Introduction

There are three Imperatives of Combat. The first is “believe in your mission;” the second is “believe in your commanders.” For the operational test community, the third imperative holds special significance: “believe in your weapons and equipment.” Our soldiers, sailors, airmen, marines, and guardians, along with DOD leadership and the Congress, count on us to tell them when and where to place that faith. We must not let them down.

As we start the third decade of the 21st century, the United States remains the world's preeminent military power, thanks to our dedicated all-volunteer force, who are committed to their oath to support and defend the Constitution, and the civilians who stand beside and behind our women and men in uniform. Our Armed Forces' intellect, creativity, and countless hours of selfless service fuel America's successful national defense. Those unparalleled intangibles are backed by the technology the Defense Department puts in their hands, which, thus far, has given them the edge necessary to protect our homeland and our allies, and to advance the United States' strategic objectives.

The acquisition and testing communities are responsible for ensuring that this technology continues to provide American forces the decisive advantage they need. On the surface, the operational tester's job may appear simple: determine a system's operational effectiveness and suitability, and the survivability of the system and its operator, in the context of the intended mission. This succinct description belies the challenge in assessing a weapon or other technology in operationally realistic conditions – with the warfighters who will use it, in the expected physical environment, under the tactical conditions and battle plan anticipated, facing threats that accurately replicate our potential adversaries. As the operational test community knows, fulfilling that mandate was never simple and the future offers no respite. U.S. systems are growing more complex; our adversaries are becoming more sophisticated and capable; and joint multi-domain operations, encompassing land, air, sea, space, and cyberspace, are now the driving operating concept. The need to execute rigorous, credible OT&E has not lessened; in fact, it may be more critical than ever. Over the past year, competitors revealed technological advances that match and outpace our own, for instance, in hypersonic missiles. In November 2021, just prior to concluding four decades of service, then Vice Chairman of the Joint Chiefs of Staff General John Hyten remarked that "probably should create a sense of urgency." DOT&E couldn't agree more.

But where should that sense of urgency steer the operational test community? Concerns about being able to conduct proper OT&E are perennial. The Annual Report for Fiscal Year 2000 noted then that "Weapon technologies are outdistancing our ability to adequately test systems as they are developed." That statement remains accurate today. The high-volume wave of new technology in DOD's acquisition pipeline, the rapidly changing threat landscape against which we must evaluate it, and the need to field systems at the ever-quickening speed of relevance will strain or exceed our current infrastructure, tools, processes, and knowledge base.

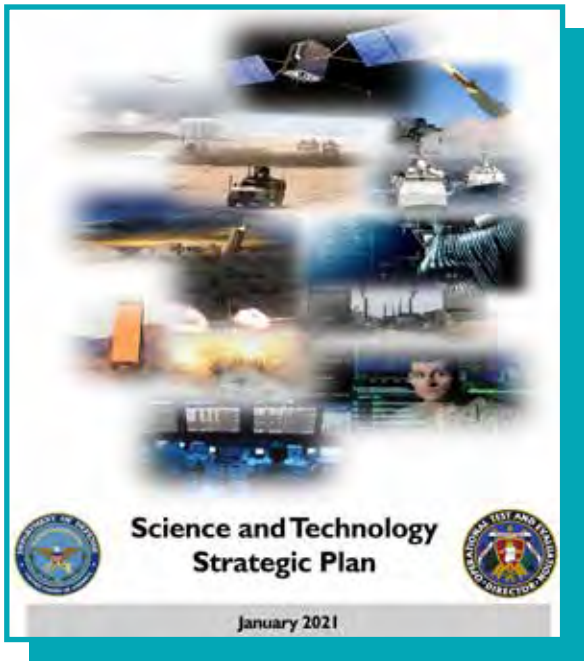
Some of the most frequently cited principles and means to improve acquisition outcomes and T&E efficacy and efficiency aren't novel, either. In 1995, then Secretary of Defense William Perry laid out five themes to guide the strategic direction for T&E. Four of them are equally valid now as they were 26 years ago: earlier involvement of operational testers in the acquisition process; more and more effective use of models and simulations; combining, where possible, different types of testing; and conducting operational testing and training exercises together. Quoting then Under Secretary of Defense (Acquisition, Technology, and Logistics) Jacques Gansler, the FY 2000 Annual Report also highlighted what is now known as the "shift left" mantra: "… serious testing with a view toward operations should be started early in the life of a program. Early testing against operational requirements will provide earlier indications of military usefulness. It is also much less expensive to correct flaws in system design, both hardware and software, if they are identified early in a program. Performance-based acquisition programs reflect our emphasis on satisfying operational requirements vice system specifications." These sentences could have been crafted today.

The nature of most organizations is to change incrementally – that is, to evolve – and the Defense Department is no exception. But the pace of evolution no longer is sufficient for national security writ large, nor operational test and evaluation in particular. Instead, to keep fulfilling our obligation to the warfighter, we need a T&E revolution.

Where the T&E Revolution Should Start

In January 2021, DOT&E released a Science and Technology Strategic Plan to help set the stage. A basic blueprint for operational T&E over the next five years, the S&T Strategic Plan has five focus areas.

Software and Cybersecurity T&E



Software and cybersecurity T&E lead the pack. The vast majority of DOD systems are extremely software-intensive. Software quality, and the system's overall cybersecurity, often are the factors that determine operational effectiveness and survivability, and sometimes lethality. The survivability aspect is especially critical. Many national security experts predict the next Pearl Harbor won't manifest as bombs destroying ships but as key strokes and hidden malware idling a fleet in home port or already at sea – an equally effective attack, with deniability, similar tactical results at lower cost for the adversary, and an unpredictable impact on public opinion due to the lack of visible carnage.

Now more than ever before, getting cybersecurity right on our weapon systems is essential to their actually being useful in the field. Warfighters, commanders, and program managers are relying on operational T&E to tell them what the cybersecurity risks, and their potential consequences, are, and to help them devise mitigation options to fight through a loss of capability. That means we must be certain that we understand the threat and can accurately emulate it during testing, and we can represent the entire attack surface, including the network and other platforms to which the system connects. The use of commercial technologies and services, such as cloud computing, adds another layer of risk to assess: are those commercial products, services, and their supply chains secure and suitable for military use?

The need for a sea change in cybersecurity OT&E is undeniable. The sheer number of systems that should undergo robust cybersecurity testing – that the Congress expects DOD to test – only intensifies that need. Cybersecurity testing must be accurate yet not endanger the operator. It must uncover whether the system is hackable and can be compromised, and what the impacts would be. Is the operator induced to make a bad choice based on spoofed system readings? Is certain offensive or defensive functionality lost, which, in turn, impedes individual or unit mission accomplishment? Or, does the platform shut down entirely?

Strengthening the engineering rigor of our testing is one place to start. We must expand cybersecurity T&E to examine whole-of-platform and systems-of-systems architectures and concepts of operations that reflect joint multi-domain operations. Broader use of automated testing methods, perhaps enhanced by artificial intelligence and machine learning, also is necessary; relying solely on people to conduct cybersecurity OT&E no longer is feasible due to the scale and scope of the testing requirement. Program schedules must accommodate an iterative approach to operationally relevant testing, with time and resources for test-fix-test cycles that begin with the minimum viable product and continue until, and perhaps beyond, a full deployment decision. The operational testing community, and DOD at large, will have to build a much larger and deeper bench of cyber expertise, both in house and outside the department to be tapped on demand, as well.

Getting cybersecurity principles right at the early stages of system design and development – long before operational testing begins – is a step the acquisition community can take to foster system resilience and posture the program for long-term success. Operational testers, and warfighters trained in offensive and defensive cyber operations, have the requisite knowledge. DOT&E is ready to assist any program office in incorporating the right cybersecurity principles that will give its platform the ability to respond to the continuously and rapidly morphing threat.

Transforming T&E to ensure cybersecurity across DOD systems and supporting supply chains will require a collaborative effort with our partners inside DOD, across the federal government, in industry and academia, and among our international allies.

Next-Generation T&E Capabilities

The quality of T&E – and ultimately warfighting capability – depends on the quality of the T&E tools, infrastructure, and processes we use. T&E must be able to handle whatever technologies are presented and it must mirror real-world environments and scenarios, to include accurate threat and countermeasure replication, in order to be thorough, operationally representative, and credible.

The T&E enterprise is not as prepared as it needs to be for the types of systems currently, or soon to be, in the development pipeline. The majority of the Department's open-air test and training ranges and laboratories are outdated and must be modernized to capture the complexities and capabilities of today's and future operational environments. We know already that artificial intelligence, autonomous and adaptive systems, space-based systems, directed energy, hypersonics, and biotechnology will challenge DOD's T&E capacity, facilities, and methodologies. To keep pace with the technological advancements expected on the modern battlefield, and to adequately test and train U.S. and coalition partner forces in complex and dynamic multi-domain operational environments, DOD requires significant and sustained investments in T&E infrastructure. The T&E Resources section of this report provides more detail regarding the critical T&E capability shortcomings that we must address to dominate the next conflict. We almost certainly will discover additional gaps as new technologies and operating concepts arise.

In late 2020, DOT&E commissioned the National Academies of Sciences, Engineering and Medicine (NASEM) to assess the ranges, infrastructure, and tools used for operational T&E. The main question was: Will the Defense Department be able to conduct the robust operational T&E our warfighters deserve on the systems and technologies anticipated in the 2025-2035 timeframe? NASEM completed and released to the public the first segment of that study in September 2021 and expects to finish the second segment, which is classified, in mid to late FY22. DOT&E will review the findings and recommendations from both portions to help inform DOD efforts to develop a holistic, enterprise-wide modernization and investment plan. With 2025 around the corner, DOT&E will press DOD stakeholders to begin implementation immediately.

DOT&E already is working on transforming one of the most important aspects of operational test and evaluation, and the acquisition process overall: data management. Capturing the right data, sharing them with the right people, and making the best data-driven decisions possible in a timely manner are fundamental to testing and fielding high-quality capabilities at the speed of need. The utility of data collected during all phases of T&E – contractor, developmental, integrated, and operational – can be much broader when analyzed as a whole, however. This vast quantity of data potentially could reveal trends in system design and performance, threat replication and emulation, test design and execution, program management, and other areas that would reshape DOD decision-making. But our ability to exploit that treasure trove is currently limited: the Defense Department lacks the means to aggregate, securely store, easily access and query, trace, and quickly analyze the cornucopia of test data it collects.

DOT&E recently engaged outside expertise to help revamp how we handle and use test data, and we would welcome information on other data analytics projects that are underway both inside and outside DOD. Our goal is to partner with other DOD stakeholders to start generating a draft data management capabilities and architecture blueprint.

Integrated T&E Lifecycle

The S&T Strategic Plan's third focus area is instituting an integrated T&E lifecycle. The concept isn't new, yet, to date, DOD has not fully implemented it. DOT&E believes that DOD's best avenue to improving T&E efficacy and efficiency is to greatly reduce, if not eliminate, the traditional contractor, developmental, and operational test silos. We need to replace that segregated, sequential approach with a process that integrates CT, DT, and OT to maximize test efficiency and effectiveness within a mission construct, whenever possible. In practical terms, that means designing test events to collect data that satisfy both DT and OT needs, when able. Additionally, program managers must involve the intended users and testers in developing system specs to ensure that

they're operationally relevant and testable; and contract language to ensure that testing requirements fulfill OT needs as early as possible and the right data are collected.

DOT&E currently is working with the Under Secretary of Defense (Research and Engineering) Developmental Test, Evaluation, and Assessments team to examine how we can expand the integrated T&E window. The goal is to enable agility, efficiency, and expediency. With that in mind, the operational test community must expand its partnership with DT and program offices to maximize integrated testing that provides operationally relevant data.

Digital Transformation

DOD is struggling to keep up with industry's and our adversaries' adoption of digital research and development and T&E capabilities. Besides the need for new tools and data management practices, perhaps the most critical shortcomings are in "digital twinning" and modeling and simulation (M&S). The test community acknowledged the need for M&S more than 20 years ago. That requirement has only become more urgent over time. For a variety of reasons, live operational testing in a threat-representative environment is not always feasible. When that occurs, we must have high-fidelity, operationally realistic M&S venues that produce enough high-confidence data to inform a determination of operational effectiveness, suitability, and survivability. These venues must be constantly refreshed and undergo continuous verification, validation, and accreditation (VV&A), particularly of the system under test and threats portrayed.

Sound VV&A, based on data collected during live (not simulated) events, is critical. The results of certain recent live operational tests diverged significantly from the outcomes predicted by M&S. Creating accurate, high-caliber M&S is a complicated endeavor but we must continue to invest in it and follow through with VV&A to ensure that our warfighters and commanders can trust operational T&E findings.

T&E Workforce: The Essential Human Element

The final focus area is the T&E workforce. T&E of complex technologies requires a tremendous amount of deep and broad cutting-edge expertise. DOD needs mechanisms both to attract more talent to government service and to obtain consistent, on-demand access to experts from academia and industry. DOT&E looks forward to working with DOD stakeholders, industry, and the Congress to improve T&E talent development, access, and management to ensure that the T&E community continues to provide outstanding support to the warfighter over the next decade.

Realigning DOT&E: Strategic Initiatives, Policy, and Emerging Technologies

To help set the conditions for T&E transformation, DOT&E initiated an internal reorganization last summer. In the spirit of integrated T&E, we folded Live Fire Test & Evaluation (LFT&E) functions and personnel into the warfighting domain divisions to better align our efforts. DOT&E's LFT&E expertise and oversight capacity remain the same; the LFT&E program will not be reduced.

To ensure that operational T&E is prepared to fulfill the warfighter's and the decision-maker's demands for credible, independent data and analysis, DOT&E has created a new division focused on the future. The Deputy Director for Strategic Initiatives, Policy, and Emerging Technologies (SIPET) will proactively look forward to identify OT&E needs, gaps, and potential solutions; craft new ways of doing business; and help Service and agency operational test organizations solve problems. Working with stakeholders across the department, SIPET also will develop and refine operational test policy guidance. The first areas SIPET will address include cybersecurity testing guidance and M&S VV&A guidance.

By dedicating personnel to the full-time mission of planning for emerging technology, digging into shared T&E challenges, and big-picture brainstorming, SIPET will foster greater agility and responsiveness in the operational

test community. The intent is that, as a result, we will create the conditions to "shift left" more often, more quickly, and with even better results than we achieve today.

DOT&E has realigned the Annual Report itself, as well. A new Executive Summary highlights major DOT&E products, contributions, and findings from this fiscal year.

Impetus and Way Ahead

Revolutionizing test and evaluation is within our grasp. It will take a concerted effort, and a steady and substantial flow of intellectual and financial resources – but we can achieve it.

Maintaining the status quo is not an option. The Defense Department's 2021 annual report to Congress on military and security developments involving the People's Republic of China noted that our primary pacing challenge "has substantially reorganized its defense-industrial sector to improve weapon system research, development, acquisition, testing, evaluation, and production." For the operational test community to fulfill its role as trusted, unbiased arbiters of a system's performance and its effect on mission accomplishment, DOD's T&E enterprise must stay ahead.