



FY 2017 Annual Report

I was confirmed by the United States Senate November 19, 2017, and appointed by the President on November 21, 2017, as the seventh Director of Operational Test and Evaluation. It is an honor to serve in this position. I know from personal experience there are three imperatives in combat: Believe in yourself and your training; believe in your mission and commanders; and believe in your equipment and weapons. As the Director, Operational Test and Evaluation, I will provide independent and objective assessments so that our soldiers, sailors, airmen, and marines believe in their equipment and weapons, and are confident they are combat ready. I am committed to independently and objectively evaluating our systems to enable the Department of Defense (DOD) to make sound acquisition and deployment decisions. I will always be mindful of the taxpayer investments in our military and the priorities of the Secretary of Defense.

Most of the content in this report is based on tests conducted and independent evaluations completed before my tenure, but I have reviewed the content. In this introduction, I have contributed my own thoughts on future focus areas, the relevance of DOT&E, the importance of our workforce, and acquisition reform.

I submit this report, as required by section 139 of title 10, U.S. Code, summarizing the operational and live fire test and evaluation activities of the DOD during fiscal year 2017.

FOCUS AREAS

As I begin to shape my initiatives as the Director, my past experience, the emergence of new technologies, and the rapid evolution of threats suggest several key focus areas for the future. These areas include testing of software intensive systems and cybersecurity implications, integrated testing, test infrastructure, and modeling and simulation (M&S).

Software Intensive Systems & Cybersecurity

Today, the building material of choice for our weapon systems is software. The amount of software source lines of code in today's weapon systems is growing exponentially. Software does not just increase the functionality of these systems, it fundamentally defines the weapon system. However, as the number of lines of code increases so does the complexity of the system and cybersecurity vulnerabilities. The implications for T&E are profound. We are now making more changes that effect system capability through software than through hardware. For example, the F-35 Joint Strike Fighter's effectiveness in combat relies on software mission data loads, which work in conjunction with the avionics software and hardware to drive sensor search parameters. These files are critical for F-35 identification and correlation of threat and friendly radar signals. This increased dependence of system capabilities on software dictates that T&E must become a continuous, risk-based process for the life cycle of the system.

As weapon systems increase their dependency on software, the potential cybersecurity attack surface also increases. DOT&E has been a steady voice in the need to improve the cybersecurity posture of our systems, networks, and human interactions with networked systems. DOT&E has advocated for improved cybersecurity testing to identify critical problems and their operational impact and is currently funding the development of automated test tools. The cybersecurity section, later in this report, provides a number of recommendations to improve the Department's cybersecurity posture based on the past efforts of this office.

The cybersecurity of our weapons and networks needs increased attention. In support of that, the Department needs to evolve how we monitor our cybersecurity posture. The two-phase Cooperative Vulnerability and Penetration Assessment (CVPA) and Adversarial Assessment (AA) approach currently outlined in DOT&E test guidance is necessary to help inform the cybersecurity posture of DOD systems, but is not sufficient. This testing has greatly improved our understanding of cyber vulnerabilities, but in addition to dedicated assessments, DOD systems must be built to include technologies to continuously monitor cybersecurity, and automatically find and patch software vulnerabilities. Periodic assessments by Red Teams alone are not adequate, because the security of system software can change at any time due to operator errors, or adversary cyber-attacks. Red Teams are critical, but by themselves will never scale to meet the enormity of the cybersecurity challenge facing the Department.

One of my top priorities will be to update cybersecurity and risk-based testing guidance to reflect best business practices. Cybersecurity testing needs to move forward in the acquisition life cycle so that it can influence the system architecture from early development. I will advocate for additional resources for the development of automated software testing tools and the threat teams who use these tools. I will continue to advocate for rigorous cybersecurity testing and include evaluations of cybersecurity vulnerabilities in my assessments of systems. In the context of the rapid pace of software development, I will look for ways to align T&E activities with the velocity of the development of software systems.

FY17 INTRODUCTION

Integrated Test and Evaluation

I am supportive of previous efforts by the Department's T&E community to integrate testing, but they have not gone far enough. During my tenure, I plan to expand on those efforts. I know from experience there are many instances where operational and live fire evaluations can benefit from data acquired during developmental testing (DT) and where DT events can benefit from greater operational realism. Incorporating operational factors in DT&E and conducting early operational assessments aids in early discovery of problems and performance shortfalls. My office has often observed that operational testing identifies system performance problems that should have been identified in DT&E. The discovery is often due to bringing together the combination of operational users and realistic environments, missions, and threats for the first time. A more integrated approach could identify these issues early, when there is still time and resources available to fix them.

The implications of integrated testing for taxpayers and the warfighter are undeniable. We must look for better approaches to coordinate the planning of developmental and operational testing with the goal of accelerating our knowledge of system capabilities while reducing discovery later in the program.

I plan to update existing DOT&E guidance to incorporate an integrated testing philosophy. In my independent assessments, I intend to use all credible information to provide the warfighter and the Congress a complete understanding of how the systems the Department acquires will improve the readiness and lethality of our military forces.

Test Infrastructure

The Department needs T&E infrastructure for the five warfighting domains: air, land, sea, space, and cyber. However, much of our test range infrastructure is over 50 years old, with some assets built prior to World War II. Twenty-eight percent of our Major Range and Test Facility Base (MRTFB) facilities are in poor or failing condition, with an estimated cost to repair of over \$1.1 Billion. The majority of threats we have on the ranges do not represent the modern capabilities of our potential future adversaries. And, the once seemingly vast space of the open-air ranges is no longer large enough to test modern weapons and sensors at the employment distances envisioned. Test infrastructure for cyber is just now beginning to be realized, while the space domain remains in its infancy. We need to modernize our test ranges. As I stated in my confirmation hearing, I will visit major DOD test ranges early in my tenure to gain first-hand knowledge of their capabilities and limitations, and make recommendations accordingly.

An alarming trend over the past 10 years is that our potential adversaries are increasing their capabilities faster than the DOD test infrastructure can adapt and realistically represent them. The Department must accelerate the speed that threat capabilities are characterized and transferred to the test base. The test infrastructure of the future cannot just focus on open-air test ranges. The Department needs a strategy that incorporates software testbeds, software and hardware-in-the-loop facilities, anechoic chambers, open-air simulators, threat emulators, effects-based M&S, and open-air facilities. Open-air facilities need the ability to incorporate aspects of the virtual and constructive simulations to improve operational realism and span the full operational environment. As we develop infrastructure, particularly in the cyber and space domains, we must leverage virtual and constructive test environments.

The need to develop new test infrastructure quickly is important in the rapidly evolving areas of cyber threats and new software-enabled threat electronic warfare (EW) capabilities. Since 2010, DOT&E has used the yearly budget review process to advocate for resources to improve both cyber test capabilities and EW test range infrastructure to support realistic testing of modern combat systems. Notably, in 2012, DOT&E convinced the Department to invest nearly \$500 Million in the Electronic Warfare Infrastructure Improvement Program (EWIIP) to upgrade open-air test ranges, anechoic chambers, and reprogramming laboratories in order to develop and understand the performance of the F-35 and other advanced air platforms against advanced near-peer threat integrated air defense systems. I will monitor those investments to see they come to fruition as the Department rapidly approaches the start of the F-35 IOT&E.

Other significant T&E infrastructure shortfalls that DOT&E has highlighted routinely include: Fifth Generation Aerial Target; Self-Defense Test Ship; multi-stage supersonic targets; torpedo and submarine surrogates for anti-submarine warfare operational testing; the Warrior Injury Assessment Manikin for assessing force protection of ground combat vehicles to underbody blast events; range sustainability; and testing of space programs against offensive space threats. I will review the adequacy of the Department's T&E infrastructure to perform the full range of T&E responsibilities of Department weapons systems and equipment and advocate for improvements for any shortfalls I identify.

Improving the use of Modeling and Simulation (M&S)

Modeling and simulation is a critical element of test and evaluation. DOD acquisition programs are progressively more complex systems that support missions in increasingly complex environments. Programs often rely on M&S to fill data gaps when testing is either too expensive or not technically feasible. Programs can use M&S to provide insights on performance over the entire operational envelope even when testing is limited to a few strategic shots. Future T&E activities will undoubtedly increase their reliance on M&S tools, especially in the domain of space. This will require the acquisition and test communities to improve upon current M&S capabilities, including verification, validation, and accreditation (VV&A) of M&S assets.

FY17 INTRODUCTION

For programs that use M&S, program managers and Operational Test Agencies (OTAs) should design system T&E programs to collect adequate data to support the validation of those models. DOT&E issued guidance in March 2016 and January 2017 on ways to improve VV&A activities. VV&A activities should include a comparison of live test data to M&S runs coupled with a quantification of the uncertainty in such assessments. The DOD acquisition community should leverage emerging research methods from academia to improve the efficiency of VV&A activities, while ensuring the methods are scientifically sound. I plan to update DOT&E guidance on the use of M&S and the VV&A of such models to reflect my views on the importance of it in operational and live fire evaluations.

The Department needs to think about a wider application for M&S tools. For example, the Joint Technical Coordinating Group for Munition Effectiveness has initiated development of M&S tools for offensive cyber effects. These cyber effect tools are a non-kinetic threat parallel to the existing Joint Munition Effectiveness Manuals used by the weaponeers and mission planners. Cyber effects models will enable the Department to assess the survivability of our systems against adversary cyber threats. To support these modeling efforts, the Department should start generating data-based network models, threat characterization models, and models to predict the cyber effects for a range of target-weapon pairings. Similar to kinetic threats, such an approach would drive the materiel developers to design for survivability to the cyber threat, it would enable a more robust and quantified review of the system vulnerabilities and vulnerability mitigation features, and would enable a more phased or building block approach to survivability evaluation that includes component, sub-system, system, and full-up system-level testing.

DOT&E RELEVANCE TO THE DEPARTMENT

DOT&E's oversight enables the Department to deliver weapon systems that work through adequate testing. In FY17, DOT&E approved 35 Test and Evaluation Master Plans (TEMPs) and 95 test plans. DOT&E's independent assessments provide objective information to the military Services describing what works and what does not work, as well as provide recommendations for improvement. This objective information informs acquisition and fielding decisions that result in a more lethal force. In FY17, DOT&E provided 46 independent assessments for the Department and the Congress. DOT&E's contributions go beyond the benefits to specific programs. DOT&E's contributions and their impact to the larger DOD community from 2017 include:

- DOT&E improved the DOD cybersecurity posture using threat-representative cyber Red Teaming of Combatant Command networks during 12 major exercises and cyber readiness campaigns. The DOT&E find-fix-verify cybersecurity assessment program approach has improved the ability of Combatant Command network defenders to withstand realistic cyber-attacks and maintain their critical missions. The success of this program resulted in three Combatant Commands instituting permanent cyber Red Team operations on their live operational networks that will continually monitor and improve their cybersecurity posture.
- DOT&E funded improvements to M&S tools to better quantify system survivability. The Joint Live Fire (JLF) program worked to expand the validation of several widely used vulnerability M&S tools and improved the ability of those tools to support the assessment of system survivability. DOT&E's JLF program funded projects that will inform programs on the factors that most affect system vulnerabilities as well as the uncertainty inherent in those predictions.
- DOT&E improved test infrastructure for testing fifth generation systems. DOT&E is leading the development and testing of the Fifth Generation Aerial Targets (5GAT). This target will support operational and live fire testing of advanced weapon systems against low-observable targets. Testing against these targets will inform warfighters on how their weapons will work against advanced adversaries. These relatively low cost targets are currently meeting all early performance and cost goals.
- DOT&E collaborated with international partners to improve testing efficiency by sharing test venues and infrastructure. For example, DOT&E developed and fielded weaponeering tools in support of U.S. Forces Korea/Combined Forces Command to have effective target planning, munitions requirement development, and weapon procurement analysis. These efforts received the attention of General Jeong Kyeong-doo, Chairman of the Republic of Korea Joint Chiefs of Staff. Additionally, DOT&E collaborated with Israel on Army network modernization. The Israeli Defense Forces invited DOT&E and Army personnel to observe the Ground Forces exercise Light of Dagan, Israel's largest military exercise in 19 years.

One way that DOT&E can provide more relevant information to the Department is by proactively encouraging integrated testing. Including operational factors in developmental testing will help identify problems earlier, when they can still influence system design. In cases where systems perform well, programs should be able to take credit for early data, reducing the required resources in IOT&E. While I cannot direct developmental testing, I plan to have my staff engage early to look for opportunities to integrate testing with the goal of facilitating early learning and reducing the overall testing required of systems, when systems perform well. One example of how I plan to be flexible with integrated testing is my engagement with F-35 test stakeholders. I am working with them to allow approval of early test events. We are currently engaging with the test team to approve pre-IOT&E activity for cold weather testing in early 2018, months prior to the official IOT&E. If the program meets system development milestones, I will consider additional pre-IOT&E activity. This early testing will reduce the data required during the formal IOT&E period, which will increase aircraft availability for the core IOT&E missions. Increased aircraft availability will reduce execution risk helping to complete testing on

FY17 INTRODUCTION

time. Additionally, because cold weather testing must occur during the winter, my approval of this pre-IOT&E activity may eliminate the need for a cold weather deployment in the middle of the dedicated IOT&E.

In 2018, I will continue to look for ways that DOT&E can use our operational and technical expertise to provide relevant and credible information to the Department and the Congress.

STATUS OF THE OPERATIONAL TEST WORKFORCE

As I mentioned during my confirmation testimony, one of my highest priorities is to assess the current DOT&E and Service Operational Test Agency (OTA) workforce. To adequately assess the operational effectiveness, suitability, and survivability of weapon systems, a skilled workforce must have a clear understanding of the current operational tactics, techniques, and procedures and the operational threats to units equipped with these systems. The workforce must understand the operational mission and the systems under test, and apply scientific, statistical, and analytical techniques to evaluate those systems.

Since 2010, DOT&E has seen reductions in our staffing from a peak in 2010 of 93 civilians, 17 military billets, and 66 contractors to 80 civilians, 17 military billets, and 28 contractors in 2017. By FY20, my staff will be reduced to 76 civilians, 14 military billets, and 28 contractors. DOT&E must maintain the right mix of expertise in both military operations and technical knowledge to independently evaluate a diverse range of systems. In FY17, there were 308 systems under DOT&E oversight; the number and diversity of these systems require a highly skilled workforce. I plan to evaluate the efficacy of the government and contractor mix in the office of the DOT&E and identify areas that may need to be complemented with individuals who are savvy with emergent technologies and current operational experience.

DOT&E also reviews the state of the overall OTA workforce. It is critical that OTA personnel have strong operational, scientific, and analytical expertise. The Services have reduced the OTA workforce during the last decade. The OTA workforce fell over 12 percent between 2006 and 2016, driven mostly by the loss of military personnel. Many of these losses are attributable to draw downs in the overall military. Nonetheless, the loss of military personnel with operational experience diminishes the ability of the OTAs to test and evaluate increasingly complex weapon systems. Since 2010, the OTA workforce has remained relatively stable at approximately 1,900 personnel. At the current level of staffing, my staff has observed that the OTAs sometime have to prioritize programs and have limited access to subject matter experts across the ranges of areas of expertise necessary to test complex military systems.

DOT&E also continues to have a concern about retirement-eligible civilians within the OTA workforce, which increased to 43 percent in 2016. The OTA retirement eligibility rates are well above the GAO predicted rates for both the DOD and overall Federal workforces that could produce mission critical knowledge gaps if left unaddressed (see U.S. Government Accountability Office report GAO-14-215, "Federal Workforce: Recent Trends in Federal Civilian Employment and Compensation," January 2014). Based on the most recent analysis completed this year, I will work with the OTAs to develop workforce strategies that:

- Monitor the number of military personnel supporting T&E so that operational expertise is not lost.
- Develop recruitment plans that prevent mission critical skills gaps from developing as skilled civilians retire and create a future workforce ready for the evolving needs of T&E. In response to evolving technologies, the OTAs should recruit individuals with cybersecurity, statistics, autonomy, machine learning, human factors, and M&S expertise.
- Collaborate on best practices for providing both educational opportunities to targeted members of the workforce and training to the broader workforce, potentially leveraging elements of each other's programs and DOT&E-sponsored training.

I will assess the adequacy of the OTA workforce over the next year and update this assessment.

DOT&E SUPPORT TO THE TEST AND EVALUATION WORKFORCE

The Department will continue to acquire sophisticated technologies and we need a test workforce that is well equipped to test those technologies. Advancing methods for T&E requires partnering with academia, industry, professional test societies, and other government agencies. To that end, my office has collaborated with former Deputy Assistant Secretary of Defense for Developmental Test & Evaluation (DASD DT&E)/Director, Test Resource Management Center (TRMC), National Aeronautics and Space Administration (NASA), and academia to support the T&E Workforce. Going forward, I will continue this collaboration with the future leadership of these organizations.

In collaboration with TRMC, my office funds the Science of Test Research Consortium, which develops new techniques, aids in the education of the T&E workforce, and provides an important link between academia and the T&E community. I look forward to expanding these research efforts to address evolving needs in cybersecurity and software testing. In collaboration with NASA, my office supports the Defense and Aerospace Test and Analysis Workshop (DATAWorks), which seeks to build a community around statistical approaches to T&E in defense and aerospace. I also support DASD DT&E's funding of the Scientific Test and Analysis

FY17 INTRODUCTION

Techniques Center of Excellence, which partners with major acquisition programs to develop scientifically sound test programs and has recently expanded to include expertise in software and cybersecurity testing.

DOT&E also supports the workforce by developing training materials and resources that are available to the wider T&E community. Resources are available on TEMP development, test planning, experimental design, software testing, reliability growth planning and testing, cybersecurity, statistical analyses, survey design and analysis, and M&S. Resources are forthcoming on the development of Live Fire Test and Evaluation Strategies.

ENSURING ADEQUATE OT&E AND LFT&E UNDER ACQUISITION REFORM

I am very supportive of acquisition reform efforts to streamline and improve the defense acquisition process to build a more lethal force. DOT&E welcomes smart reforms that make the acquisition system more efficient and allow the DOD to provide warfighters with timely new capabilities that work. The statutory responsibilities provided to this office are essential to ensure adequate and realistic testing. I will collaborate with the Department acquisition community to accelerate the speed of acquisition when prudent to significantly reduce the time to deliver war-winning capability to our soldiers, sailors, airmen, and marines.

I am supportive of efforts to improve testing through the increased use of M&S and integrated developmental and operational testing. I firmly believe that testing early and often will improve acquisition outcomes by informing better system design and providing decision makers with information on system capabilities as they are developed. I support integrated testing to demonstrate system capabilities early and inform what data are needed from operational testing. I will be mindful of the balance between M&S, integrated testing, and operational and live fire testing when approving the adequacy of operational test plans. Additionally, I will focus on ensuring that operational and live fire testing remains adequate, while supporting the intent of acquisition reform, to streamline the DOD acquisition process.

We need to revisit the best practices of competitive system prototyping incorporating new technologies to improve future acquisition outcomes. Competitive system prototyping, followed by live experimentation on the ranges and quick iterations on system design has helped weapon systems evolve quickly. Then, a final fly-before-buy period provided the government and the Congress with the assurances that they needed before fielding systems. On the B-2 program in particular, we used a maturity model to manage expectations and show progress towards acquiring full operational capability. At major milestones, we reviewed progress, compared to the system maturity model, and updated expectations for the future.

In order to improve future acquisitions, we need to integrate testing into the system engineering process starting with early system prototypes. We must acknowledge the degree at which software defines system capabilities and the rate of software updates in our testing philosophy. Rapid prototyping, including digital prototyping using M&S, should focus on rapidly developing stable hardware designs. Once the hardware is stable, using iterative incremental development supports rapid software co-development and testing, and subsequent incremental fielding of software-defined capabilities. As systems become more complex, integrate autonomous capabilities, and software defines system capability, we must engage in testing early and often to inform decisions and manage expectations.

CONCLUSION

I have spent my professional career preparing for the opportunity to be the Director of Operational Test and Evaluation. I have employed weapons in combat environments and have been both a developmental and operational test pilot. And, for the last 5 years, I have been immersed into the technical areas of software engineering and cybersecurity. I will use these experiences to establish a professional core of operational testers that will provide equipment and weapons that are effective, suitable, and survivable in combat. We owe our warfighters and taxpayers nothing less.



Robert F. Behler
Director

FY17 INTRODUCTION