EXECUTIVE SUMMARY

This page intentionally left blank.

MAJOR PRODUCTS

In FY23, DOT&E designated 45 new DoD systems for OT&E and LFT&E oversight and removed 22 systems from the T&E Oversight List. As of September 2023, DOT&E had 266 DoD systems on the T&E Oversight List for OT&E and/or LFT&E. In FY23, DOT&E:

- Reviewed and approved 32 T&E strategies/ Test and Evaluation Master Plans (TEMPs) and disapproved 1 TEMP.
- Approved 59 individual test plans and disapproved 1 test plan.
- Published 46 reports, including 34 reports on the independent evaluation of test adequacy and operational performance of DoD systems and 12 reports on cyber assessments or response to congressional taskers.

DOT&E completed 4 of the 14 assigned congressional taskers and is on track to complete the remaining tasks summarized in Table 1 in accordance with the agreed-upon timelines between DOT&E and Congress.

Table 1. Summary of DOT&E Congressional Activities

Source	Title	Status
FY22 NDAA		
*Sec. 115	Limitation on availability of funds pending report on the Integrated Visual Augmentation System	Complete
*Sec. 223	Development and implementation of digital technologies for survivability and lethality testing	Complete
Sec. 1529	Demonstration program for automated security validation tools	Ongoing
Other FY22 Congressional Taskers		
SASC Report pg. 191-192	Electronic Health Record interoperability between DoD and Veterans Affairs	Ongoing
FY23 NDAA		
Sec 217	Competitively awarded demonstrations and tests of electromagnetic warfare technology	Ongoing
Sec 242	Study and report on sufficiency of operational test and evaluation resources supporting major defense acquisition programs	Ongoing
Sec 1514	Operational testing for commercial cybersecurity capabilities	Ongoing
Sec 1656	Persistent cybersecurity operations for ballistic missile defense systems and networks	Ongoing
Other FY23 Congressional Taskers		
*Omnibus	Certification of funding for test infrastructure and test event resources	Complete
*Omnibus	Certification of test strategies on Middle-Tier Acquisition and Rapid Prototyping programs	Complete
HASC Report pg. 77	Assessment of contractor-provided test and evaluation capabilities	Ongoing
HASC Report pg. 77	Battery testing infrastructure	Ongoing
HASC Report pg. 77	Development and testing of body-worn equipment	Ongoing
HASC Report pg. 78	Equipment shortfalls within the test and evaluation community	Ongoing

Acronyms: HASC – House Armed Services Committee; NDAA – National Defense Authorization Act; SASC – Senate Armed Services Committee; USD(A&S) – Under Secretary of Defense for Acquisition and Sustainment; USD(R&E) – Under Secretary of Defense for Research and Engineering

* These activities resulted in reports to Congress in FY23 which are reflected in the Appendix.

DOT&E published a DOT&E Strategy Implementation Plan (I-Plan) formally endorsed by the Under Secretary of Defense for Research and Engineering, the Under Secretary of Defense for Acquisition and Sustainment, and the Military Service Secretaries. The I-Plan outlines key actions and deliverables intended to contribute to the transformation of T&E infrastructure, tools, processes, and workforce in response to emerging changes in acquisition, technology, and warfighting.

DOT&E drafted the DoD Instruction for OT&E and LFT&E and the following DoD Manuals: (1) TEMP/ T&E Strategy; (2) Modeling and Simulation (M&S) Verification, Validation, and Accreditation (VV&A) for OT&E and LFT&E; (3) OT&E and LFT&E of Software; (4) OT&E and LFT&E of Artificial Intelligence (AI)-Enabled and Autonomous Systems; and (5) Full-Spectrum Survivability and Lethality T&E. These policies are intended to: (1) enable OT&E and LFT&E stakeholders to inform acquisition contracts; (2) optimize the use of all data, intelligence, and program artifacts across the acquisition life cycle; and (3) increase OT&E and LFT&E efficiency and agility by increasing the use of digital engineering, digital tools, and technologies (e.g., M&S, model-based engineering, smart documentation, data repositories, data analytics, modern predictive analytics tools using AI and machine learning).

OT&E AND LFT&E OVERSIGHT OF DoD SYSTEMS

» ENSURED ADEQUATE OT&E AND LFT&E PLANNING AND EXECUTION

In FY23, DOT&E evaluated the adequacy of the TEMPs or T&E strategies, and OT&E and LFT&E plans based on the degree that they will provide: (1) data to support credible evaluation of operational effectiveness and suitability; (2) coverage of the battlespace and threats; (3) adequate verification and validation of M&S; (4) complete assessments of system survivability and lethality against mission-relevant kinetic and non-kinetic threats; (5) production-representative test articles; (6) operational realism; and (7) sufficient funding required to support test execution.

In FY23, DOT&E approved all but one TEMP and all but one test plan. Common DOT&E pre-requisites for approval included improving: (1) testing of the supply chain and inclusion of all potential attack vectors in contested cyberspace; (2) coverage of the operational environment and threats; (3) testing of all possible system variants; (4) M&S verification and validation (V&V); (5) use of latest software versions; and (6) data collection processes or equipment to support an evaluation of operational effectiveness, suitability, survivability, and lethality (as applicable).

In FY23, for the 34 programs that executed OT&E and/or LFT&E, DOT&E assessed 70.5 percent (24 of 34) as adequate, 20.5 percent (7 of 34) as partially adequate, and 9 percent (3 of 34) as not adequate, as shown in Figure 1. By comparison, over the last 7 years (FY16-22), DOT&E assessed 66.5 percent (143 of 218) of the executed OT&E and LFT&E as adequate, 25 percent (54 of 218) as partially adequate, and 9.5 percent as not adequate. The inadequacy or partial adequacy of OT&E and LFT&E were caused by: (1) lack of operational testing prior to early fielding, (2) delays in testing and early test termination, (3) lack of production-representative hardware or software, (4) data collection shortfalls, and/or (5) lack of testing under all pertinent threats and conditions. For the 34 programs that executed OT&E and/or LFT&E, DOT&E highlighted at least one test limitation in 26 of them, including but not limited to:

- Lack of assessment of all relevant cyberattack vectors or paths due to limitations imposed to protect the system from damage and ensure operator safety.
- Lack of access to production-representative hardware, software, supporting systems, and relevant documentation during early tests.
- Lack of an available threat simulator or surrogate.
- Lack of an environment that replicates the most challenging scenarios and limited or no testing in a contested electromagnetic environment.
- Inadequately validated or accredited M&S, or M&S results that did not capture all operationally

relevant conditions or did not model all important interactions.

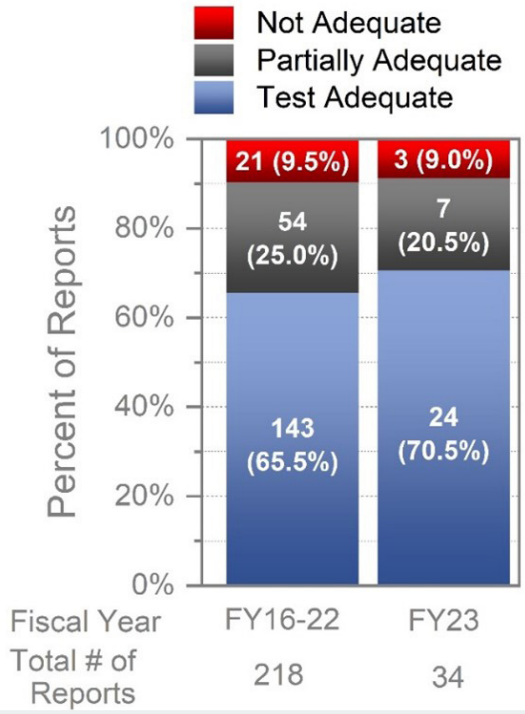

Figure 1. Test Adequacy in FY23 and Prior Years

Test Adequacy Recommendation Trends

DOT&E reports included recommendations for improving test adequacy including but not limited to:

- Completing testing with production-representative assets before fielding. A large fraction of test adequacy problems result from incomplete testing prior to fielding.
- Conducting operational testing across the relevant missions sets, operating conditions, and threats. Execution of robust testing continues to reveal important shortfalls that can be addressed prior to fielding.
- Testing all relevant cyberattack paths.
- Evaluating system suitability and cyber survivability early in the design to increase test efficiency, discover problems early, and improve outcomes in OT&E and LFT&E.
- Developing robust and independent V&V for all M&S to be used in OT&E and LFT&E.

Programs Pursuing the Middle Tier of Acquisition Pathway

In FY23, for the 97 programs approved by the Service Acquisition Executives to pursue Middle Tier of Acquisition pathways, DOT&E received and reviewed 55 test strategies and certified 43 of those to be appropriate. Test strategies were not certified as appropriate primarily due to inadequate resources for OT&E and/or LFT&E to evaluate the required capability in an operationally representative contested cyberspace and contested, congested, and constrained electromagnetic spectrum environments.

Adequacy of Resources for Programs with Approved TEMPs or T&E Strategies

In FY23, DOT&E assessed the adequacy of OT&E and LFT&E resources required to execute the agreed upon OT&E and LFT&E, scheduled in the current year and future years defense planning. This assessment could only be made for those programs on the DOT&E oversight that had approved TEMPs or T&E strategies.

- Fifty-five percent (72 of 131) of the eligible programs were assessed to have adequate funding to support the remainder of the planned test execution. Five percent (7 programs) were identified as having funding shortfalls, while 17 percent (22 programs) required updated TEMPs or T&E strategies due to program changes that may require new or altered testing or resource requirements. Eleven percent (15 programs) have fully executed all required testing and require no current or Future Year Defense Program funding. Fifteen additional programs were not assessed despite being eligible for this assessment because funding data were not provided by the Services.
- The identified OT&E and LFT&E resource shortfalls required to support adequate testing were primarily related to: (1) threat and target representation in contested environments; (2) representative digital representation of DoD systems; (3) physical and virtual range capabilities required to support testing of hypersonic weapons, integrated fires, and force-on-force operational performance; and (4) workforce skills

and capacity to accelerate the use of credible digital tools, automation, space-based OT&E and LFT&E and increase the availability of cyber and software scientists and engineers to include National Security Agency-certified Red Teams.

» PROVIDED INDEPENDENT EVALUATION OF OPERATIONAL PERFORMANCE

In FY23, DOT&E published 34 reports summarizing the adequacy of OT&E and LFT&E and a preliminary evaluation or final evaluation of the operational performance of the system. Of the 34 system reports, all included an assessment of test adequacy, 17 provided a final evaluation of operational effectiveness, 16 provided a final evaluation of operational suitability, and 12 provided a final evaluation of survivability. The remaining reports included a preliminary evaluation of operational performance not included in the operational performance trends discussed below and in Figure 2.

Operational Effectiveness Trends

In FY23, DOT&E reported 65 percent (11 of 17) of the evaluated programs to be operationally effective. By comparison, over the last 7 years (FY16-22), DOT&E reported 52 percent (71 of 137) to be operationally effective. DOT&E assessed two FY23 programs as not operationally effective and four programs as being partially effective because the system could either not complete one or more of its primary missions or had poor operational effectiveness in some operationally relevant conditions. For example, one system was able to complete missions in a permissive environment but could not complete missions in a contested environment because of poor survivability.

Operational Suitability Trends

In FY23, DOT&E reported 56 percent (9 of 16) of the evaluated programs to be operationally suitable. By comparison, over the last 7 years, DOT&E reported 52 percent (61 of 131) to be operationally suitable. DOT&E assessed six programs as not operationally suitable and one program as being partially operationally suitable. These seven programs, without exception, experienced shortfalls in reliability, availability, and/or maintainability. Other common suitability limitations included human systems integration challenges related to workload, usability, or training; transportability challenges; and immaturity of the logistical supply system.

Survivability Trends

In FY23, DOT&E reported 25 percent (3 of 12) of the programs to be survivable and 33 percent (4 of 12) to be partially survivable. By comparison, over the last 7 years, DOT&E assessed 31 percent (34 of 109) as survivable and 27 percent (29 of 109) as partially survivable, primarily due to a significant number of mission critical vulnerabilities in contested

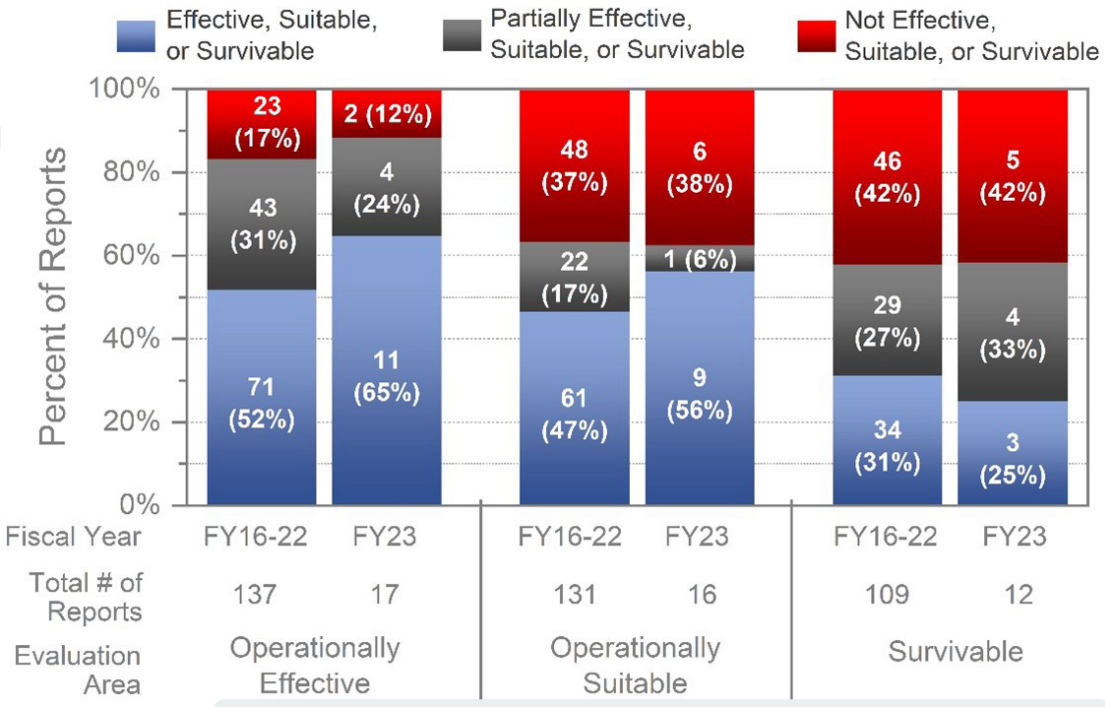

Figure 2. Operational Performance Trends in FY23 and Prior Years

cyberspace. Some systems also demonstrated unexpected vulnerabilities to kinetic threats and while operating in a contested electromagnetic spectrum environment.

Recommendation Trends

DOT&E reports include practical recommendations to fix the identified deficiencies, improve the operational performance of the DoD systems in expected operational scenarios, identify conditions to minimize risk to warfighters, and maximize probability of mission success. Examples of common recommendations are related to immature software, poor reliability, not survivable against cyberattacks, poor system performance, deficient human systems integration, and insufficient training and technical manuals.

RESPONDED TO WARFIGHTER REQUIREMENTS AND ADVANCED OT&E AND LFT&E PRACTICES

In FY23, DOT&E managed the Cyber Assessment Program (CAP) and the following field activities: (1) Center for Countermeasures (CCM), (2) Joint Aircraft Survivability Program (JASP), (3) Joint Technical Coordinating Group for Munition Effectiveness (JTCG/ME) that includes the Joint Live Fire (JLF) program, (4) Joint Test and Evaluation (JT&E), and (5) T&E Threat Resources Activity (TETRA). DOT&E also initiated the execution of the DOT&E Strategy Implementation Plan. Collectively, these activities made progress in responding to urgent warfighting requirements and transforming the DoD T&E infrastructure, tools, processes, and workforce in response to emerging changes in acquisition, technology, and warfighting. Details can be found in the DOT&E-Managed Activities and the Strategy Implementation Plan Update sections of this Annual Report. In summary, these activities:

- Improved the threats and operational realism in test. As an example, TETRA continued the development, validation, and delivery of 10 radio frequency and 10 infrared high-priority threat models, as well as over 25 high-fidelity, closed-loop, electronic-warfare-capable, emulative threat models needed for OT&E and LFT&E. JASP supported the delivery of new electronic attack techniques against advanced radar threats and improved hardware-in-the-loop capabilities for man-portable air-defense systems. CAP worked with the combatant commands to improve the operational realism of cyber assessments and emulate advanced threats.
- Advanced the use of credible digital tools in OT&E and LFT&E. For example, JTCG/ME continued to develop new digital tools (e.g., the Next Generation Enterprise Maritime Lethality Tool) and reduce the uncertainty in existing tools (e.g., Submarine Vulnerable Effects Model, Navy Enhanced Sierra Mechanics, and Dynamic System Mechanics Advanced Simulation) required to support the delivery and fielding of weaponeering tools against maritime targets while also supporting the survivability and lethality evaluations of U.S. Navy ships and submarines in contested environments. JTCG/ME also continued the critical VV&A and uncertainty quantification advancement efforts in coordination with the U.S. Army, U.S. Air Force, U.S. Navy, and Lawrence Livermore National Laboratory representatives.
- Supported the advancement of efficiency and agility of OT&E and LFT&E. For example, DOT&E supported the development of a prototype software application to optimize test sizing in a dynamic way using modern statistical inference methods to enable adaptive, integrated testing. DOT&E also supported the development of a prototype of a smart word processing and content management application intended to expedite the development and review of acquisition program (model-based) documents.
- Responded to urgent warfighter requirements. For example, JT&E supported new, data-based concepts of employment for long-range hypersonic weapons, nuclear command and control operations, and improved cyber survivability. JTCG/ME generated 13 reach-back packages for weaponeering, collateral damage estimates in support of current operations.

- Helped address the workforce challenges. For example, DOT&E supported an internship program – in partnership with the Army's Program Executive Office for Simulation, Training, and Instrumentation – resulting in 20 cyber experts poised to earn 20 Security+ certifications and 9 Certified Ethical Hacker certifications and join the T&E workforce.

CONTINUED TO SUPPORT GLOBAL T&E PARTNERSHIPS

In FY23, DOT&E continued to maintain multiple bilateral and multilateral agreements with international partners through the International T&E Program (ITEP) expediting the development and fielding of advanced warfighting technologies and supporting T&E infrastructure. Through ITEP, DOT&E finalized 12 new project agreements and is monitoring 24 ongoing projects. These projects are intended to improve capabilities and instrumentation among U.S. allies in areas including electronic warfare, autonomy, and survivability.

This page intentionally left blank.