

Shift Left!

Steven J. Hutchison, Ph.D.

Office of the Secretary of Defense/Acquisition Technology and Logistics, The Pentagon, Washington, D.C.

Last year, ITEA partnered with the University of Memphis Systems Testing Excellence Program to conduct the annual Technology Review in conjunction with the University's sixth International Research Workshop on Advances and Innovations in Software Testing. During the event, Mr. Dave Miller, vice president for software quality assurance at FedEx, gave a presentation entitled "Partnering—the Shift Left." Dave and his team are innovators in software testing, and I've given considerable thought to his message and its applicability to the broader landscape of Department of Defense (DoD) testing. Thus, I want to thank Dave for planting the seed of what follows.

Much of what we do in the DoD test and evaluation community to support decision makers is, frankly, *late to need*. We have to change the paradigm that has pushed critical test activities to the right in our acquisition process; it is time to *Shift Left!* Why is this important? Simply stated, testing late means finding problems late, and the later this occurs in the life cycle, the more costly it is to fix. Late discovery then leads to either delayed deployment or to accepting the shortcoming and fielding the system as is to our Warfighters. Neither outcome is acceptable to me, professionally or personally, but especially when we turn a development problem into a Warfighter problem.

Developmental testing in the DoD acquisition process

In today's acquisition process, critical test activities to determine if the new equipment satisfies the needs of the Warfighter occur after the decision to begin producing the new equipment. Our programs begin production prior to Initial Operational Test and Evaluation (IOT&E), Joint interoperability testing, and, increasingly more important today, cybersecurity testing. Because these events will occur later anyway, Program Managers (PMs) frequently trade off developmental testing ("we'll do that in OT") for near-term buying power. That is a strategy that almost never pays off.¹ The bottom line—if we want to get improved capability into the hands of the Warfighter more rapidly and at less cost, we have to get the development right

and verify it through rigorous Developmental Test and Evaluation (DT&E) before we commit to production.

The evolution of our acquisition process has right-shifted critical test and certification activities. Instead of using T&E to support the decision to *begin* production, we begin production to support testing! This is codified in statute and reflected in policy: 10USC §2399 establishes considerations for OT&E, including determination of the quantity of articles required for operational testing. The statute then goes on to establish conditions for proceeding *beyond* Low Rate Initial Production (LRIP). What 10USC does not do, however, is codify similar considerations for DT&E or establish conditions for proceeding *into* LRIP. This idea is not new; the U.S. Government Accountability Office (GAO) reported this finding almost 20 years ago:

*"Current legislation and DOD's acquisition policies permit LRIP to start before any OT&E is conducted. The consequences have included procurement of substantial inventories of unsatisfactory weapons requiring costly modifications to achieve satisfactory performance and, in some cases, deployment of substandard systems to combat forces.... In GAO's view, the key decision as to whether to proceed with production should be made at the start of LRIP because, in many cases, it is also the de-facto full-rate production decision."*²

The same GAO report went on to suggest:

"Congress may wish to require that all defense acquisition programs (major and nonmajor) conduct enough realistic testing on the entire system or key subsystems to ensure that key performance parameters are met before LRIP is permitted to start. The objective of GAO's recommendations is to avoid the premature commitment to production and thereby avoid fielding systems that do not meet requirements and need costly and time-consuming retrofits."

In other words, the GAO was recommending a shift left. And it hasn't just been the GAO; there have been countless blue ribbon commissions, defense science board panels, National Research Council studies, Inspector General reports, industry reports, and more,

that have all reached the same conclusion—when it comes to testing, earlier is better.

Take a look at the timing of test events in our acquisition process. *Figure 1* highlights test and evaluation as depicted in the well-known acquisition “wall chart.” The wall chart is a detailed systems engineering-based depiction of activities and critical decisions described in the DoD 5000 series directive and instruction. This template for the acquisition life cycle is reflected in the acquisition strategy, the test and evaluation master plan (TEMP), and other key planning documents. It is a script that our program managers—and testers—follow, typically without deviation.³ This image illustrates what initially appears to be a good DT&E strategy as the program moves up the right-hand side of the systems engineering “V” in preparation for Milestone C. Yet, on closer inspection, it is incomplete; note the timing of Joint interoperability testing and the issuance of the interoperability certification relative to Milestone C. Note also that cybersecurity testing (information assurance) and the issuance of an authority to operate as prescribed under the DoD Information Assurance Certification and Accreditation Process (DIACAP)⁴ are not specifically included in the details of this image. Both interoperability and cybersecurity certification involve *critical test activities* essential to helping programs set the conditions for successful production and deployment. But they are late to need.

I am describing the situation based on a picture of course; so the question is, how does this play out in the real world and what outcomes are we achieving? Where interoperability and cybersecurity are concerned, we now have considerable data showing that unresolved issues continue to be discovered in operations. These data come from 10 years of executing the congressionally directed program of interoperability and information assurance assessments during combatant command and Service exercises led by the Director, Operational Test and Evaluation (DOT&E).⁵ After almost a decade of assessments, the results are remarkably consistent: fielded systems exhibit interoperability shortcomings and information assurance vulnerabilities. The FY2012 DOT&E Annual Report, for example, stated: “Overall, the DOT&E [Information Assurance and Interoperability] IA/IOP program observed cyber effects caused by unresolved interoperability deficiencies, coupled with low-to-moderate level threats that were sufficient to adversely affect the quality and security of mission critical information in a way that could (and where permitted did) degrade mission accomplishment significantly.”⁶

Since this program assesses the interoperability and information assurance posture of *operational* systems, it provides value in the form of acquisition hindsight; in other words, it lets us see what got through the acquisition process into the field. For testers, this hindsight can be confirmation of what we discovered in testing and the subsequent decision to accept the risk for fielding; or (this is not as palatable), it can illuminate deficiencies we missed in testing. Regardless, we can be certain that the current process is permitting numerous defects to get to the field, and it is time to break that paradigm.

Shift Left!

It is time to Shift Left! The basic premise of *Shift Left* is to “do better DT&E” and fix the problems before entering production. There are three key elements to Shift Left: earlier mission context with user input, earlier interoperability testing, and earlier cybersecurity testing. Part of doing better DT&E requires us to get past the old-school notion that DT&E is only about validating technical compliance.⁷ If we limit DT&E to compliance with specifications and other technical matters, we will miss the sense of whether the capability satisfies the Warfighter need. If, however, we test in a mission context, not only will we obtain that critical operator feedback early in the life cycle, but we will also be able to answer the technical questions—that’s a two-for-one deal no PM should pass up! Robust DT&E should also include all of the elements of interoperability and cybersecurity testing and bring the right resources to bear to provide confidence in the decision to enter production. End-to-end testing of a Joint mission thread will provide this confidence. This does not mean PMs have to conduct DT involving large-scale deployment of troops to the field; but getting the new capability out of the lab to see how it will actually be used should be considered an important part of DT. Ideally, this is what the PM, chief developmental tester, and lead DT&E organization intend to accomplish when they assemble the T&E Working Integrated Process Team (WIPT) and write the TEMP.

So what are the offices of the Deputy Assistant Secretary of Defense (DASD) for DT&E and Director, Test Resource Management Center (TRMC) planning to do to facilitate this Shift Left? First and foremost, in our engagement with programs, we are going to assist PMs, chief developmental testers, and the lead DT&E organizations in developing and executing a comprehensive DT&E strategy that reflects the mission context and includes early testing in support of interoperability and cybersecurity certification. We will help programs craft the wording

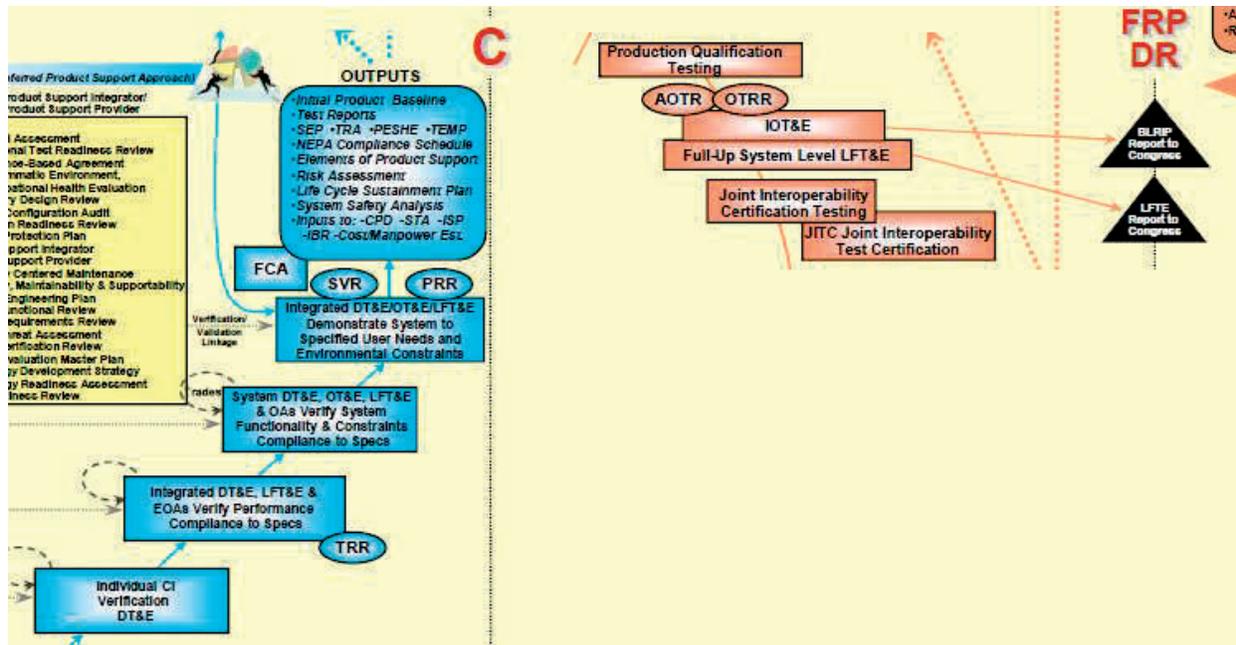


Figure 1. Test and evaluation in the defense acquisition system.

in TEMPs and other documents to reflect a sound DT&E strategy that will set the conditions for entry into production and successful OT&E. We will assist programs in provisioning the necessary infrastructure resources, such as the Joint Mission Environment Test Capability (JMETC) and cyber range, to execute the tests in the most efficient manner. Finally, during decision-making forums, such as Defense Acquisition Board meetings, we will raise the bar for DT&E. To do this, we will execute an internal Shift-Left process change and shift from the Assessment of Operational Test Readiness (AOTR), which is too late to really be meaningful, to a “DT&E Assessment” to accurately characterize the performance, reliability, interoperability, and cybersecurity status of the system to better support the decision to begin production at Milestone C (MS C).

Interoperability

The latest version of the Joint interoperability certification process reflects a new role for DASD(DT&E) and the Joint Interoperability Test Command (JITC). The Chairman of the Joint Chiefs of Staff Instruction (CJCSI) 6212.01F⁸ states:

“(A.2.c) DoD Components will ensure the Component Developmental Test and Evaluation (DT&E), Operational Test and Evaluation (OT&E) processes include mission-oriented [Net-Ready Key Performance Parameter] NR KPP assessments...

“(A.7.b) [Defense Information Systems Agency] DISA will ensure JITC leverages previous, planned

and executed DT&E and OT&E tests and results to support joint interoperability test certification and eliminate test duplication. DASD(DT&E) shall approve Developmental Test and Evaluation plans in support of Joint Interoperability Test Certification as documented in the TEMP. JITC shall advise DASD(DT&E) regarding the adequacy of test planning in support of Joint Interoperability Test Certification.”

When this new language was released, we applauded the first paragraph to make efficient use of existing test activities for interoperability test purposes but scratched our heads wondering how we were going to implement the relationship directed in the second paragraph. In meeting with JITC, we determined that the best path forward was not to introduce a burdensome new test plan approval process; rather, we decided to work with program offices to add, where appropriate, relevant interoperability test measures and data collection activities during DT&E, and reflect the interoperability test objectives in the DT&E event descriptions and required resources in the TEMP.

Cybersecurity

We have a similar effort underway in the revision to the DoD 8500 series instructions for information assurance (IA). The 8500 series is being revised based on new guidance from the National Institute of Standards and Technology (NIST) and the Committee on National Security Systems Instruction (CNSSI) and may include changes to current processes such as

adopting the term “cybersecurity,” implementing the Risk Management Framework (RMF) (as opposed to the current Mission Assurance Category [MAC] and Confidentiality Level [CL]), and updating terminology such as the following:

1. Certification and Accreditation (C&A) becomes Assessment and Authorization (A&A).
2. Designated Approving Authority (DAA) becomes Authorizing Official (AO).
3. Certifying Authority (CA) becomes Security Control Assessor (SCA).

More important though, as was highlighted in the discussion on *Figure 1*, our current process fails to adequately highlight cybersecurity testing as a critical test activity during DT&E. Security test and evaluation⁹ has been all but lost under DIACAP scorecards and Program of Actions And Milestones (POAMs) and is overseen by a DAA who probably does not understand test. We are working with all stakeholders to recognize that in order to certify, you have to test, and that these security test activities are (a) essential elements of a program’s overall T&E strategy; (b) that we need to include the cybersecurity test organization in the T&E WIPT; and (c) as with interoperability testing described above, we need to plan for and execute cybersecurity testing in conjunction with other test activities to reduce duplication.

In the areas of interoperability and cybersecurity, DASD(DT&E) is working with all stakeholders to insert needed testing early and to define the right way to oversee these processes. Words are important, and it is important that we be clear in our intent: our objective is to establish a value-added process to oversee and improve the developmental test activities that support certification, not to set requirements or oversee certification, as those functions are the responsibility of the respective offices in the Joint Staff and DoD Chief Information Officer.

A cybersecurity DT&E methodology

To better understand what cybersecurity DT&E entails, and thereby help programs craft improved cybersecurity testing as part of their overall T&E strategy, DASD(DT&E) and TRMC staff held a series of strategic planning sessions to begin development of a comprehensive cybersecurity DT&E roadmap, encompassing policy, methodology, workforce skills, and infrastructure. Since cybersecurity testing potentially exposes systems and networks to representative capabilities of the advanced persistent threat, we recognize the need to provide a secure, isolated representation of the operational environment to avoid collateral damage to live networks, systems, and data sources.

There are several variants of “cyber ranges” today; however, none were explicitly designed to support cybersecurity T&E for programs of record or have the capacity to do so. Recently, the National Cyber Range was transferred to the TRMC and will become an essential part of a distributed cyber test and training infrastructure. Much work needs to be done to provide this capability for the enterprise, and DASD(DT&E) and TRMC are working aggressively to define the requirements, determine investments needed, engage stakeholders, and develop the infrastructure.

To support robust cybersecurity DT&E, DASD(DT&E) developed a four-step cybersecurity DT&E methodology:

- Step 1. Understand cybersecurity requirements.
- Step 2. Characterize the cyber attack surface.
- Step 3. Understand the cyber kill chain.
- Step 4. Conduct cybersecurity DT&E.

The four-step process will not require additional documentation; it will use existing documents such as the Program Protection Plan, to assist in identifying critical areas to focus cybersecurity DT&E. The process is intended to be iterative throughout the life cycle; updating information as requirements and operational concepts evolve. The first step in developing a cybersecurity DT&E strategy is to understand the requirements and concept of operations for cybersecurity; thus we begin very early in the acquisition life cycle, at milestone A or B, to identify critical DT&E activities and cybersecurity metrics. Step 2 factors in the environment in which the system interoperates. In this way, we can characterize the *attack surface*; that is, understand the avenues by which a potential adversary may gain access to the system. The third step, once we understand the attack surface, is to understand the cyber *kill chain*. The kill chain is a construct that describes potential adversary actions, such as monitoring data exchanges, escalation of privileges, or embedding malicious software. When we understand attack surface and kill chain, developers and network defenders are better able implement measures to improve resilience. Finally, as step 4 indicates, we believe programs and decision makers will derive tremendous benefit from rigorous DT&E with a capable cyber threat representation, in a range intended for that purpose. By understanding the requirements, attack surface, and kill chain, developmental testers can identify the right set of metrics and craft a robust cybersecurity DT&E strategy that will provide decision makers essential information and reduce the potential for discovery when it is too late to fix and a development problem becomes a Warfighter problem.

Summary

Developmental test and evaluation is a tool used throughout the program life cycle to provide early and continuous feedback; it is the key to an informed decision to begin production and helps set the conditions for acquisition success. DT&E provides the knowledge to measure progress on performance, assess safety, as well as characterize capabilities and limitations. As DoD acquisition programs become increasingly complex, DT&E must leverage all resources and venues as potential data sources, to include testing in system integration labs and ranges, use of modeling and simulation, and where practical, leverage training exercises, experimentation, and operations. DT&E should exploit the power of the network—resources such as the JMETC—as a means to bring test resources together to reduce cost, gain efficiency, and improve realism. We need to bring the mission context to DT&E, and we must bring essential elements of interoperability and cybersecurity into DT&E to get that critical, early assessment to better inform the decision-making process. This is a priority initiative for the DASD(DT&E) and TRMC. The way I see it, if we want to gain some acquisition agility and get capabilities into the Warfighter’s hands more rapidly within the framework of the existing DoD 5000 process, we have to Shift Left! □

DR. STEVEN J. HUTCHISON is the Acting Deputy Assistant Secretary of Defense for Developmental Test and Evaluation (DASD[DT&E]) and Acting Director, Test Resource Management Center (TRMC). Dr Hutchison has been a member of the T&E community for the past 15 years, holding positions including T&E Executive at Defense Information Systems Agency (DISA), net-centric warfare systems analyst in DOT&E, and both assistant technical director and evaluator at U.S. Army

Test and Evaluation Command (ATEC). Dr. Hutchison recently completed serving as a member of the ITEA Board of Directors. E-mail: steven.hutchison@osd.mil

Endnotes

¹In July 2000, the GAO wrote: “Despite good intentions and some progress by the Department of Defense (DOD), weapon system programs still suffer from persistent problems associated with late or incomplete testing.” GAO, Best Practices: A More Constructive Test Approach Is Key to Better Weapon System Outcomes, July 2000 (<http://www.gao.gov/assets/160/156809.pdf>).

²GAO, Weapons Acquisition: Low-Rate Initial Production Used to Buy Weapon Systems Prematurely, November 1994 (<http://www.gao.gov/assets/160/154796.pdf>).

³The National Research Council recently wrote about the ability to tailor the DoD 5000 process: “Although the DOD’s current governance and oversight structure permits tailoring and provides the flexibility needed for a milestone decision authority and program manager to adjust how the process is applied to specific programs, there is no established best practice or accepted template for tailoring.” NRC, Achieving Effective Acquisition of Information Technology in the Department of Defense. National Academies Press, 2010 (http://www.nap.edu/catalog.php?record_id=12823).

⁴DoDI 8510.01, DoD Information Assurance Certification and Accreditation Process (DIACAP), November 28, 2007.

⁵In the FY03 Defense Appropriations Bill, Congress directed DOT&E to “establish a process using OT&E of the systems on his oversight list and exercises conducted by Combat Commands and the Services to monitor the Department’s on-going efforts to improve interoperability and information assurance.”

⁶FY2012 DOT&E Annual Report (<http://www.dote.osd.mil/pub/reports/FY2012/>).

⁷On the back of the acquisition wall chart is a definition of Developmental Test and Evaluation: “A technical test conducted to provide data on the achievability of critical system performance parameters.”

⁸CJCSI 6212.01F, Net Ready Key Performance Parameter (NR KPP), 21 March 2012.

⁹The DITSCAP process (formerly DoDI 5200.40), which preceded DIACAP, featured robust Security Test and Evaluation as part of security certification and accreditation and included the following definition: Security Test and Evaluation (ST&E). Examination and analysis of the safeguards required to protect an IT system, as they have been applied in an operational environment, to determine the security posture of that system.