightweight Torpedo	FOT&E	Navy		×				
Manpack Radio	FOT&E	Army		×	×			

The following discussion involves the discovery of new problems in two of the programs listed in Table 7. The programs are: (1) the DOD Automated Biometric Identification System (ABIS) and (2) the Ballistic Missile Defense System (BMDS). These two programs illustrate the value of OT regarding new problem discovery. Specifically, the DOD ABIS provides a powerful example of the benefits of testing in an operationally realistic environment. The BMDS is an example of an exceedingly complex weapon system that discovers problems during operational flight testing (costing hundreds of millions of dollars), that should have been found in a more cost-effective fashion through comprehensive ground testing.

DOD Automated Biometric Identification System (ABIS)

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.



ABIS has been fielded and supported as an Army Quick Reaction Capability since 2009. Since it was not a formal program of record, ABIS has not had an approved Test and Evaluation Master Plan to guide the developmental or OT&E of this system. After an initial deployment attempt in August 2013, which was unsuccessful, the biometrics program undertook a set of user tests that, while not fully conducted to the rigor of a formal DT, were more rigorous than previous regression testing.

In August 2014, the Army Test and Evaluation Command performed a two-phased OT on ABIS version 1.2. This was the first OT conducted on the system. The first phase was conducted August 7 – 28, 2014. The second phase, which was supposed to begin directly following the first phase, was delayed to address the five effectiveness problems discussed below. The second phase of OT was conducted October 17 – 22, 2014.

The first phase of OT was structured to allow comparison between the then current Authoritative Database (ABIS 1.0) and the system under test (ABIS 1.2) by streaming all live data into both systems. To mitigate operational risk, only ABIS 1.0 sent responses back to the field. Since the ABIS 1.2 system was not the authoritative system, Phase 1 of the OT could have been conducted as an operationally realistic DT event.

During the first phase of OT, the following problems were noted in a DOT&E memo to Army acquisition leadership. If a rigorous DT using an operationally realistic environment had been conducted prior to the OT, the problems detailed in the DOT&E memo after Phase 1 of the OT would have likely been identified. The issues are detailed as follows:

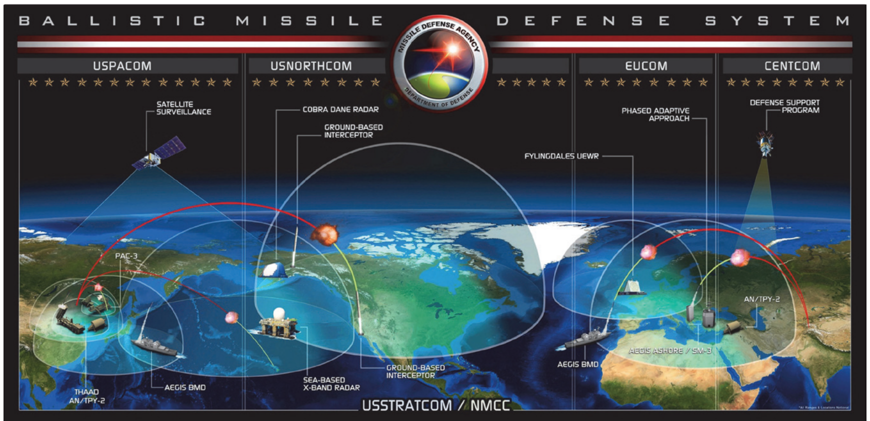
- The National Ground Intelligence Center (NGIC) puts all ABIS biometric match results into its Biometric Identity Intelligence Repository (BI2R). NGIC also fuses the data from worldwide biometric collection systems into and out of ABIS. BI2R is used by DOD Intelligence agencies to identify persons that should be added to the watchlist. During Phase 1 OT, NGIC observed thousands of discrepancies between match results returned from ABIS 1.0 and 1.2. Other problems included incorrect email addresses for sending alerts. Without alerts, no actions can be taken when a person on a watchlist is identified by the system. The mission impact is the potential loss of actionable intelligence when encountering persons of interest throughout the world. OT was necessary to uncover this problem because the number and complexity of live interfaces with real-world biometric submitters could not be adequately simulated in a DT.
- A latent fingerprint is one taken from an object in the field, such as an improvised explosive device. Latent (fingerprint) examiners at the Biometrics Identity Management Activity noted a key identifier (Grand ID) was missing from Latent Examination tools in ABIS 1.2. This capability was available in ABIS 1.0. The Grand ID enables latent examiners to link different latent images with a single forensic case. The problem was not discoverable prior to entering OT because the user cases that were exercised required external interfaces with biometric and latent submitters that could not be simulated in the DT environment.
- ABIS 1.2 responses to biometric submissions failed to meet the specifications required by the Federal Bureau of Investigation Integrated Automated Fingerprint Identification System (IAFIS) preventing acceptance by the IAFIS

interface. An operational environment with the actual production equipment receiving submissions in parallel with the legacy operational system was essential to allow discovery of such issues.

- One of four custom watchlists had over 1,800 discrepancies between the responses from ABIS 1.0 and 1.2. Custom watchlists are used by military personnel in the field to determine courses of action when a person is detained in a particular geographic area. Custom Biometrically Enabled Watchlists could have been assessed before the Phase 1 OT began while the live submissions were streaming into both systems.

Ballistic Missile Defense System (BMDS)

The BMDS is designed to protect the United States, deployed forces, allies, and friends against ballistic missiles of all ranges and in all phases of flight. The BMDS is a distributed system currently comprised of five elements (four shooters and one command and control element) and five sensor systems (four radar systems and one space-based system).



Aegis BMD – Aegis Ballistic Missile Defense
NMCC – National Military Command Center
SM-3 – Standard Missile-3
THAAD – Terminal High-Altitude Area Defense
UEWR – Upgraded Early Warning Radar
USCENTCOM – U.S. Central Command
USEUCOM – U.S. European Command
USNORTHCOM – U.S. Northern Command
USPACOM – U.S. Pacific Command
USSTRATCOM – U.S. Strategic Command

The first OT of the BMDS, referred to as Flight Test Operational – 01 (FTO-01) occurred in September 2013 and demonstrated a layered upper-tier regional/theater BMDS defense against a raid of two simultaneously-launched and threat-representative medium-range ballistic missiles threatening a shared defended area. Although a layered defense was demonstrated in this test, true system integration was not demonstrated due to system network configuration errors, interoperability limitations, and component failures.

FTO-01 was an extremely complex flight test event—it was the second most complex flight test ever attempted by the Missile Defense Agency (MDA) to date. A major difficulty in finding problems such as those uncovered during FTO-01 is that the BMDS can be instantiated in many ways using different combinations of shooters, sensors, and operational laydowns. Despite this variability, some of these findings could have been discovered prior to executing the flight test. In particular, some of the network configuration errors could have been discovered through comprehensive ground testing and analyses. The MDA has taken action to correct the problems uncovered during FTO-01. Details of the problems and specific actions are classified.

Known problem re-observations

Some problems are observed in OT that are already known from prior testing; known problems were observed in 25 of the 48 programs that underwent OT in FY14. As noted earlier, many programs that re-observed known problems also experienced new problem discovery in OT; these are highlighted in grey in Table 8.

TABLE 8. PROGRAMS WITH KNOWN PROBLEM RE-OBSERVATIONS

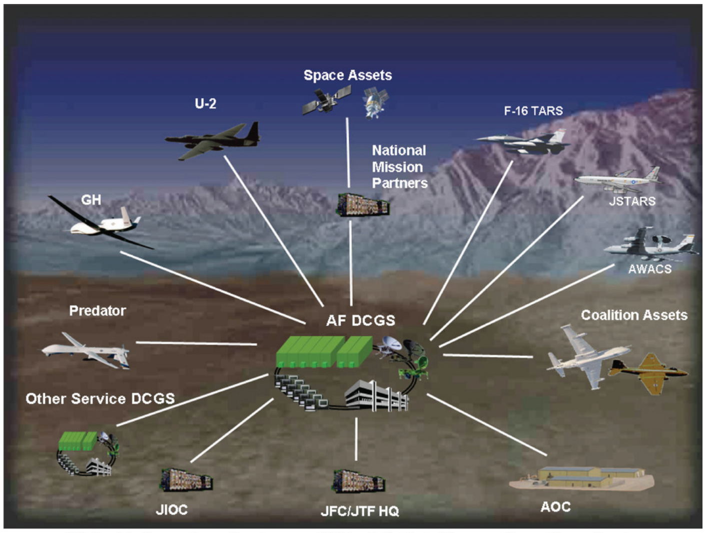
Program	OT Event Type	Service	Known Problem		
			Effectiveness	Suitability	Cybersecurity
AN/PRC-117G	Pre-IOT&E	Army	×		
AN/SQQ-89A(V)15 Integrated Undersea Warfare (USW) Combat System Suite	Pre-IOT&E	Navy		×	
CVN-78 *Gerald R. Ford* class Nuclear Aircraft Carrier	Pre-IOT&E	Navy	×	×	
Defense Enterprise Accounting and Management System (DEAMS)	Pre-IOT&E	Air Force	×	×	×
Infrared Search and Track (IRST)	Pre-IOT&E	Navy	×		
Aegis Ballistic Missile Defense (BMD)	IOT&E	Navy, MDA		×	
F-15E Radar Modernization Program (RMP)	IOT&E	Air Force		×	
Joint Battle Command – Platform (JBC-P)	IOT&E	Joint		×	
Littoral Combat Ship (LCS)[1]	IOT&E	Navy	×	×	×
Miniature Air Launched Decoy (MALD) and MALD – Jammer (MALD-J)	IOT&E	Air Force		×	
Multi-Static Active Coherent (MAC) System	IOT&E	Navy	×		
Q-53 Counterfire Target Acquisition Radar System	IOT&E	Army	×		
QF-16 Full-Scale Aerial Target (FSAT)	IOT&E	Air Force	×		
RQ-21A Blackjack (formerly Small Tactical Unmanned Aerial System (STUAS))	IOT&E	Navy		×	
Surveillance Towed Array Sensor System (SURTASS) and Compact Low Frequency Active (CLFA) Sonar	IOT&E	Navy	×		
Air Force Distributed Common Ground System (AF DCGS)	FOT&E	Air Force	×		
Air Operations Center – Weapon System (AOC-WS)	FOT&E	Air Force			×
Battle Control System – Fixed (BCS-F)	FOT&E	Air Force			×
F/A-18E/F Super Hornet and EA-18G Growler	FOT&E	Navy	×	×	
Joint Warning and Reporting Network (JWARN)	FOT&E	Joint		×	
Mark XIIA Identification Friend or Foe (IFF) Mode 5	FOT&E	Navy	×		
MK 54 Lightweight Torpedo	FOT&E	Navy	×		
MV-22 Osprey	FOT&E	Joint		×	
P-8A Poseidon Multi-Mission Maritime Aircraft (MMA)	FOT&E	Navy	×		
Manpack Radio	FOT&E	Army	×	×	

[1] . Two survey entries for separate oversight programs with separate problems, both discussed in the LCS section of the annual report.

The following discussion involves the re-observation of known problems in three of the programs listed in Table 8. The programs are: (1) Air Force Distributed Common Ground System (AF DCGS); (2) Multi-static Active Coherent (MAC) System; and (3) Small Tactical Unmanned Aerial System (STUAS) Tier II. These three programs illustrate the value of OT in highlighting the operational implications of known problems. In these three cases, program management either decided to accept the risk that their known problems would not affect the OT assessment, or let schedule drive the program into OT in spite of known shortcomings.

Air Force Distributed Common Ground System (AF DCGS)

The AF DCGS provides software tools for operators to task, process, exploit, and disseminate Intelligence, Surveillance, and Reconnaissance information. AF DCGS consists of multiple ground systems at dispersed operational sites. AF DCGS participates in the DOD intelligence enterprise via the DCGS Integration Backbone, which uses a metadata catalog and discovery service to enable sharing of information among participants.



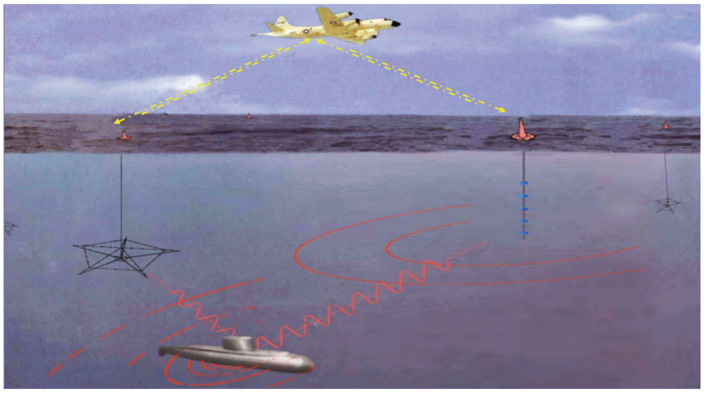
AOC - Air Operations Center
AWACS - Airborne Warning and Control System
GH - Global Hawk
HQ - Headquarters
JFC - Joint Forces Command
JIOC - Joint Intelligence Operations Center
JSTARS - Joint Surveillance Target Attack Radar System
JTF - Joint Task Force
TARS - Tactical/Theater Airborne Reconnaissance System

AF DCGS Bulk Release 10B failed both developmental and regression testing and did not meet the entrance criteria for the OT phase known as the Force Development Evaluation. Despite not meeting the OT entrance criteria (the system had two known Category I and four Category II software deficiencies that were open and unresolved), the Air Force Intelligence, Surveillance, and Reconnaissance Agency approved entrance into OT. In January and June 2014, the 605th Test and Evaluation Squadron conducted Phases 1 and 2 of a two-phase Force Development Evaluation to assess the operational effectiveness and suitability of AF DCGS Bulk Release 10B. Two new software applications that were part of the Geospatial Intelligence upgrade known as Bulk Release 10B had major performance problems. They caused such significant slowdowns in workflow that the Air Force made the decision to stop using the new applications, and operators reverted to using the legacy manual processes during the test. The system did not meet any of its reliability requirements because of critical failures and downtime. While users can execute their missions with AF DCGS under normal load conditions, performance under heavy loads could not be determined. Heavier loads are expected in the future when new sensors are deployed and the number of simultaneous external users is increased.

In part because the Air Force placed AF DCGS in the sustainment vice development phase, the program lacks a strategy for testing and evaluation, documented performance requirements for planned enhancements, accurate software maturity trend information, and an approved system-engineering plan. By developing and following these key programmatic guidance documents, the Air Force would likely improve AF DCGS performance.

Multi-static Active Coherent (MAC) System

The MAC system is an active sonar system composed of two types of sonobuoys (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.



The Navy completed OT of the MAC Phase 1 system on P-3C Multi-mission Aircraft (MMA) in October 2013. OT consisted of 3 DT events conducted off the coast of Jacksonville, Florida; 7 dedicated OT events conducted in the Southern California Fleet Operating Areas; and 14 events in the Narragansett Bay Operating Areas. After the series of OT in January 2013 near San Diego, the Navy knew the system did not work in some environments. However, the program requirement criterion for MAC was a roll-up of test detection results, and testing was not stopped until a series of test events in May 2013 in the Narragansett Bay test area, where performance appeared below threshold, even after counting DT results prior to 2013.

During the May 2013 series of tests, it became clear operators were not recognizing valid target returns as targets because the target-signature criteria they had been trained to use did not cover the new environment. There was also a need to fix some materiel information technology problems with the aircraft used for the test. OT was halted for a period of four to five months because of a combination of materiel problems with the aircraft used to employ the system and training of operators to use the system. More thorough DT might have minimized or eliminated this delay. The operators were retrained to recognize new target signature features that enabled them to distinguish between valid target returns and clutter returns more effectively. In the October 2013 test series following the re-training, the operators were able to recognize valid targets more accurately, but not by a margin that could be clearly distinguished from previous rates under the confidence limits of the data collected.

RQ-21A Blackjack (formerly Small Tactical Unmanned Aerial System (STUAS) Tier II)

Marine Corps commanders will use the RQ-21A Blackjack (formerly Small Tactical Unmanned Aerial System (STUAS)) to provide units ashore with a dedicated persistent battlefield Intelligence, Surveillance, and Reconnaissance (ISR) capability that will reduce their dependence on higher headquarters for ISR support. The persistence of the system allows commanders greater coverage of their areas of interest, while providing the capability to concentrate for longer periods of time on a specified target of interest. The Marine Corps is developing RQ-21A as an organic asset in an effort to wean itself off the contractor-owned, contractor-operated systems currently under contract. In order to transition from ISR services contracts to an organic ISR asset, the Program Office decided to enter IOT&E in spite of the low reliability demonstrated during an earlier operational assessment.

The Navy started the RQ-21A IOT&E in January 2014. Testing consisted of a land-based IOT&E phase (with concurrent ship-based DT) intended to be followed by a ship-based IOT&E phase aboard an LPD-17 class ship. During the land-based phase of IOT&E at Marine Corps Air Ground Combat Center, Twentynine Palms, California, operators flew 188 flight hours during 31 flights. The first flight ended in a mishap and loss of the air vehicle. Post-mishap investigation suspended OT flights for 10 days. The RQ-21A demonstrated a Mean Flight Hours Between Abort (MFHBA) of 15.8 hours, well below the MFHBA threshold criterion of 50 hours. Low reliability adversely affected the ability of operators to support ground units in a timely manner. Many of the reliability problems identified during the land-based IOT&E appear to result from poor quality control during the production process. The Program Office is working with the manufacturer to increase quality control processes with sub-vendors, improve acceptance testing of spare parts, and review their acceptance procedures.

Concurrent with the land-based phase of IOT&E, the Navy conducted RQ-21A ship-based DT aboard an LPD-17 class ship. This ship testing identified interference between the ship's degaussing system and the air vehicle's magnetometer. Without realistic shipboard testing, this deficiency would not have been identified. This deficiency necessitated software upgrades and regression testing, which delayed the scheduled ship-based phase of IOT&E until December 2014. Based on poor system performance during the land-based phase of IOT&E and software update to correct a GPS deficiency associated with shipboard operations, the Navy conducted a second land-based phase of IOT&E in June at Marine Corps Base Camp Lejeune, North Carolina. Operators flew 20.9 hours during eight flights. Analysis of results is ongoing.

For programs with upcoming OT events in the next three years, I found that slightly more than one-third (15 of 42) of the programs currently do not exhibit performance problems significant enough to jeopardize successful performance in OT. Table 9 shows these results by type of problem and phase of testing.

PROGRAMS WITH UPCOMING OT EVENTS

For programs with upcoming OT events in the next three years, I found that slightly more than one-third (15 of 42) of the programs currently do not exhibit performance problems significant enough to jeopardize successful performance in OT. Table 9 shows these results by type of problem and phase of testing.

Upcoming pre-IOT&E test events are far more likely to be delayed to correct problems compared to both upcoming IOT&E and FOT&E. In fact, Table 9 shows that for the programs covered in this Annual Report, there were no upcoming FOT&E events that were delayed to correct problems. Of the programs for which potential problems exist for upcoming OT events, 5 of 7 of the pre-IOT&E events were delayed to address at least one issue, 2 of 10 of the IOT&E events were delayed, and 0 of 6 of the FOT&E (or post-IOT&E) events were delayed.

TABLE 9. PROGRAMS COMMENCING OT WITHIN THE NEXT THREE YEARS

Category	Number of Programs[1]	Number and Type of Problems (program count)[2]		Phase of Testing (program count)		
				Pre-IOT&E	IOT&E	FOT&E
No problems for upcoming OT	15			4	5	6
Problems have delayed upcoming OT	7	Effectiveness	11	3	0	0
		Suitability	9	4	2	0
		Cybersecurity[3]	5	0	0	0
			Total	**5**	**2**	**0**
Problems have not delayed upcoming OT	19	Effectiveness		2	7	4
		Suitability		1	5	4
		Cybersecurity[3]		1	1	1
			Total	**4**	**8**	**6**
Other problems threaten upcoming OT	4			2	2	0

1. Forty-two programs will undergo an OT in the next three years. The number of programs adds up to more than 42 because some programs have problems that delayed their upcoming OT as well as problems that did not delay OT.
2. The number of programs summed across type of problems adds up to more than the number of programs because some programs have multiple problems or more than one type of problem..

For programs that have not delayed their upcoming OT and have known problems, the distribution between effectiveness and suitability problems is about the same (13 compared to 10), with three cybersecurity problems. Table 10 expands further upon programs that have not delayed their upcoming OT and shows that FOT&E events are considerably more likely to have a fix implemented (and in many cases tested) going into the OT, regardless of type of problem, compared to both pre-IOT&E and IOT&E. The data are currently insufficient to determine whether the differences between the rate of implementing and testing fixes prior to FOT&E compared to IOT&E is a trend or simply random.

TABLE 10. ACTIONS TAKEN TO ADDRESS PROBLEMS FOR UPCOMING OT EVENTS FOR WHICH OT HAS NOT (YET) BEEN DELAYED

Category	Type of Problem	How has the problem been addressed to date (program count)?								
		pre-IOT&E			IOT&E			FOT&E		
		Fix Not Identified	Fix Documented	Fix implemented (tested)	Fix Not Identified	Fix Documented	Fix implemented (tested)	Fix Not Identified	Fix Documented	Fix implemented (tested)
Problems have not delayed upcoming OT[1]	Effectiveness	2	0	1 (0)	4	4	2 (1)	2	0	2 (2)
	Suitability	0	1	0	4	1	1 (0)	0	1	3 (2)
	Cybersecurity	1	0	0	0	0	1 (0)	0	0	1 (1)

1. Numbers in parentheses are the number of problems that have already tested fixes prior to the upcoming OT.

Specific Programs that have upcoming OT in the next three years

No problems for upcoming OT

Fifteen of the 42 programs with upcoming OT events have not yet exhibited problems considered to significantly jeopardize performance in upcoming OT events. Such programs could be examples where the program development process, including DT and OT, appears to be moving along well. Alternatively, the testing to date might not have been sufficiently stressing to surface any problems. These programs are listed below in Table 11.

Problems delayed upcoming OT

For some programs, early testing has uncovered problems and entry into the upcoming OT has already been delayed to provide the program an opportunity to correct them. Seven programs fall into this category and are given in Table 12. Note that some programs that have problems that delayed their upcoming OT also have problems that did not delay the OT; the programs with both types of problems are highlighted in grey.

Problems have not delayed upcoming OT

Some programs have uncovered problems in early testing that, if not satisfactorily corrected, could result in my assessing the system as not being operationally effective or suitable. Unlike the above, the OT has not (yet) been delayed to correct these problems. These programs are also shown in Table 12. Note that some programs that have identified problems that did not delay their upcoming OT have also identified other problems that did delay the OT; the programs with both types of problems are highlighted in grey.

TABLE 11. PROGRAMS WITH NO PROBLEMS FOR UPCOMING OT
AIM-9X – Air-to-Air Missile Upgrade
Air and Missile Defense Radar (AMDR)
C-17 Increase Gross Weight (IGW) and Formation Spacing Reduction (FSR)
Common Aviation Command and Control System (CAC2S)
E-2D Advanced Hawkeye
F-22A Advanced Tactical Fighter
Family of Advanced Beyond Line-of-Sight Terminals (FAB-T)
Guided Multiple Launch Rocket System – Alternative Warhead (GMLRS-AW) XM30E1[1]
Integrated Defensive Electronic Countermeasures (IDECM)
Joint Information Environment (JIE)[2]
Joint Standoff Weapon (JSOW)
Massive Ordnance Penetrator (MOP)
Rifleman Radio
Small Diameter Bomb (SDB)
Standard Missile-6 (SM-6)

[1] Emerging Results from the recent GMLRS-AW IOT&E indicate its lethality is insufficient.

[2] The JIE has not had any OT to date pending development of governance processes for the Joint Regional Security Stack transport infrastructure.

TABLE 12. PROGRAMS WITH PROBLEMS THREATENING UPCOMING OT

Program	OT Event Type	Service	OT Delayed			OT Not Delayed		
			Effectiveness	Suitability	Cybersecurity	Effectiveness	Suitability	Cybersecurity
AC-130J Ghostrider	Pre-IOT&E	USSOCOM	×					
Distributed Common Ground System – Army (DCGS-A)	Pre-IOT&E	Army		×				×
F-35 Joint Strike Fighter (JSF)	Pre-IOT&E	Joint	×	×				
Key Management Infrastructure (KMI)	Pre-IOT&E	NSA		×				
MQ-4C Triton Unmanned Aircraft System	Pre-IOT&E	Navy				×		
Public Key Infrastructure (PKI)	Pre-IOT&E	Joint					×	
Remote Minehunting System (RMS)	Pre-IOT&E	Navy	×	×		×		
AN/SQQ-89A(V)15 Integrated Undersea Warfare (USW) Combat System Suite	IOT&E	Navy					×	
Ballistic Missile Defense System (BMDS)	IOT&E	MDA				×		
Defense Enterprise Accounting and Management System (DEAMS)	IOT&E	Air Force				×		×
Infrared Search and Track (IRST)	IOT&E	Navy				×		
LHA-6 New Amphibious Assault Ship	IOT&E	Navy				×		
Littoral Combat Ship (LCS)	IOT&E	Navy		×		×	×	
Mobile Landing Platform (MLP) Core Capability Set (CCS) and Afloat Forward Staging Base (AFSB)	IOT&E	Navy					×	
Patriot Advanced Capability-3 (PAC-3)	IOT&E	Army				×	×	
Precision Guidance Kit (PGK)	IOT&E	Army		×				
RQ-4B Global Hawk High-Altitude Long-Endurance Unmanned Aerial System (UAS)	IOT&E	Air Force				×	×	
AGM-88E Advanced Anti-Radiation Guided Missile (AARGM)	FOT&E	Navy				×		
Air Force Distributed Common Ground System (AF DCGS)	FOT&E	Air Force				×		
Defense Medical Information Exchange (DMIX)	FOT&E	Joint				×	×	
M829E4 Armor Piercing, Fin Stabilized, Discarding Sabot-Tracer (APFSDS-T)	FOT&E	Army					×	
MQ-9 Reaper Armed Unmanned Aircraft System (UAS)	FOT&E	Air Force					×	
Warfighter Information Network – Tactical (WIN-T)	FOT&E	Army				×	×	×

The following discusses the problems that potentially jeopardize successful performance in upcoming OT in three of the programs listed in Table 12. The programs are: (1) AC-130J Ghostrider; (2) Remote Minehunting System (RMS); and (3) Warfighter Information Network – Tactical (WIN-T). These programs illustrate the types of problems that jeopardize successful performance in upcoming OT and should be addressed to the maximum extent possible prior to OT.

AC-130J Ghostrider

The AC-130J is a medium-sized, multi-engine, tactical aircraft with a variety of sensors and weapons for close air support. U.S. Special Operations Command is developing AC-130J through the integration of a modular Precision Strike Package (PSP) onto existing MC-130J aircraft. The PSP provides a 30 mm side-firing gun; wing-mounted, GPS-guided Small Diameter Bombs; Griffin laser-guided missiles; two electro-optical/infrared sensor/laser designator pods; a synthetic aperture radar pod; and multiple video, data, and communication links.

There have been problems with integration of the PSP weapon kit onto the aircraft that continue to delay portions of DT by prohibiting weapons employment and hindering system effectiveness. First, the visual acuity of the electro-optical/infrared sensors installed on the AC-130J is not sufficient for accurate target identification and designation because of excessive vibration on the new aircraft as compared to the legacy AC-130W aircraft on which the PSP was previously installed. Second, electrical/radio-frequency interference between aircraft systems and the hand controllers used by crewmembers to direct the sensors and weapons has caused erratic sensor movements. This inhibits target tracking and is a safety hazard (risk of fratricide) during weapon employment. The program is working on correcting the sensor vibration issue by collecting flight test data that can be used by the subsystem contractor to develop mechanical and software updates to reduce the effect of vibration. Similar efforts are underway to characterize and correct electrical interference with the controllers. The program has reported some progress in the laboratory environment on both fixes, but definitive solutions have not yet been demonstrated on the aircraft.

The program has accomplished 36 test flights out of approximately 130 flights planned for a total of 97 flight hours. Initial DT is now expected to be completed in May 2015. Delays in DT have delayed the planned operational assessment by the 18th Flight Test Squadron by approximately four months, and IOT&E has been delayed until October 2015. This schedule does not allow much time to developing and implementing fixes to problems already observed in the DT.

Remote Minehunting System (RMS)

The RMS is a system-of-systems designed to detect and classify mine-like objects throughout the water column and to identify bottom objects in shallow waters. The Navy expects to employ the system with both variants of the Littoral Combat Ship (LCS) as a key component of the Mine Countermeasures (MCM) mission package.

DOT&E disapproved the Navy's plan to conduct an operational assessment of the RMS in 2QFY14 because the assessment would have been a wasted effort for the following reasons:

- The proposed test article was not representative of the system the Navy plans to employ in the first increment of the LCS MCM mission package (it was an earlier version without planned upgrades) and therefore, would not provide data necessary to augment the IOT&E of an LCS equipped with that mission package;
- Test limitations would have precluded an operational evaluation of some phases of the end-to-end mission; and
- Conduct of the test would have delayed vehicle upgrades necessary to support testing of the system the Navy expects to field.

The RMS program has not yet demonstrated that the system can meet its detection and classification requirements against moored and bottom mines spanning the portion of the shallow water regime not covered by the Airborne Laser Mine Detection System (ALMDS). The program anticipates that the AN/AQS-20B sensor will permit the system to cover the portion of the water column below that covered by the ALMDS. The new sensor will be tested in FY15.

RMS radios have had difficulty establishing reliable communications with the LCS during DT, and once communications are established, the current communications systems do not support Remote Multi-Mission Vehicle (RMMV) mine identification operations beyond the horizon. RMMV will need to operate beyond the horizon to support efficient MCM operations in long shipping channels while LCS remains in an area clear of mines. This problem arose when the Navy decertified the MH-60S helicopter for towing MCM devices, including the AN/AQS-20A/B sensor. The range limitation did not exist when the sensor was towed by the helicopter. The Navy has not subsequently developed a solution to this problem.

The combined results of shore-based and LCS-based testing conducted since the program was recertified following a Nunn-McCurdy breach in 2010 have not demonstrated that an LCS equipped with an MCM mission package that includes two RMMVs and three AN/AQS-20A sonars will be able to support the sustained area coverage rate the Navy has established for the Increment 1 MCM mission package. The program believes that RMMV reliability improvements and an upgraded version of the minehunting sensor, designated AN/AQS-20B, will resolve many of the program's identified problems.

The reliability of the version 4.2 (v4.2) RMMV during combined developmental and integrated testing completed in FY14 was 31.3 hours Mean Time Between Operational Mission Failure (MTBOMF), which is well below the required reliability. DT completed in 1QFY15 provides a point estimate for v6.0 vehicle reliability of 34.6 hours MTBOMF. Statistical analysis of all test data indicates the result is not sufficient to conclude that reliability has actually improved since a Nunn-McCurdy review of the program in 2010. Therefore, test data currently available (including early testing of the v6.0 vehicle) do not support the Navy's assertion that vehicle reliability has improved. Moreover, the current estimate of RMS reliability, once all of the other components of the system are considered, is no more than 20 hours MTBOMF, which is well-short of what is needed to complete MCM missions in a timely fashion and meet the Navy's desired mission timelines.

The results of combined DT/integrated testing completed in FY14 continued to show that the RMS's AN/AQS-20A sensor does not meet Navy requirements for contact depth localization accuracy (the difference in depth between reported contact position and ground truth target position) or false classification density (number of contacts erroneously classified as mine-like objects per unit area searched). The sensor also continues to have problems meeting the Navy's detection and classification requirements in shallow waters, and RMS has difficulty guiding the sensor over bottom objects for identification in deep water. Because the first phase of the LCS IOT&E with an embarked MCM mission package was delayed, the Navy was afforded more time to develop an upgraded sensor and implement other system changes that it expects will correct these problems. The program believes that the new sensor, AN/AQS-20B, will correct or greatly mitigate the depth localization and false classification problems; however, the AN/AQS-20B prototypes received from the vendor performed poorly during acceptance and early characterization testing and thus required rework. Testing will continue in FY15.

Warfighter Information Network – Tactical (WIN-T)

WIN-T Increment 2 is a two-tiered communications architecture (celestial and terrestrial) that serves as the Army's high-speed and high-capacity tactical communications network. It is designed to provide reliable, secure, and seamless communications for units operating at theater level and below. It supports both mission command and situational awareness through native WIN-T applications and existing and future battle command applications.

M-ATV Point of Presence

M-ATV Soldier Network Extension

Stryker Point of Presence

1 - Net-Centric Waveform Antenna
2 - High-Band Networking Waveform Antenna

M-ATV - Mine Resistant Ambush Protected (MRAP) All-Terrain Vehicle (M-ATV)

Tactical Comms Node

WIN-T has executed its last three OTs as part of the Army's Network Integration Evaluations (NIEs). This includes a May 2012 IOT&E and May 2013 FOT&E for which DOT&E prepared an IOT&E report and operational assessment report, respectively. A second FOT&E was executed in November 2014 and analysis is ongoing. The NIEs provide access to a full brigade equipped with a complete set of battle command applications to drive traffic on the WIN-T network. The complete brigade is necessary for OT to ensure the WIN-T transport layer can realistically support the data needs of a brigade with a complete set of battle command applications. While laboratory testing of these is possible, it is difficult to execute, and DOT&E has not yet seen a DT for WIN-T that included the full breadth of these applications. The only way to ensure thorough OT is to use a fully equipped and trained brigade combat team.

A concern with the NIEs, from an OT perspective, is their inherent schedule-driven nature; NIEs are very complex events, which are held twice each year. Planning for the NIEs begins 12 to 18 months prior to execution and systems are inserted into the event after planning has begun. Relevant Army programs plan their test schedule around fitting into the NIE, rather than ensuring their system is truly ready for test. The NIE is but one example of external events driving the OT schedule vice scheduling tests to verify fixes implemented to correct problems observed in earlier testing.

WIN-T has executed four OT to date (including a Limited User Test executed in May 2009, prior to the existence of NIEs). Some performance problems identified in the IOT&E have remained constant throughout WIN-T testing. Many of these problems could not have been observed during the WIN-T DTs. Sometimes this was due to the limited scope of the DTs, sometimes because the observation requires a representative unit facing a representative threat in an operational environment. The problems include:

- Poor performance of the line-of-sight Highband Networking Waveform (HNW) – HNW is required to offer 27 megabits per second (Mbps) at-the-halt at a 12-kilometer distance and 18 Mbps on-the-move at a distance of 2 kilometers. OT has shown the HNW is not capable of providing this capability to a dispersed brigade. This could have been identified in DT, but was not because of the limited scope and benign conditions of DT.
- Poor performance of the Soldier Network Extension (SNE) – The SNE is a company-level vehicle kit that includes a satellite transponder and computer for connection to the WIN-T network. At the IOT&E and FOT&E, it had major usability and reliability problems that were only discoverable in OT. Identification of these problems required the evaluation of the ability of representative trained operators in an operational setting to execute their mission.
- Lack of Network Operations capability – Outside of the central Network Operations and Security Center, there is very limited capability for the unit to monitor and manage the WIN-T network. This was only observable in OT. This would have been difficult to identify in DT because it requires an assessment of a representative unit's ability to monitor and manage a dispersed network reacting to a realistic operational scenario.
- Poor reliability – The WIN-T Increment 2 configuration items were not reliable. They did not meet the Army's requirements or serve the needs of operators and commanders. The consequences of reliability problems on the unit's ability to complete its mission are discoverable in OT but not DT. The context of OT provides programs and users the magnitude of the mission consequences.
- Cybersecurity vulnerabilities – The Army's brigade-level network has a significant number of cyber vulnerabilities. These vulnerabilities can only be put into context and evaluated properly when tested using a representative computer network defense and threat employed during OTs, such as the NIEs. Additionally, cybersecurity assessments require the presence of the complete set of battle command applications (hardware and software) and the support of external computer network defense organizations to create a representative environment, which is only available through OT. The WIN-T Program Office has combined their efforts with the cooperative and adversarial cybersecurity assessment teams to identify vulnerabilities and initiate fixes.

RELIABILITY

Of the 81 total programs surveyed, 21 had reliability problems serious enough to either negatively affect the suitability assessment in an FY14 OT report or jeopardize successful performance in OT. Programs are taking corrective actions throughout the acquisition cycle to address reliability problems, but for some systems, reliability remains a concern even after IOT&E or FOT&E. Table 13 summarizes this information. (One system had reliability problems during a pre-IOT&E test that will be tested in an upcoming FOT&E.)

TABLE 13. FY14 OT RELIABILITY RESULTS BASED ON NUMBER OF PROGRAMS

Category	Number of programs with reliability problems[1]	Pre-IOT&E	IOT&E	FOT&E
Known problem re-observations	9	1	6	2
New problem discovery	4	1	2	1
Problems delayed upcoming OT	4	2	2	0
Problems have not delayed upcoming OT	5	1	2	2

1. Twenty-one programs had reliability problems serious enough to either affect a suitability assessment or jeopardize the successful performance in an upcoming OT event. I identified one program that had a reliability problem in a pre-IOT&E event, and this item will be tested in an upcoming FOT&E. This program contributes both to the number of programs that have conducted an OT and to the number programs with upcoming OT events. Thus, the sum of the number of programs is 22.

For the most part, programs are either delaying OT to address reliability and/or are implementing fixes to address reliability prior to entering an OT event; Table 14 summarizes program responses to reliability problems. Of the eight programs that re-observed reliability problems during an IOT&E or FOT&E, six implemented fixes to address reliability prior to the OT event. Similarly, of the nine programs with known reliability problems that jeopardize successful performance in an upcoming OT event, four have delayed OT to address reliability (Table 14). Early OTs are the most likely to be delayed. All five of the programs with reliability problems that have not delayed an upcoming OT have, at a minimum, identified a fix, and four have implemented fixes.

TABLE 14. PROGRAM RESPONSES TO RELIABILITY PROBLEMS

Category	pre-IOT&E			IOT&E			FOT&E		
	Fix Not Identified	Fix Documented	Fix implemented (tested)	Fix Not Identified	Fix Documented	Fix implemented (tested)	Fix Not Identified	Fix Documented	Fix implemented (tested)
Known problem re-observations (IOT&E or FOT&E)	--	--	--	2	--	4 (1)	0	--	2 (2)
Problems have not delayed upcoming OT	0	1	0	0	1	1 (0)	0	0	2 (1)

Despite program attempts to address reliability, some programs continue to observe reliability shortfalls during IOT&E and FOT&E. In part, this reflects the iterative nature of reliability improvement; programs go through multiple cycles of testing and implementing fixes. Nevertheless, a decrease in reliability problems observed during IOT&E and FOT&E, as opposed to earlier phases of testing, might be possible with further improvement in reliability growth plans. See the discussion on reliability in my introduction to this Annual Report for further details.

PROGRESS UPDATES ON DISCOVERIES REPORTED IN THE FY13 DOT&E ANNUAL REPORT

In FY13, I identified 12 systems that had significant problems in IOT&E that should have been discovered and resolved prior to commencement of OT. They are listed in Table 15 below.

I also identified 10 programs that re-observed known problems in IOT&E, shown in Table 16.

One of the programs in Table 16, the Mission Planning System (MPS) is no longer under oversight. The status of the remaining systems is shown below.

TABLE 15. FY13 SYSTEMS THAT HAD SIGNIFICANT NEW PROBLEM DISCOVERY IN OT

IOT&E with New Problem Discovery	OT other than IOT&E with New Problem 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)
Surveillance Towed Array Sensor System (SURTASS) and Compact Low Frequency Active (CLFA)	Warfighter Information Network – Tactical (WIN-T)

All fixes implemented and demonstrated in OT
- AIM-9X Air-to-Air Missile Upgrade
- Global Command and Control System – Joint (GCCS -J)

Some (or all) fixes implemented but new problems discovered or known problems re-observed in OT
- F-15E Radar Modernization Program (RMP)
- Manpack Radio
- Joint Battle Command – Platform (JBC-P)
- Miniature Air-Launched Decoy (MALD) and MALD-Jammer (MALD-J)
- P-8A Poseidon Multi-Mission Maritime Aircraft (MMA)

Some fixes (potentially) implemented; Currently in OT or planning additional OT
- AIM-120 Advanced Medium-Range Air-to-Air Missile (AMRAAM)
- Defense Enterprise Accounting and Management System (DEAMS)
- DOD Automated Biometric Identification System (ABIS)
- E-2D Advanced Hawkeye
- H-1 Upgrades – U.S. Marine Corps Upgrade to AH-1Z Attack Helicopter and UH-1Y Utility Helicopter
- 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)

TABLE 16. FY13 SYSTEMS THAT HAD KNOWN PROBLEMS RE-OBSERVED IN IOT&E

Known Problem Re-Observations in IOT&E
AIM-120 Advanced Medium-Range Air-to-Air Missile (AMRAAM)
Cooperative Engagement Capability (CEC)
E-2D Advanced Hawkeye
F-15E Radar Modernization Program (RMP)
Global Broadcast System (GBS)
Global Command and Control System – Joint (GCCS-J)
H-1 Upgrades – U.S. Marine Corps Upgrade to AH-1Z Attack Helicopter and UH-1Y Utility Helicopter
Manpack Radio
Mission Planning System (MPS)/Joint Mission Planning Systems – Air Force (JMPS-AF)
P-8A Poseidon Multi-Mission Maritime Aircraft (MMA)

No Fixes Planned
- Mk 54 Lightweight Torpedo

Not reported on in this year's Annual Report because no OT took place this year
- Acoustic Rapid Commercial Off-the-Shelf (COTS) Insertion for Sonar AN/BQQ-10 (V) (A-RCI) and the AN/BYG-1 Combat Control System
- Cooperative Engagement Capability (CEC)
- Global Broadcast System (GBS)

In FY13, I also identified 16 systems that had significant issues in early testing that should be corrected before IOT&E. They are listed in Table 17.

The following provides an update on the progress these systems made in implementing fixes to those problems. Two of these programs are not reported on in this year's Annual Report because no significant OT activity occurred and the Integrated Electronic Health Record (iEHR) program is no longer on the oversight list.

TABLE 17. FY13 SYSTEMS THAT HAD SIGNIFICANT ISSUES IN EARLY TESTING	
CVN-78 *Gerald R. Ford* class Nuclear Aircraft Carrier	LHA-6 New Amphibious Assault Ship
Defense Enterprise Accounting and Management System (DEAMS)	Littoral Combat Ship (LCS)
DOD Automated Biometric Identification System (ABIS)	M109 Family of Vehicles (FoV) Paladin Integrated Management (PIM)
Manpack Radio	Next Generation Diagnostic System (NGDS)
Rifleman Radio and Nett Warrior	Public Key Infrastructure (PKI)
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)

Fixes tested in OT – New problems discovered

- DOD Automated Biometric Identification System (ABIS)
- Surface Ship Torpedo Defense (SSTD) System: Torpedo Warning System (TWS) and Countermeasure Anti-torpedo (CAT)

Fixes tested in OT – Known problems re-observed

- CVN-78 *Gerald R. Ford* class Nuclear Aircraft Carrier
- Defense Enterprise Accounting and Management System (DEAMS)
- Joint Warning and Reporting Network (JWARN)

Fixes tested in OT – Both new problems discovered and known problems re-observed

- Manpack Radio
- Q-53 Counterfire Target Acquisition Radar System

Upcoming testing with no problems identified

- Rifleman Radio and Nett Warrior
- Integrated Defensive Electronic Countermeasures (IDECM)

Upcoming testing with problems identified

- LHA-6 New Amphibious Assault Ship
- Littoral Combat Ship (LCS)
- Public Key Infrastructure (PKI)
- RQ-4B Global Hawk High-Altitude Long-Endurance Unmanned Aerial System (UAS)
- M109 Family of Vehicles (FoV) Paladin Integrated Management (PIM)

Not reported on in this year's Annual Report because no OT took place this year

- Next Generation Diagnostic System (NGDS)