NOT FOR PUBLIC RELEASE
UNTIL APPROVED BY THE
COMMITTEE ON ARMED SERVICES
U.S. SENATE

STATEMENT

BY

RAYMOND D. O'TOOLE, JR.

ACTING DIRECTOR, OPERATIONAL TEST AND EVALUATION

OFFICE OF THE SECRETARY OF DEFENSE

BEFORE THE

SENATE ARMED SERVICES COMMITTEE

READINESS SUBCOMMITTEE

NOT FOR PUBLIC RELEASE
UNTIL APPROVED BY THE
COMMITTEE ON ARMED SERVICES
U.S. SENATE
SASC – APRIL 28, 2021

Raymond D. O'Toole, Jr.
Acting Director, Operational Test and Evaluation (DOT&E)
Office of the Secretary of Defense

Chairman Kaine, Ranking Member Sullivan, and distinguished Members of the Committee –

Thank you for this opportunity to discuss the performance of Department of Defense (DOD) acquisition programs and acquisition reform. This is my first appearance before this Committee and it is an honor to be here to testify with Ms. Stacy Cummings, who is performing the duties of the Undersecretary of Defense for Acquisition and Sustainment, and Ms. Shelby Oakley from the Government Accountability Office.

DOT&E's Role and Perspective

As specified in Title 10 of the U.S. Code, DOT&E provides independent oversight of operational and live fire test and evaluation of DOD acquisition programs. Test and evaluation (T&E) is critical to the acquisition process: It assesses a system's operational performance and identifies system issues, offering program leads the opportunity to correct them before the final acquisition or fielding decision is made.

DOT&E is tracking 234 acquisition programs across the Department, which does not account for highly classified programs. Among the competing priorities of program cost, schedule, and performance, DOT&E is focused on delivering an authoritative assessment of system performance in combat. To do this, we ensure that the test is conducted in operationally realistic and representative conditions with trained operators,

in a mission-ready system configuration, and with representative threats; and that the test is comprehensive enough to capture the factors that may affect credible assessment of operational effectiveness, suitability, survivability, and/or lethality in theater. Our findings inform acquisition decisions and help our military forces understand good and bad aspects of their system's performance so that they can plan and execute their mission within that context. For programs under DOT&E oversight, we provide our assessment of the results of operational testing to the Secretary of Defense and Congress, in accordance with Title 10.

Attributes and Practices That Promote Program Success

Every acquisition program is unique but some key attributes can help influence whether a program succeeds, including technical complexity and maturity, resource availability, contract strategy, and the skills of the government and contractor personnel associated with the program. Based on DOT&E's evaluations of a wide range of DOD programs, I offer three insights on how acquisition program managers can achieve better outcomes and provide timely delivery of the required capability. Program managers should understand: (1) the value of T&E, which is critical to determining mission performance; (2) the value of integrating developmental and operational T&E, which enables earlier discovery of problems; and (3) the value of credible modeling and simulation (M&S) to augment and enhance, and in some cases replace, traditional "live" testing. I will illustrate these insights with a few examples from acquisition programs that have either embraced these principles or set them aside.

Understanding the Value of T&E

A good program must start with a realistic baseline of cost, schedule, and performance to ensure enough margin to adapt as the program evolves. In this balancing act, operationally realistic T&E is essential to understand the performance of the unit equipped with that system. T&E is the only way to demonstrate system performance, to include mission effectiveness, suitability, survivability, and lethality, prior to fielding. When conducted early in a program's development and when adequately resourced across the acquisition cycle, operationally realistic T&E offers a unique opportunity for the program office to not only identify but also solve problems before the system matures. Early problem discovery may allow the program to better manage cost and schedule later in the process, when retrofits and problem solutions become more complex, expensive, and time-consuming to implement. Most importantly "fixing" problems early in the T&E process mitigates the risk of discoveries in operational test, the field, or, worse, combat.

The Amphibious Combat Vehicle program serves as a good example of prudent planning and the benefits of early, operationally realistic testing. The program office understood that T&E would identify problems, provided the resources required to solve those problems, and was well-positioned to respond to problems discovered in early, developmental and limited user tests that supported a successful Milestone C acquisition decision. Early understanding and correction of deficiencies led to improved operational performance, demonstrated in a successful Initial Operational Test and Evaluation, which supported an informed full-rate production decision.

On the other hand, the KC-46A aerial refueling tanker program was years late in delivering test aircraft to the Air Force due to several reasons, including inadequate schedule margin for early identification of deficiencies through T&E, followed by failure to rapidly develop and demonstrate deficiency solutions. Fortunately, the KC-46A program has improved. Last year, the vendor shifted from a position of "what's good enough" to "what's the best we can do", spurring development of a new remote visual system design critical for unrestricted air refueling. So far, it appears that the new subsystem – which is based on significant research and excellent technologies – will contribute to the tanker's eventually fulfilling its primary mission.

As cybersecurity threats become more ubiquitous and sophisticated, DOD's acquisition and T&E communities need to address cybersecurity more comprehensively. Unfortunately, some programs do not properly plan for cybersecurity assessments. More critically, due to poor system hardening against dynamic cyber threats, driven by lack of workforce cyber capacity, talent and tools within the program offices, virtually none of the programs assessed in FY20 were survivable against relevant cyber threats.

A good example of recognizing the importance of cybersecurity is the Ground-Based Strategic Deterrent (GBSD) program, which is the replacement for the Minuteman III Intercontinental Ballistic Missile program. To ensure an effective cyber defense for GBSD, the program manager is funding an integrated Mission Defense Team to provide overall security for the program, including cybersecurity, physical security, and nuclear safety. The program manager started building this team in parallel with early development of the rest of the program. This early cybersecurity capability, coupled with

early cybersecurity testing, increases the likelihood that cyber defenses will be ready to protect the GBSD program when it is deployed, although future GBSD cybersecurity testing will demonstrate the effectiveness and any potential shortfalls of this approach.

Understanding the Value of Integrated Test and Evaluation

Integrated test and evaluation (IT&E) begins with collaborative developmental, live fire and operational test planning and execution during early phases of the acquisition program. Involving operational testers and the intended system users in the earliest stages of program development and test planning helps to set the conditions for a successful operational test, to discover mission-relevant problems early, and to reduce the cost of fixing problems. When adequately planned and resourced, integrated T&E can increase T&E efficiency by eliminating unnecessary test redundancies, and enable leveraging of data and lessons learned across the acquisition cycle.

The AIM-120D Advanced Medium-Range Air-to-Air Missile program has directly benefited from early developmental and operational test integration. The test teams ensured AIM-120D test shots were relevant and useful for both developmental and operational test, shortening test timelines and mitigating the possibility of transferring undiscovered operational utility risk to the user. Despite initial delays due to technical challenges, the AIM-120D team has established an efficient and collaborative test battle rhythm that has generated significant improvements, accelerating the fielding of better capabilities to the warfighter.

While integrated testing continues to produce T&E efficiencies, it currently represents only a small portion of overall T&E activities within DOD. Moreover, much of the success of integrated testing is attributed to individual programs' establishment of integrated test teams. DOT&E has been working with USD(R&E) to advance the integrated T&E concepts, policy, and guidance needed to further leverage the potential benefits; additional changes may be necessary to fully support integrated T&E implementation. For example, effective integrated T&E requires mission-relevant, testable requirements that can be assessed in the context of mission outcomes throughout the acquisition cycle, rather than just technical specification requirements. Integrated T&E also requires sharing T&E-relevant data across the acquisition cycle; to do so, DOD must improve data collection processes, instrumentation, access to contractor data, and data storage approaches. While current collection and storage practices do not routinely facilitate such sharing of data, to include advanced data analysis and analytics, many programs achieve this in a more ad hoc fashion.

The Armored Multi-Purpose Vehicle (AMPV) program exemplifies the value of data sharing, even in its current manual instantiation. Data sharing between the test teams and the program office has been exceptional AMPV's testers understood the performance requirements and their rationale early, which allowed them to scope the test early; as a result, the final contract included test assets necessary to support all phases of testing. The exchange of data during operational tests also enabled the program to understand the significance of the problems identified by the Army Operational Test

Center and DOT&E in earlier operational T&E, which they were then more inclined to fix.

Understanding the Value of Credible Modeling and Simulation

Modeling and simulation (M&S) is necessary for development, integration, and mission-level evaluation due to the complexity of the systems DOD is acquiring, the increasing importance and difficulty of representing complex operating environments, and the growing sophistication of our adversaries' weapon systems. To have confidence in M&S-based evaluations, we must ensure that each M&S environment is supported by an independent and agile verification, validation, and accreditation (VV&A) process that uses credible and relevant data for accreditation.

The Tomahawk Weapon System (TWS) program recognized the value of adequately validated M&S and developed an M&S representation of the shipboard TWS computer and communication architecture. The program office committed to recurring validation of this M&S capability with live flight data, allowing M&S to be used to evaluate operational performance with high confidence. This resulted in the reduction of flight time and associated resource expenditures, which translated to significant cost savings compared to a test program that would have employed only live testing.

In some cases, independently accredited M&S provides critical supplemental data to evaluate a system's performance. For example, safety limitations preclude testing manned Navy surface ships' self-defense capability against some anti-ship cruise missiles. An adequate test campaign to evaluate various combat, radar, and weapon

systems against these threats requires live test data, a capable unmanned asset to support this live testing, and accredited M&S. The Navy currently does not have a well-defined strategy or funding to provide any of these three capabilities, creating an unacceptable risk in our ability to evaluate the operational effectiveness and survivability of future ships in combat.

Adaptive Acquisition Framework

The Adaptive Acquisition Framework consists of six Acquisition Pathways recently developed by USD(A&S) for use by DOD program managers. DOT&E, in coordination with USD(R&E), is developing the T&E guidance for the Adaptive Acquisition Framework to enable the T&E community to support the six Acquisition Pathways effectively without compromising the ability to characterize effectiveness, suitability, survivability, and lethality of our weapon systems.

My assessment of the effectiveness of the Adaptive Acquisition Framework is based on Middle Tier of Acquisition (MTA) and Software Acquisition Pathway programs. The MTA Pathway has been widely adopted by program managers and DOT&E currently oversees 28 MTA programs. Per the explanatory statement accompanying the FY21 appropriations act, USD(A&S) and the Service acquisition executives have approved certain acquisition programs to use "prototyping or accelerated acquisition authorities." In accordance with the same legislation, DOT&E is assessing the available test strategies for these programs for appropriateness and risk to test execution.

The Services use the MTA Pathway for a wide range of systems and warfighting capabilities. In some cases, the MTA programs modestly upgrade an existing system. In other cases, MTA programs, such as the Future Long-Range Assault Aircraft and the ORCA (Extra Large Unmanned Undersea Vehicle/XLUUV), provide advanced new capabilities via emerging technologies. Approximately 75 percent of MTA programs are used for rapid prototyping while others are used for rapid fielding.

The agile acquisition approach utilized by some MTA programs exacerbates some existing acquisition challenges. For example, MTA test strategies frequently lack well-defined resources to plan and execute operational testing, or to train operators, maintainers, and cyber defenders. Some lack the rigor typically required to demonstrate operational effectiveness, suitability, survivability, and lethality. Certain MTA programs have wisely incorporated integrated test approaches with rapid test-fix-test cycles but doing so has begun to stress the Service operational test agencies and developmental test organizations, to include relevant oversight organizations, which currently are not resourced, staffed, or trained for the continuous level of effort and reporting required by such approaches.

While DOT&E fully supports the MTA concept of faster acquisition and fielding in order to get capability to warfighters more quickly, MTA programs still need to be positioned to assess and demonstrate operational performance – what the system can and cannot do, and whether employment and unit tactics, techniques, and procedures can remediate system shortcomings. An adequate operational demonstration, or an otherwise tailored operational test, must be executed to provide an opportunity to "fly before you

buy" – with the operational user behind the proverbial wheel – before the initial production or fielding decision is made in order to mitigate risk to the user. Any increase in tolerance for performance risk in pursuit of acquiring emerging technologies must be characterized, if not quantified, in the context of the actual capability delivered to warfighters and their ability to win and survive wars.

Test and Evaluation Authorities, Responsibilities, and Capabilities

It is important that the same rigorous oversight DOT&E provides be applied to the earlier developmental T&E phases of a program. Certain acquisition programs have a strong DOT&E presence, with DOT&E providing oversight for 234 acknowledged programs. In contrast, USD(A&S) is the Milestone Decision Authority for 11 programs, providing oversight across the entire acquisition cycle. USD(R&E) provides Developmental Test, Evaluation and Assessment oversight of 11 programs, in accordance with previous Deputy Secretary of Defense guidance. Because initial operational testing represents a fraction of the overarching T&E program, and tends to occur at the end of a system's development cycle, there is an opportunity for A&S and R&E to provide more and earlier T&E oversight. This is especially true if we expect to take full advantage of adaptive acquisition, integrated testing, and early deficiency discovery and remediation, all of which can lead to faster and less costly development of more effective and survivable systems.

Program offices, in an effort to balance cost, schedule, and performance, are sometimes drawn to truncating developmental test efforts to maintain schedule or cost

objectives. Developmental testing may be cut short, or problems that developmental testing uncovers may be left unaddressed in order to keep the program moving forward. This recently occurred in the Bradley A4 Engineering Change Proposal program. Developmental testing had discovered indications that the system was overcharging turret batteries but the Army did not identify this as a fault or safety hazard and did not address it. Later in the program, operational testing identified a significant safety issue; the system overcharged the turret batteries and released hazardous toxic fumes into the crew compartments. Improved oversight of developmental testing likely would have prevented this problem from persisting until soldiers were exposed to a safety hazard during operational testing.

As discussed above, acquisition outcomes could be improved if the T&E community were positioned to more effectively leverage the benefits of integrated T&E. To support that, contracts should be negotiated to require operationally relevant, mission-level goals during developmental test, rather than focusing only on technical specification compliance. In addition, as the use of integrated T&E expands, it would be helpful to codify in the law, and otherwise enable inclusion of, operational test representatives in decisions regarding execution of developmental and integrated test events. On several occasions, DOT&E had intended to obtain data via integrated T&E or simply to use developmental test data, only to see the test event canceled without input from the operational test community.

The T&E community plays a large role in assuring test adequacy and shepherding programs to operational test success and, ultimately, fielding. As a result, the T&E

community needs to be equipped with state-of-the art tools and capabilities to meet emerging needs and the needs of the future. Earlier this year, DOT&E laid out a Science and Technology Strategy to provide a basic framework to guide T&E modernization and to keep up with changing weapon system capabilities – both ours and that of our adversaries. The strategy comprises five focus areas.

The first focus area is software and cybersecurity T&E. We are finding cyber issues and vulnerabilities in nearly every program we oversee. Given the volume and complexity of cybersecurity and software testing, it is clear that people-centric T&E approaches are not sufficient. Instead, the T&E community needs automated solutions for both testing and continuous monitoring of system cybersecurity and software. This needs to be fortified by a workforce trained and equipped to combat cybersecurity threats.

The second focus area is next-generation T&E capabilities. The quality of T&E – and ultimately warfighting capability – depends on the quality of T&E tools, infrastructure, and processes. DOD's T&E enterprise must be able to adequately assess emerging capabilities and threats, such as systems using artificial intelligence, space-based systems, and directed-energy and hypersonics programs – and must mirror real-world environments and scenarios. DOT&E recently commissioned the National Academies of Sciences, Engineering and Medicine (NASEM) to assess DOD's T&E capabilities and capacity, and to provide actionable recommendations to shape the Department's investment strategy over the next five to 10 years.

The third focus area is more widely instituting the integrated T&E lifecycle. DOD can make T&E more effective, and likely more efficient, by mitigating the adverse

effects of traditional contractor, developmental, and operational test silos. The segregated, serial approach should be replaced with a process that integrates all test phases – from contractor testing to developmental testing to operational testing – within a mission construct. This will require advanced tools and methods for designing test events that collect data that satisfy both developmental and operational needs across the acquisition cycle. As part of the integrated T&E lifecycle, we also must institutionalize inclusion of the intended users and testers in development of system specifications and contract requirements to ensure that they are operationally relevant and testable.

The fourth focus area is digital transformation. T&E must respond to industry's and adversaries' adoption of digital technologies and capabilities. T&E needs automated, even AI-enabled, data collection and analysis tools. We also must build easily shared – yet cybersecure – data repositories for better data analysis and analytics. In addition, more programs should incorporate credible digital twinning in their design and testing efforts. We need to prioritize the development of sophisticated modelling environments that undergo constant refresh and continuous agile verification, validation, and accreditation, as well.

The final focus area is workforce expertise and partnerships. T&E of complex technologies requires cutting-edge expertise. The ability to attract more talent to government service and to obtain consistent, on-demand access to experts from academia and industry is key. Equally important are more structured, rigorous, and continuous training programs to help the acquisition and T&E workforce meet future needs.

I appreciate the invitation to be here today and I would welcome the opportunity to meet in person or virtually with any member of the committee or your staff to talk further about the value of operational testing to the DOD acquisition process.