OFFICE OF THE SECRETARY OF DEFENSE
1700 DEFENSE PENTAGON
WASHINGTON, DC 20301-1700

OPERATIONAL TEST
AND EVALUATION

**INFO MEMO**

JAN 14 2015

FOR: SECRETARY OF DEFENSE

FROM: J. Michael Gilmore, Director, Operational Test and Evaluation

SUBJECT: Director, Operational Test and Evaluation Fiscal Year 2014 Annual Report

- The report at TAB A provides my annual assessment to you and the Congress of programs under my oversight and a summary of associated operational and live fire testing in Fiscal Year (FY) 2014 as required by section 139 of Title 10, United States Code. On Friday, January 16, 2014 I will begin providing my Annual Report to the Congress and on Tuesday, January 20, 2014, I will make the report available on the DOT&E public webpage. The report provides detailed discussions of testing conducted for 92 defense programs during the past fiscal year, any of which may generate interest and discussion within the Congress, Services, and the trade press. The report also discusses findings from my office's cybersecurity testing and assessments, as well as my concerns regarding test resources.

- In the current environment of constrained defense budgets, the value of testing writ large, and of operational testing in particular, is being questioned. It is also being asserted that testing is a major cause of delays in defense programs. My Annual Report also provides a detailed rebuttal of these claims.

- One of the primary purposes of operational testing, and a key value of such testing, is to identify critical problems that can be seen only when systems are examined under the stresses of realistic combat conditions, prior to the full-rate production decision. This identification permits corrective action to be taken before large quantities of a system are procured and avoids expensive retrofit of system modifications. The assertion that testing causes delays misses the essential point: fixing deficiencies causes delays, not the testing. Furthermore, taking the time to correct serious problems is exactly what we desire in a properly-functioning acquisition system. For example, operational testing of the Navy's Cooperative Engagement Capability (CEC) on the E-2D Hawkeye aircraft revealed several deficiencies. The CEC created many more dual tracks compared to the baseline CEC system, exhibited interoperability problems with the E-2D mission computer, and there was degradation in CEC's ability to maintain consistent air tracks compared to the baseline E-2C version. As a result of these discoveries in operational testing, the Navy's acquisition executive decided to delay the full-rate production decision until the root causes for these deficiencies could be found and fixed. The Navy is now implementing fixes to address these problems, and operational testing will be conducted to verify these fixes have corrected the problems. The value of such testing is abundantly clear if one considers the alternative: discovering these problems for the first time in combat, when it is too late to correct them.

- As discussed in my report, this year, my office updated a previous study that we conducted with the Under Secretary of Defense for Acquisition, Technology and Logistics (USD(AT&L)) in 2011 on the causes of program delays. This year's analysis examined case studies for 115 acquisition programs which were selected because they had experienced a delay of six months or more and had a full-rate production decision since the year 2000. Not surprisingly, consistent with the prior study, the analysis found the most common reason that contributes to a delay is a performance problem discovered during developmental or operational testing that acquisition executives decide must be addressed before a program moves forward.

- In the remainder of this memorandum, I provide a synopsis of my report's discussions of four acquisition programs I judge to be particularly noteworthy: the Littoral Combat Ship (LCS), the Joint Strike Fighter (JSF), the DDG-51 Flight III destroyer equipped with the Air and Missile Defense Radar (AMDR), and the Army Manpack radio. I also provide a synopsis of my report's discussion of Cybersecurity.

***Littoral Combat Ship (LCS).***

- Developmental testing and operational assessments conducted during the past year have revealed numerous problems with all elements of the Mine Countermeasures (MCM) package to be employed ou LCS. There is little time remaining to correct these problems before at-sea operational testing of the package scheduled to be conducted in 2015. If these problems remain uncorrected, the ability of the MCM package to fulfill even the minimal Increment I requirements promulgated by the Navy for mine reconnaissance is in question.

- A more detailed discussion follows immediately below.

    - The MCM Mission Package has not yet demonstrated sufficient performance to achieve the Navy's minimal Increment 1 requirements. Although the ship's and its crew's ability to launch and recover Remote Multi-Mission Vehicles (RMMVs) has improved, LCS has had difficulty establishing and maintaining reliable communications with the RMMV, and the RMMV continues to exhibit reliability problems. The current communications systems also do not support bottom mine identification beyond the horizon, further limiting the ability to clear mines quickly over a large area. Attempts to demonstrate the sequence of events necessary for an LCS to complete end-to-end mine clearance operations have been limited by low operator proficiency, software immaturity, system integration problems, and poor Remote Minehunting System (RMS)/RMMV reliability.

    - During a shore-based assessment, the Airborne Mine Neutralization System (AMNS) did not meet the Navy's requirement for mine neutralization success. Failures of the host MH-60 aircraft's systems and its associated Airborne MCM kit severely limited AMNS availability. Frequent loss of fiber-optic communications between the aircraft and the neutralizer was the primary cause of unsuccessful attack runs. Both problems increase the time needed to conduct

LCS-based AMNS operations and reduce the ship's sustained area coverage rate for mine clearance.

- Equipment reliability problems have degraded the operational availability of both LCS variants. The Navy reports that recent reliability improvements made to the affected seaframe components have led to improved operational availability; however, that improvement has not been verified in operational testing.

- During FY 2014, the Navy conducted both developmental testing and operational testing of the Freedom class Littoral Combat Ship (LCS) seaframe and Increment 2 Surface Warfare (SUW) Mission Package aboard USS Fort Worth (LCS 3). The 2014 operational testing identified shortcomings in air defense, reliability, and endurance, and significant vulnerabilities in cybersecurity. When equipped with the Increment 2 SUW Mission Package, LCS 3 was able to defeat a small number of Fast Inshore Attack Craft under the particular conditions specified by the Navy's reduced incremental requirement and after extensive crew training and tailoring of the tactics described in Navy doctrine; however, testing conducted to date has not been sufficient to demonstrate LCS capabilities in more stressing scenarios consistent with existing threats.

- The core combat capabilities of the Independence class variant seaframe remain largely untested. Developmental testing has focused on evaluating the performance of the seaframe and the Increment 1 Mine Countermeasures (MCM) Mission Package, with multiple deficiencies identified.

- LCS is not expected to be survivable in high-intensity combat because its design requirements accept the risk that the ship must be abandoned under circumstances that would not require such an action on other surface combatants. Although the ship incorporates capabilities to reduce susceptibility to attack, previous testing of analogous capabilities demonstrates it cannot be assumed LCS will not be hit in high-intensity combat.

- In my Annual Report, I also state that I prepared in October 2014 an independent written report on the combat capabilities and survivability of the alternative concepts for a new small surface combatant (SSC) developed by the Navy's SSC Task Force. In that report, I concluded that the Navy's SSC Task Force's results indicate, of the alternatives it considered, the multi-mission combat capabilities and survivability design features of a modern frigate could be provided only by a new ship design or a major modification to the LCS design – the so-called large plug insertion developed by the Task Force. While offering some improvements in combat capability and survivability (primarily via reduced susceptibility) relative to LCS, the minor modifications to LCS considered by the Task Force and recommended by the Navy Leadership do not satisfy significant elements of a capability concept developed by the Task Force for a modern frigate. (The Task Force developed a number of capability concepts incorporating various mixes of capabilities consistent with a frigate. After consulting with the Task Force's lead,

my assessment used one particular concept as representative of a modern frigate's capabilities. Also, "major modification to LCS" and "minor modification to LCS" are the characterizations used by the Task Force of its alternatives.) Notwithstanding potential reductions to its susceptibility relative to LCS, my assessment is that minor modifications to LCS will not yield a ship that is significantly more survivable than LCS.

***Joint Strike Fighter (JSF).***

- Testing of Block 2B mission systems, which provides initial, limited combat capability to support the Marine Corps' declaration of initial operational capability, will complete early this calendar year as opposed to October 2014, as previously planned. The engine fire that occurred in June 2014 caused some of the delay; but, discoveries of deficiencies as testing proceeded were also causal. Block 2B testing would have completed many months later in 2015 if not for the program office's decision to defer/eliminate a substantial number of test points relative to its previous plans. The extent to which these deferrals and eliminations will cause the need for additional work during testing of Block 3F mission systems (which will provide full combat capability) is not yet known. In my view, what is clear is that Block 2B will finish with deficiencies remaining that will affect operational units. Contrary to previous planning, fixes for these deficiencies will be deferred to testing of Blocks 3i and 3F.

- In March 2014, I recommended to the USD(AT&L) that the Operational Utility Evaluation (OUE) of Block 2B missions systems planned to be conducted during summer 2015 be cancelled because the aircraft would not be ready to support training of operational pilots and successful completion of a comprehensive operational evaluation. Moreover, I expressed the concern that attempting to conduct the OUE would have a detrimental effect on progress completing Block 3F development and testing. The USD(AT&L) accepted my recommendation.

- Overall suitability (including reliability, maintainability, and availability) continues to be less than desired by the Services, and relies heavily on contractor support and unacceptable workarounds, but has shown some improvement in calendar year 2014.

- Serious deficiencies in the hardware and software used to develop mission data files for the JSF have been confirmed by a study conducted by the program office during the past year. The deficiencies now acknowledged are more numerous and serious than those identified by my office in late calendar year 2012. Mission data files enable the aircraft to identify radiofrequency emitters, both friendly and enemy, and are essential to conducting effective combat operations against advanced enemy air defenses, a key reason JSF is being fielded. Unless immediate action is taken to remedy these deficiencies, the ability of the JSF to successfully complete operational testing in 2018 and be effective in combat is at substantial risk.

- A more detailed discussion follows immediately below.

- In spite of a focused effort, the program was not able to accomplish its goal of completing Block 2B flight testing by the end of October. Slower than planned progress in mission systems, weapons integration, and F-35B flight sciences testing delayed the completion of the testing required for Block 2B fleet release. The program now projects this to occur by the end of January 2015, instead of the end of October 2014 as was previously planned.

- Restrictions imposed on the test fleet as a result of the engine failure in June reduced test point availability and slowed progress in mission systems and flight sciences testing from July through November. For example, the effect on mission systems testing was approximately 17 percent loss of productivity in accomplishing test points, from 210 points accomplished per month prior to the engine restrictions to approximately 175 points per month.

- In addition, discoveries of deficiencies continued to occur in later versions of Block 2B software. For example, completion of weapons delivery accuracy events lagged the plans for calendar year 2014 and was put on hold in August when the program discovered a deficiency in the F-35 navigation system. Through the end of November, 10 of 15 weapon delivery events had been completed; all events were planned to be completed by the end of October. However, the program must transition development and flight test resources to Block 3 in order to preserve an opportunity to complete the System Design and Development phase as planned in 2018. Block 2B will finish later than planned, with deficiencies remaining that will affect operational units. Contrary to previous planning, fixes for these deficiencies will be deferred to Blocks 3i and 3F.

- In the FY 2013 Annual Report, I estimated that the program would complete Block 2B testing between May and November 2015 (7 to 13 months late), depending on the level of growth experienced, while assuming the program would continue test point productivity equal to that of the preceding 12 months. Since the end of October 2013, the program has made several adjustments to reduce the delay estimated in my FY 2013 report:

    - In February 2014, while finalizing the 2014 annual plan, the program consolidated test points from plans of earlier blocks of mission systems (Blocks 1A, 1B, and 2A) with those from the Block 2B test plan and decided to account for only those test points needed for Block 2B fleet release, eliminating approximately 840 points. All of these points were planned to be accomplished as of my FY 2013 Annual Report. This reduction amounts to approximately four months of testing. Further adjustments to the baseline number of test points needed for Block 2B fleet release were made in June 2014, resulting in an additional net reduction of 135 points. The extent to which these reductions in Block 2B testing will necessitate additional testing, if any, in Block 3 is not currently known.

- Based on test point accomplishment rates experienced since October 2013, the program will complete Block 2B development in February 2015. This estimate assumes no further growth in Block 2B testing, test point productivity at the current rate, and that operating restrictions stemming from the engine failure are relieved for the test aircraft such that all blocked test points are made available. Completion of Block 2B development by the end of January will, therefore, require an increase in test point productivity and/or elimination of additional test points.

- In April, USD(AT&L) and the program accepted my recommendation that the Block 2B OUE, which was being planned for mid-2015, should not be conducted and that instead, resources should be focused on conducting limited assessments of Block 2B capability and re-allocated to assist in the completion of development and testing of Block 3i and Block 3F capabilities. This recommendation was based on my review of Block 2B progress and assessment of the program's inability to start the Block 2B OUE as planned without creating a significant impact to Block 3F development.

- The Program Office, JSF Operational Test Team, and Service representatives then began working to "re-scope" use of operational test aircraft and operational test activities in lieu of the OUE—detailed planning is ongoing. The scope of the operational test activities will be limited until the flight restrictions imposed due to the engine failure are removed from the operational test aircraft. Moreover, availability of the operational test aircraft will continue to be affected in 2015 and 2016 by the depot time required for modifications to make these aircraft production-representative.

- Overall suitability continues to be less than desired by the Services and needs the program's continued attention to improve in order to meet requirements. Aircraft availability was flat over most of the past year, maintaining an average for the fleet of 37 percent for the 12-month rolling period ending in September – consistent with the availability reported in the FY 2013 DOT&E report of 37 percent for the 12-month period ending in October 2013. However, the program reported an improved availability in October 2014, reaching an average rate of 51 percent for the fleet of 90 aircraft and breaking 50 percent for the first time, but still short of the program objective of 60 percent set for the end of calendar year 2014. Measures of reliability and maintainability that have Operational Requirements Document (ORD) requirements have improved since last year, but all nine reliability measures (three for each variant) are still below program target values for the current stage of development. The reliability metric that has seen the most improvement since May 2013 is not an ORD requirement but a contract specification metric, mean flight hour between failures scored as "design controllable" (which are equipment failures due to design flaws). For this metric, the F-35B and F-35C are currently above (better than) program target values, and F-35A is slightly below (worse than) the target value but has been above the target value for several months during the last year.

***DDG-51 Flight III and the Air and Missile Defense Radar (AMDR).***

- My assessment continues to be that the operational test programs for the AMDR, Aegis Modernization, and DDG 51 Flight III Destroyer programs are not adequate to fully assess their self-defense capabilities in addition to being inadequate to test key Navy-approved AMDR and DDG 51 Flight III requirements.

- The AMDR Capability Development Document describes AMDR's integrated air and missile defense mission, which requires AMDR to support simultaneous defense against multiple ballistic missile threats and multiple advanced anti-ship cruise missile (ASCM) threats. The Capability Development Document also includes an AMDR minimum track range Key Performance Parameter.

- The DDG 51 Flight III Destroyer has a survivability requirement directly tied to meeting a self-defense requirement threshold against ASCMs described in the Navy's Surface Ship Theater Air and Missile Defense Assessment document of July 2008. It clearly states that area defense will not defeat all the threats, thereby demonstrating that area air defense will not completely attrite all ASCM raids and that individual ships must be capable of defeating ASCM leakers in the self-defense zone.

- Use of manned ships for operational testing with threat representative ASCM surrogates in the close-in, self- defense battlespace is not possible due to Navy safety restrictions because targets and debris from intercepts pose an unacceptable risk to personnel at ranges where some of the engagements will take place. The November 2013 mishap on the USS Chancellorsville (CG 62) involving an ASCM surrogate target resulted in even more stringent safety constraints.

- In addition to stand-off ranges (on the order of 1.5 to 5 nautical miles for subsonic and supersonic surrogates, respectively), safety restrictions require that ASCM targets not be flown directly at a manned ship, but at some cross-range offset, which unacceptably degrades the operational realism of the test.

- Similar range safety restrictions will preclude manned ship testing of eight of the nine ASCM scenarios contained in the Navy-approved requirements document for the Aegis Modernization Advanced Capability Build 16 Combat System upgrade, as well as testing of the AMDR minimum track range requirement against supersonic, sea-skimming ASCM threat-representative surrogates at the land-based AMDR Pacific Missile Range Facility test site.

- To overcome these safety restrictions for the LHA-6, LCS, DDG 1000, LPD-17, LSD-41/49, and CVN-78 ship classes, the Navy developed an Air Warfare/Ship Self Defense Enterprise modeling and simulation (M&S) test bed that uses live testing in the close-in battlespace with targets flying realistic threat profiles and manned ship testing for other battlespace regions and softkill capabilities to validate and accredit the M&S test bed. It is ironic the Navy agrees this realistic testing is needed for the self-defense capabilities of its aircraft carriers, amphibious ships, DDG-1000, and LCS, but not for the DDG-51

destroyers. The irony arises because if the DDG-51s cannot defend themselves, they will not survive to provide the area air defense attrition of ASCMs absolutely necessary for all other surface combatants to have a hope of surviving.

***Army Networking Radios.***

- The Army has been working for more than a decade to develop and field a number of radios providing tactical voice and digital data communications at multiple echelons, including at company and platoon; the Manpack is one such radio. In April through May 2014, the Army conducted a Manpack follow-on operational test as part of the Network Integration Evaluation 14.2 at Fort Bliss, Texas. I assessed the Manpack radio as not operationally effective when employed in dismounted operations, operationally effective for mounted operations, and not operationally suitable. The higher frequencies at which newer data radios operate in order to provide greater throughput relative to existing radios limit the newer radios' range. This limitation is severe for dismounted units, which do not employ a sufficient number of the Manpack radios to assure the ubiquitous lines-of-sight among radios needed for many short-range "hops" to successfully transfer data between units located beyond immediate line-of-sight. The radios are also not suitable for dismounted operations because they are heavy and require a large number of batteries that must be carried by dismounted soldiers and frequently re-charged. Testing continues to reveal the numerous challenges associated with realizing the Army's goal of fielding a "battlefield Internet."

***Cybersecurity.***

- During 2014, cybersecurity testing of more than 40 systems showed improvements must occur to assure secure and resilient cyber capabilities. One important conclusion from my 2014 review of defense programs was that operational testing still finds exploitable cyber vulnerabilities that earlier technical testing could have mitigated. These vulnerabilities commonly include unnecessary network services or system functions, as well as misconfigured, unpatched, or outdated software, and weak passwords. Developmental testing over the course of the program, including the process to grant a system the authority to operate on defense networks, could have found most of these vulnerabilities; yet, such vulnerabilities are still found during Initial Operational Test and Evaluation. My review of these systems also identified the need to increase the participation of network defenders and assessment of mission effects during threat-representative, adversarial assessments.

- My office continues to emphasize the need to assess the effects of a debilitating cyber-attack on the users of these systems so that we understand the impact to a unit's mission success. A demonstration of these mission effects are often not practicable during operational testing due to operational safety or security reasons. I have therefore advocated that tests use simulations, closed environments, cyber ranges, or other validated and operationally representative tools to demonstrate the mission effects resulting from realistic cyber-attacks. Representative cyber environments hosted at cyber ranges and labs provide one means to accomplish the above goals. Such cyber ranges and labs provide realistic network environments representative of warfighter systems,

network defenses, and operators, and they can emulate adversary targets and offensive/defensive capabilities without concern for harmful effects to actual in-service systems/networks. For several years, I have proposed enhancements to existing facilities to create the Department of Defense (DOD) Enterprise Cyber Range Environment (DECRE), which is comprised of the National Cyber Range (NCR); the DOD Cybersecurity Range; the Joint Information Operations Range; and the Joint Staff J-6 Command, Control, Communications, and Computers Assessments Division. The need and use of these resources is beginning to outpace the existing DECRE capabilities. As an example, the NCR experienced a substantial increase in customers in FY 2014, and the Test Resource Management Center, which oversees the NCR, has initiated studies to examine new capabilities to further expedite the planning, execution, and sanitization of NCR events.

- Also in 2014, my office conducted 16 cybersecurity assessments in conjunction with Combatant Command and Service exercises. A notable improvement over previous years was the increased participation of higher-echelon computer network defense service providers and local defenders, resulting in a more comprehensive assessment of cyber defensive postures. Despite the improved defenses, my office found that at least one assessed mission during each exercise was at high risk to cyber-attack from beginner to intermediate cyber adversaries. I have placed emphasis on helping Combatant Commands and Services mitigate and reduce those persistent cybersecurity vulnerabilities observed from assessment to assessment. My continuing focus is on finding problems, providing information and assistance to understand and fix problems, and following up to verify cybersecurity status and ability to conduct operations in a contested cyberspace environment. At the request of several Combatant Commands, I have implemented more frequent operational site assessments during day-to-day operations on live networks to provide feedback on specific areas of interest such as status of patching or defense against specific attacks (e.g., phishing) and cybersecurity implications of physical security. Additional continuing efforts include working with the intelligence community to improve cyber threat realism, and to develop a persistent cyber opposition force with the capability to operate across several Combatant Commands.

- My report to the Congressional defense committees may be accompanied by such comments as you wish to make. I will provide copies to the Under Secretary of Defense for Acquisition, Technology and Logistics; the Vice Chairman of the Joint Chiefs of Staff; each of the Service Secretaries; and the Chairmen and Ranking Members of the Congressional defense committees.

COORDINATION: NONE

Attachment: TAB A

Prepared By: J. Michael Gilmore, Director, Operational Test and Evaluation, (b)(6)