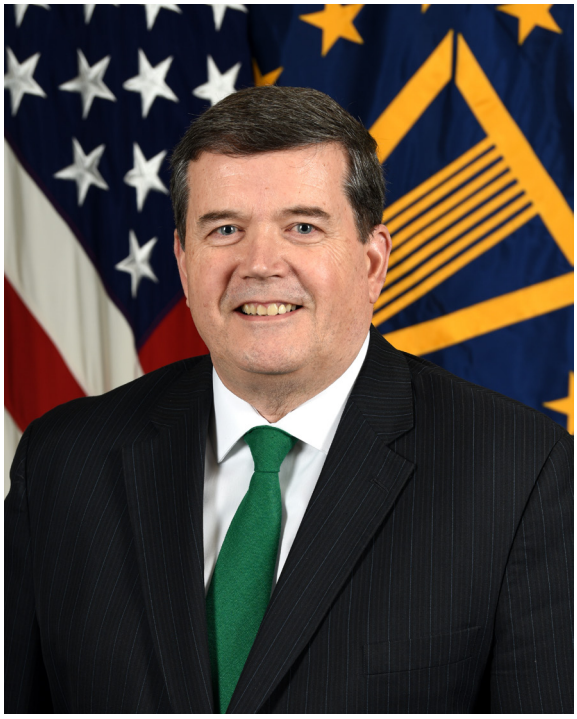
Director's Introduction



U.S. military Service members deserve to be equipped with combat-credible systems that meet the demands of today's evolving threats and tomorrow's unpredictable challenges. In a rapidly changing global security environment, where adversaries continue to develop sophisticated capabilities across all domains, our commitment to providing timely, rigorous, and independent evaluations has never been more critical. The U.S. military operates in an era of accelerating technological advancements, increasingly complex operational environments, and evolving global challenges. The DoD must rapidly and rigorously test and evaluate its systems to determine if they are not only operationally effective and suitable, but also survivable and lethal across contested domains.

To achieve these goals, DOT&E advanced our Strategy and Implementation Plan (I-Plan) this year. This plan lays out a clear path to make a strategic shift in test and evaluation (T&E) processes and builds on years of effort. These efforts started under Honorable Robert Behler, when he looked at the science and technology (S&T) of T&E and released the S&T Plan in January 2021. That work was codified under Honorable Nickolas Guertin in June 2022 with five strategic pillars and championed by Honorable Douglas Schmidt during his first six months tenure, after he was sworn in as Director, Operational Test and Evaluation on April 8, 2024. The I-Plan capitalizes on the latest advances in S&T to modernize our professional skillsets, enable our agility and efficiency, and inspire trust and confidence in system performance under wartime conditions. The plan's five pillars represent DOT&E's commitment to testing in ways that reflect operational realities, adapting to new technologies, and ensuring that our warfighters are equipped with the best possible capabilities.

CODIFYING STRATEGY IN POLICY AND GUIDANCE

This rapid pace of change from a technological, operational, and global geopolitical scale, demands that we must embrace a strategic shift in how we conduct T&E. Toward that end, in December 2024 Honorable Douglas Schmidt signed out a new policy for operational test and evaluation (OT&E) and live fire test and evaluation

(LFT&E). This policy – a DoD instruction accompanied by five corresponding DoD manuals – will drive the DoD forward by making strategic shifts in how we execute the T&E mission. I am honored to have played a role in implementing this culmination of a multi-year effort to update policy and guidance. These documents encourage early engagement from OT&E and LFT&E stakeholders and engagement in OT&E and LFT&E activities across the acquisition life cycle. The policy calls for:

- Using the latest advances in science and technology to both plan tests and evaluate outcomes,
- Requiring the integration of all relevant information into OT&E and LFT&E planning and assessment activities,
- Mandating the consideration of risk as part of the test planning process,
- Requiring that OT&E and LFT&E planning start in parallel with the initiation of the program,
- Considering the time and resources required to correct deficiencies identified in test, and,
- Ensuring T&E against the full domain of kinetic and non-kinetic threats to address the rapidly evolving threat landscape.

This policy and the corresponding manuals encourage us to lean forward in a measured way. We need to research, pilot, and inform how our future T&E practices leverage digital transformation, digital engineering models, and data collected from across the acquisition life cycle. The intricacies of software and artificial intelligence (AI) models make these practices imperative because the complexity they impose cannot be adequately tested in dedicated operational tests alone.

The I-Plan's goals, key challenges, and how the policy advances the current state of T&E is spelled out in each of the five pillars presented below.

Pillar 1 – Test the Way We Fight

Realistic operational conditions are the cornerstone of defensible T&E. Test conditions that systems are exposed to in operational tests routinely reveal new vulnerabilities and failure modes that should be remediated to avoid failure in combat. To reflect the evolving battlefield, there is a critical need for modeling and simulation (M&S) to undergo verification, validation, and accreditation (VV&A) with live data.

Pillar 1 focuses on developing T&E frameworks that reflect joint and coalition operations, particularly in highly contested and congested environments, such as cyber and electromagnetic spectrum (EMS) realms. DOT&E's new policy and guidance emphasizes the importance of using all credible sources of information to inform OT&E and LFT&E plans and assessments.

To improve our ability to test in operationally realistic environments, we are leveraging training and exercises for data. Retrieving data from theater assessments to support the continual VV&A of complex M&S is crucial to this effort's success. Early engagement by operational testers will ensure that program requirements incorporate testability considerations, so systems are instrumented to provide this critical data. Another challenge is testing in commercial cloud environments to ensure our cyber capabilities are robust enough to withstand attacks while adapting to new operational and economic realities.

In early 2024, I signed out the long anticipated F-35 Initial Operational Test and Evaluation (IOT&E) report as we concluded a multi-year test program. I shared the results with DoD leaders and Congressional defense committee members, informing the full-rate production decision for this $2 trillion program. The development of the Joint Simulation Environment (JSE) enabled adequate testing of the F-35. JSE enabled realistic scenarios with high-density threats that could not be completed in open air. Its development was critical to an adequate OT&E for the F-35 program and was championed by DOT&E. It now provides a great asset to the T&E enterprise for future F-35 and other aircraft testing.

This year I signed a memorandum of understanding (MOU) with Japan, allowing T&E projects between the two countries. This achievement and our ongoing efforts with the F-35 program illustrate how we

are "testing the way we fight." We will continue to champion testing of systems as they will be deployed in conflict through partnerships, advocacy, and other investments.

Pillar 2 – Accelerate the Delivery of Weapons That Work

In an era of rapid development, our ability to identify issues early in the life cycle of systems is crucial. Our new policy encourages shifting mission realism left and conducting OT&E and LFT&E activities throughout an acquisition program's life cycle. By pushing more test activities into earlier phases of system development, we reduce the likelihood of discovering problems late in the process, when fixes are more expensive and time-consuming.

Digital models are a key element of shifting left, but they must be coupled with live data collection to validate and update models to reflect operational realities. We are progressively shifting towards automated test tools for operational data collection and the investment in more agile processes aimed at reducing the T&E timeline while maintaining high standards of performance.

DOT&E worked with the maritime autonomous system community, and other communities utilizing emerging technologies, to identify test infrastructure requirements and develop test methods this year that will accelerate evaluation and contribute to operator confidence when applying these technologies in modern warfare.

Pillar 3 - Improve the Survivability of the DoD in a Contested Environment

Systems today must operate seamlessly in increasingly contested domains, including cyber, the EMS, and space. Building and testing systems that are resilient to these threats while also integrating them into larger networks and federated systems is hard. A key element of our new policy is testing for full-spectrum survivability and lethality, which is an integral part of the modern battlefield.

In future conflicts, the DoD and our partners and allies will face significant threats in cyber and congestion in the EMS. The advantage in these conflicts will accrue to whichever side can fix and improve their software most rapidly and reliably. The ability to rapidly reprogram, recode, and minimize total system downtime and integration downtime will be key, and must be tested intentionally and explicitly to ensure the survivability of DoD systems *and* kill webs. As we work through the implementation of the new policy, we will undoubtably uncover additional needs for digital models and live, virtual, constructive range infrastructures.

This year, DOT&E completed operational testing of the Mounted/Dismounted Assured Positioning, Navigation, and Timing Systems (MAPS/DAPS), which are among the Army's first programs designed to provide assured position navigation and timing to tactical units in a contested and congested electronic warfare environment. DOT&E also began oversight of programs that provide both military and intelligence capabilities, to help ensure these programs are operationally effective, suitable, and survivable in the hands of military personnel under representative combat conditions.

Pillar 4 – Pioneer T&E of Weapon Systems Built to Change Over Time

We no longer have the luxury of developing static, one-time-use systems. Many of our new capabilities, particularly those involving AI and software, will evolve over time. Testing for these systems requires a shift in mindset – we must embrace testing across the acquisition life cycle, even into operations and sustainment. We have to accelerate using adaptable test processes and methods that reflect prior knowledge. As a result, DOT&E is pioneering new testing methods for systems that will be updated and modified continuously throughout their life cycle.

Even when every process is followed correctly, the complexity of systems can lead to significant delays. However, by aligning our testing strategies with the evolving nature of these technologies, we can reduce the risk and ensure that these capabilities reach the warfighter when they are needed.

To support the new OT&E and LFT&E policy, Honorable Douglas Schmidt signed out the first ever DoD manual on testing AI-enabled and autonomous systems, which provides an initial starting point for the DoD to develop best practices for testing AI.

Pillar 5 – Foster an Agile and Enduring T&E Workforce

Our workforce is the backbone of the T&E enterprise. However, we are seeing increasing difficulties in recruiting and retaining the talent we need. As DoD systems become more complex, our workforce must be equipped to handle these advancements, from AI-based assessments to cybersecurity.

Many of these positions – especially in software development for operational testing – require highly sought after expertise in the private sector. To compete with companies like Google or Amazon for the best talent, we must find innovative ways to incentivize public service.

DOT&E is proud to partner across the DoD in developing innovative programs to educate and recruit the next generation of the T&E workforce. For example, we partnered with the Office of the Secretary of Defense (OSD) on the Pathfinder Program this year to develop next-generation cyber-T&E talent. We also partnered with the Office of the Under Secretary of Defense for Acquisition and Sustainment (OUSD[A&S])'s Defense Civilian Training Corps (DCTC), which focuses on developing the next generation of acquisition and T&E experts. DOT&E also continually provides growth opportunities for our current workforce, including implementing Learning Journeys to empower our workforce to develop new areas of expertise.

Another milestone this year that aligns with many of our I-Plan's pillars was the choice by the current Operational Test Agencies' (OTAs') commanders to re-validate their commitment to their six Core Test Principles: Early Operational Test Involvement; Tailor to the Situation; Continuous and Cumulative Feedback; Streamline Processes and Products; Integrated and Combined Collection/Test; and Adaptive. The OTA commanders recently signed a new memo supporting these test principles. DOT&E also continues to support these principles, which were originally captured in 2019.

DOT&E CONTRIBUTIONS TO THE T&E ENTERPRISE

The value of operational and live fire testing goes beyond informing a single decision point. We live in an era where data drives operational decisions through AI and automation. We must carefully consider how our mission needs to evolve to support the DoD at large and of course ultimately the Service members that execute its mission.

Several opportunities exist. By partnering with the acquisition community and developing requirements for testability, DOT&E can improve T&E automation and take advantage of other venues for T&E, like training, large-scale exercises, and even operations. The need for T&E is not shrinking and if anything is growing to handle the complex threat space. Unfortunately, we cannot realistically continue to grow our T&E workforce to scale given various programmatic and pragmatic limitations.

As Honorable Douglas Schmidt said, "as a researcher, I know the potential that AI can have for our T&E processes and practice, including Generative AI (GenAI), which can create certain types of images, text, videos, and other media in response to prompts." DOT&E is closely tracking today's technology transitions and leveraging them to continue improving and scaling our practice of rigorous T&E. We are also exploring how combining GenAI with live data capture can help provide more robust test data sets. Coupled with templates, large language models (LLMs) can also accelerate our analysis and communications.

SUCCEEDING THROUGH TEAMWORK

Ensuring essential operational properties of DoD systems is a team sport since DOT&E by itself can't simply "test our way to success." Moreover, the new policy and initiatives cannot be accomplished by DOT&E alone. We therefore need a holistic life cycle

view on how we acquire the best capabilities for the DoD. This vantage point requires us to collapse conventional stovepipes and partner across the DoD to succeed.

WAY AHEAD

As mentioned above, T&E is a team sport. Here are some ways DoD can work together to help to deliver weapons that work, faster:

- We must enhance the cyber resilience of products and information technology infrastructure by maintaining awareness of cyber threats, vulnerabilities, and intelligence; enforcing industry best practices for cyber defense; securing our software development environments; contracting for best security features in the cloud; and understanding and addressing supply chain vulnerabilities and dependencies.
- Initial T&E activities usually start with contractor testing and then transition to developmental testing, and ultimately, to operational testing. However, the more all this testing can be conducted in a representative operational environment, under realistic operational conditions, with as much actual threat information as possible, the more accurate and useful the results will be.
- We must continue to innovate by enabling more effective digital-physical fusion using live, virtual, constructive training environments; digital engineering and digital twins; and uncertainty quantification. These capabilities transform and enhance the value proposition of T&E by shifting the focus from what is required to what delivers the most decisive military advantage.
- Finally, we must hasten the adoption of measures that enable AI and machine learning in weapon performance evaluation, such as expanding data warehouse capabilities; automating data collection and large-scale analytics, processes, and administrative activities; and developing methods and tools that can leverage generative augmented intelligence at scale for T&E.

In an era where time, expertise, and resources are limited, our I-Plan's five pillars and their desired outcomes, coupled with our new policy, offer a solid roadmap for how we will address these challenges. We have already laid the groundwork for the T&E community to innovate and lean forward leveraging new technologies. I am confident this community can work together to develop new and innovative practices that couple new technologies with the rigorous methods and tools we have always leveraged, to ensure DoD systems are adequately tested before fielding. Together, we will ensure that the systems we deliver to our warfighters are tested thoroughly, efficiently, and in ways that reflect the future battlefield.

On January 10, 2025, I became the Director, Operational Test and Evaluation (Acting) for the third time. It is a great honor to serve as the senior advisor to the Secretary of Defense on OT&E and LFT&E of DoD weapon systems. DOT&E is immensely grateful to Congress for their continued support, and for encouraging us to innovate. Our global allies and partners in T&E transformation, and DoD liaison officers deserve our gratitude as well. Thank you to the DOT&E staff and our warfighters for their dedication to working as a formidable team. As a result of all of your efforts through hard work and timeliness dedicated to service, I am confident DOT&E is well-positioned to work successfully with the new administration and the new Secretary of Defense to continue to defend our nation.

Dr. Raymond D. O'Toole, Jr.
Director (Acting)

This page intentionally left blank.