OFFICE OF THE SECRETARY OF DEFENSE
1700 DEFENSE PENTAGON
WASHINGTON, DC 20301-1700

OPERATIONAL TEST
AND EVALUATION

JAN 10 2017

MEMORANDUM FOR CHAIRMAN, DEFENSE BUSINESS BOARD

SUBJECT: Briefing on "Best Practices for the Business of Test and Evaluation" is Fatally Flawed and Unsuitable for Decision-Making

The purpose of this memorandum is to provide my feedback on the Defense Business Board (DBB) briefing slides dated October 20, 2016 and titled, "Best Practices for the Business of Test and Evaluation." While Operational Test and Evaluation (OT&E) and Live Fire Test and Evaluation (LFT&E) are only parts of the Department's Test and Evaluation activities, I felt it necessary to provide my feedback as Director, Operational Test and Evaluation to the report's content as it pertains to activities under my purview. The briefing slides make numerous claims of fact absent any supporting data. Perhaps this is because these "facts" are, in fact, false. For example, the first finding on slide 3 states "true costs are unknown." This is not true. As we have documented, we do know the cost of operational testing, and it is a relatively small percentage (typically less than 1 percent) of a program's acquisition costs. A more important question, left unasked by the study, is "what is the cost of losing in combat when our weapons don't work as required or when required because testing was inadequate?" As discussed in detail in the remainder of this memorandum, there are so many instances of false findings that the briefing's recommendations, such as we can understand them, are not useful at all for decision-making or for action by either the Department's leadership or the Congress. This unfortunate situation could have been avoided if the study group had accepted our invitation to engage in more than the single interaction with DOT&E that occurred.

Background

In 1983, Congress created the position of Director, Operational Test and Evaluation (D,OT&E) within the Office of the Secretary of Defense, and the Director was given specific authorities in Title 10 U.S. Code. The Congressional concerns that led to the establishment of this office were many, but included: poor performance of weapon systems, inaccurate reports from the Services, shortcuts in testing because of budget pressure, and a lack of realistic combat conditions and threats in testing. The unique independence of this office, free from conflicts of interest or pressure from Service senior leadership allows us to:

- Highlight problems to DOD and Congressional Leadership to inform their decisions before production or deployment
- Tell the unvarnished truth
- Ensure operational tests are adequately designed and executed

As Director, OT&E, I do not make acquisition decisions but inform those who make them about weapon system performance under combat conditions. My staff of 60 action officers includes 17 active duty military officers from all Services in addition to civilians with advanced

engineering and science degrees. Our mission is to inform acquisition officials before the systems are purchased en masse, and combatant leadership before the systems are deployed operationally, about how weapons will work in combat, including live fire survivability and lethality, before the systems are deployed.

Furthermore, the workforce of the Services' operational test communities is relatively small compared to the approximately 26,000 personnel mentioned on slide 5. In fact the Army, Air Force, Marine Corps, and Navy operational test agencies in total employ less than 2,000 people, and, as shown in Figure 1, that total has declined more than 11 percent in the past 10 years, primarily in the number of military personnel. This is in contrast with trends reported for the broader acquisition workforce, which has grown by 24 percent since FY08.[1]

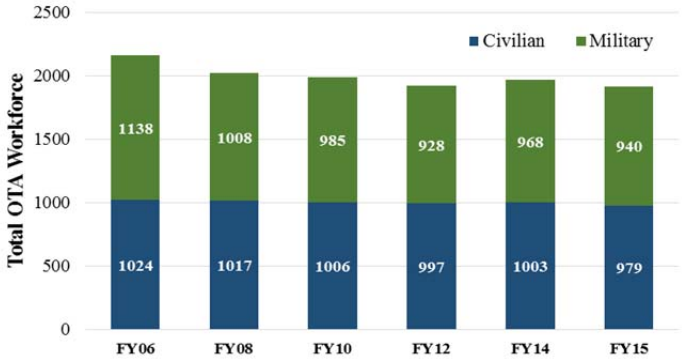

Figure 1. Proportion of Civilian vs. Military Personnel in the OTA Workforce Growth FY06-FY15

Below, I discuss the three principal areas covered by the report: the process of defense test and evaluation, its supporting infrastructure, and possible lessons to adapt from industry.

[1] Schwartz, Francis, and O'Connor (2016). The Department of Defense Acquisition Workforce: Background, Analysis, and Questions for Congress, https://www.fas.org/sgp/crs/natsec/R44578.pdf

Process of Defense Test and Evaluation

Process Findings (slide 10-11)

"**Overall T&E costs are unknown and not viewed as a priority: Lack of incentive to reduce cost because cannot track them directly**"

This is not true; we do know the costs of operational and live fire test and evaluation. The Test and Evaluation Master Plan (TEMP), required for every acquisition program under OT&E or LFT&E oversight, has a section devoted entirely to resource allocation for both developmental and operational testing, and I have provided clear guidance and examples for this section of the TEMP in my TEMP guide, which has been published online since 2011.[2] The program manager, in coordination with all the T&E stakeholders, must identify and plan for all T&E resources needed to adequately support Developmental Test and Evaluation (DT&E), OT&E, and LFT&E. The first step is to develop data requirements in the evaluation framework and from those data requirements determine the resources needed at each stage of the program. The Resource Summary section of the TEMP should flow directly from these analyses and identify the test resources to conduct the tests described earlier in the TEMP.

In my Fiscal Year 2011 Annual Report, I examined cost data from recent acquisition programs. Existing data show that T&E cost is only a fraction of the acquisition costs of a program (typically less than 1 percent). This small relative cost stands in stark contrast with the potential savings from problems identified that can be corrected before full-rate production and the likely result that the system will work when called upon in combat. We evaluated marginal cost of operational test and evaluation to programs as a percentage of total acquisition cost. A review of 78 programs in the Army, Air Force, and Navy showed that the average marginal cost of OT&E is approximately 0.65 percent of the total acquisition cost. Few programs that we reviewed (7 out of 78) required more than 1.5 percent of program acquisition costs for OT&E. For those programs with above average OT&E costs, a relatively low program acquisition cost was the dominant cause of larger proportional OT&E cost (e.g., AIM-120C Electronic Protection Improvement Program with $87M acquisition costs). Expense of test articles and their expendability was another major driver. Figure 2 shows the distribution of the marginal cost of OT for the 78 programs we examined.

T&E costs can be inflated by adding in development costs for programs which were eventually abandoned, though investments in T&E early in system development can create savings there by identifying problems with systems before good money is sent after bad. The Decker-Wagner Report, commissioned in 2010 by the Secretary of the Army, gives concrete evidence for the need for early operational testing to illuminate problems by examining the Army's failure rate of initiating and then canceling new development programs.[3] The study found that between 1990 and 2010, the Army terminated 22 Major Defense Acquisition Programs (MDAPs), and that 15 of those terminations occurred since 2001. Further, excluding the Future Combat System (FCS), the Army spent more than one billion dollars *per year* since 1996 on programs that were eventually canceled before completion. The study cited many

[2] http://www.dote.osd.mil/tempguide/

[3] Army Strong: Equipped, Trained and Ready, Final Report of the 2010 Army Acquisition Review, January 2011.

reasons for the failed programs including unconstrained requirements, weak trade studies, and erosion of the requirements and acquisition workforce. However, none of the reasons cited included T&E. Earlier and more robust T&E would have revealed problems and solutions sooner, when they would have been less costly to fix or allowed decision makers to cancel or restructure programs and avoid wasting billions of dollars.

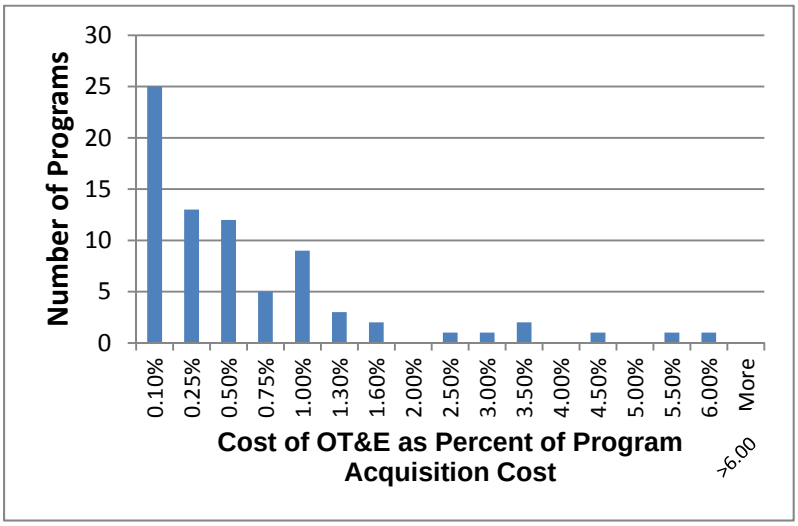

Figure 2. Marginal Cost of OT&E relative to program acquisition cost

"Philosophical difference between testing to original requirements vs. new adversary capabilities"

Programs often complain that DOT&E requires testing beyond threshold requirements, or even threshold key performance parameters (KPPs). However, our observation is that KPPs continue to provide an evaluation metric describing a level of performance that, while necessary, fails to sufficiently meet anticipated operational needs. KPPs often lack the context of the complex operational environment, including current threats. For example, the P-8A Poseidon is a maritime patrol aircraft that will replace the P-3C Orion and conduct Anti-Submarine Warfare (ASW) and other missions. However, the KPPs required only that the P-8A be reliable, be equipped with self-protection features and radios, and carry a requisite number of sonobuoys and torpedoes, but not actually demonstrate an ability to find and prosecute submarines. DOT&E, working with the Navy's Operational Test Agency, focused the testing on examining quantitative mission-oriented measures, beyond the limited KPPs, in order to characterize the aircraft's ASW capabilities. The Department of Defense Inspector General specifically criticized the Navy and Joint Staff for not promulgating requirements for the P-8 capturing the key aspects of the combat missions it is meant to conduct.[4]

Similarly, the F-35 Joint Strike Fighter, the department's largest ever acquisition program, is a fifth-generation fighter expected to operate in highly contested peer or near-peer threat environments. Between the F-35 program origins, which began from the Joint Advanced

[4] Naval Air Systems Command Needs to Improve Management of Waiver Requests, DODIG-2015-122, May 2015.

Strike Technology (JAST) program in 1993, and today, our adversaries have developed, fielded, and exported advanced integrated air defense systems – systems the F-35 was not originally designed to defeat. While the program office has a "philosophical" difference with DOT&E about testing against these threats that have in fact been fielded for a decade, it is DOT&E's statutory responsibility, and is simply common sense, to test systems in realistic operational conditions expected in actual combat. F-35 capability is critical to the U.S. prevailing in combat against existing and emerging threats worldwide. Performance of the F-35, and all combat systems, must be characterized across its operational environment so that military leaders can make informed decisions regarding their employment.

If programs were tested solely to their KPPs and ambiguous requirements documents, we often would not be able to evaluate whether systems can accomplish their primary missions. While we must always pay attention to requirements documents, we also have to interact with the operators and understand the concepts of operation and the war plans. Additionally, the Services update a program's threats on a regular basis from the intelligence information on the latest threats. All of these things will tell us how to do an operational test under the circumstances the system will actually be used in combat and enable to characterize the performance of systems across their operational envelope – not just at one key parameter. Operational testing is not a bureaucratic game to be played to obtain a "good grade." Rather it is a deadly serious activity whose purpose is to discover whether systems will perform as needed, consistent with our war plans, before they are used in combat when failure cannot be tolerated.

"Accountability for risk trade-off decisions is not clear or consistent"

A large portion of this briefing is focused on the risk calculus, which should be considered in T&E to adequately scope the T&E program. However, this study is focused on the wrong aspect of it --- it mainly targets the Services' perceived diminished role in the process of designing adequate combat testing – causing pressure on cost and schedule. However, the risk of weapons not working as required or when required in combat is not addressed. If future studies wish to truly investigate how to adequately balance resources for T&E activities, they must look at both sides of the scale. Moreover, DOT&E's reports, including its most recent Annual Report, document many significant and substantial shortfalls in operational effectiveness that never would have been known if Service proposals to "balance risk," which almost always amount to false reasons for minimizing testing, had been accepted.[5]

"Perception that drive for zero risk is too strong and uncompromising"

Testing never strives for nor achieves zero risk. DOT&E reports annually on problem discovery during operational testing – the tests conducted at the end of development before fielding. This testing reveals that nearly two-thirds of programs find significant problems during operational testing (significant enough to impact operational effectiveness, suitability, or survivability). Finding and addressing these problems before production and deployment is critical; otherwise they would be discovered in combat.

DOT&E has advocated for rigorous and scientific approach to test planning and resourcing; this scientific approach uses design of experiments (DOE) to cover the operational

[5] http://www.dote.osd.mil/annual-report/index.html

environment as efficiently as possible, allowing for us to balance risk and information. In fact, many of the tests conducted allow a high level of risk of incorrectly concluding that a system is performing adequately---which we use rigorous analytical techniques to quantify--- due to limited data. Often resources are limited up-front (when testers are not included in early resource decisions) and testers use DOE to get the most out of what they are given. DOT&E recently has been trying to define test scope/resources well-ahead of the MS B request for proposals from industry.

Furthermore, the claim that testers strive for zero risk before fielding new systems is false because DOT&E is not responsible for fielding decisions. In fact, fielding decisions are authorized by the Services, not DOT&E. It is common practice for Services to field systems with major deficiencies, transferring risk from the acquisition community to the warfighter and the nation.

"Services can feel left out of final decisions and lack authorities to make final "calls" on suitability and effectiveness"

Title 10 U.S. Code defines the responsibilities of the Director, OT&E. It is the Director's independent authority to report the operational effectiveness, operational suitability, survivability and lethality of weapons systems. The Services each have their own independent operational test agency that uses the same data but may report different overall conclusions. It is up to the Services to decide whether to do anything in response to DOT&E's findings, and as previously stated they often field the program despite the significant problems that DOT&E reports. Moreover, DOT&E works closely with the Services' operational test staffs and leadership to design realistic test. In the vast majority of cases, these interactions result in agreement between DOT&E and the Service's on the operational testing that needs to be done.

"Lack of resources when additional testing is required creates pressure on budget and schedule"

A primary purpose of operational testing is to identify critical problems that can be seen only when systems are examined under the stresses of realistic combat conditions, prior to the full-rate production decision and fielding to combat units. This identification permits corrective action to be taken before large quantities of a system are procured and avoids expensive retrofit of system modifications. The assertion that testing causes delays misses the essential point: fixing the deficiencies causes delays, not identifying them. DOT&E conducted a study of over 100 programs on DOT&E oversight that experienced delays and found that less than 25 percent of cases did a problem with test conduct delay the program; the most common reason contributing to delays was performance problems discovered during testing that had to be addressed before continuing development and fielding.[6] Furthermore, taking the time to correct serious performance problems is exactly what we desire in a properly functioning acquisition system.

"DoD and industry are engaged in addressing the emerging challenges and generally work well together; but many view DoD process as bureaucratic and overly complex"

[6] http://www.dote.osd.mil/pub/presentations.html, "Causes of program delays"

Some may view DOD process as bureaucratic and overly complex compared to industry; however, the systems DOD is developing, procuring, and fielding are the most complex and expensive in the world. Acquiring capabilities critical to the national interest – and doing so with limited public funds – necessitates a level of careful oversight above and beyond what might be required in industry. If future reports focus on identifying inefficiencies or bureaucratic processes which can be eliminated without diminishing the effectiveness of the T&E process to find and fix problems, investigating the problem instead of merely complaining about its scale, we would welcome such a study.

"Strategic Capabilities Office (SCO) offers alternatives to the current test approach: New take on the "development/test/deployment" method"

The strategic capabilities office, by definition, does not produce integrated weapon systems to be operated and sustained by a warfighter. Our experience with programs developed by the SCO is that they undergo rigorous test and evaluation just as other major combat systems do, but their deployment or early fielding is stymied by the lack of trained operators, maintainers, spare parts, and concepts of operation. Moreover, we have engaged in numerous examples of rapid testing to support rapid fielding, with the Mine Resistant Ambush-Protected Vehicle program being a notable---but by no means the only---notable example.

"Misaligned incentives drive process participants to act in certain ways: Budget (PPBE), Requirements (JCIDS), Program Management, T&E"

Programs clearly have an incentive to denounce testing as unfair when it reveals performance problems. Cost and schedule overruns, especially those that are the direct result of poor program management, reflect poorly on program managers and program executive officers. There's a terrible fear that exists that a negative DOT&E report will kill a program; however, it is much more likely that performance problems reported by DOT&E lead to a greater allocation of resources and time to fix them. However, by engaging in bureaucratic games, rationalizing problems, and minimizing testing, acquisition officials do a great disservice for the people for whom we work – the men and women in combat whose lives depend on the systems we give them to achieve their mission and the nation whose security depends on that mission's success.

<u>Process Recommendations (slides 15-17)</u>

"Create appropriate and direct accountabilities for T&E community, aligned with Program Management"

Independence is the long pole holding up DOT&E's tent – every other virtue (i.e., Accuracy, Integrity, Objectivity) flows from it. None of these would be possible if DOT&E were required to subvert its expertise and scientific investigations to bow to budgets and schedule considerations. DOT&E already takes budgets into account by using Design of Experiments to most efficiently cover the operational envelope and by using modeling and simulation to complement physical tests when doing so is scientifically justified. As stated earlier, tests are routinely planned with risk levels higher than most industry standards with substantial probability of concluding that a system is performing adequately, when it is not. By decreasing the power and thereby increasing the risk of accepting poorly performing systems, eventually the test will be a coin-toss. Is a test with only 50 percent power worth conducting?

The ultimate outcome of making T&E accountable to program management is to turn T&E into a simple box-checking exercise. Consider the contractor-run KC-46A Aerial Refueling Tanker test program. This system was purchased under a firm fixed-price contract where the contractor, Boeing, ran the integrated test program. Boeing consistently produced unrealistic and overly-optimistic schedules as analyzed in DOT&E annual reports since 2011. Reality drove the schedules past planned calendar dates leading Boeing program management to cut testing. As noted in DOT&E's KC-46A OA-2 report, Boeing's original plan was to have 75 percent of testing completed by Milestone C, but in the end, a mere 25 percent of testing was completed even though the milestone was delayed by over 12 months. Boeing only accomplished the bare minimum of events needed to check off the items written in the Milestone B Acquisition Decision Memorandum. While demonstrating the last of these items, refueling using the boom with three different receivers, a serious problem emerged. Loads within the boom were approaching structural limits while pushing and pulling on the receiver. Boeing drove to continue the demonstrations and proceed to Milestone C and the Low-rate Initial Production (LRIP) portion of the contract since they did successfully offloaded fuel. After pressure from the 412th Test Wing, Air Force Operational Test and Evaluation Center (AFOTEC), the KC-46A Government Program Office, and the A-10 program office which withdrew their clearance to test letter, Boeing postponed testing and redesigned the boom. In the process, Boeing had a safety stand-down day to internally discuss how they missed early indications of the excessive loads in collected data while they sped ahead to complete testing.

"Definitions of "success" should be agreed between PM and T&E community from the beginning and with changes that occur over time"

During the development of requirements and the concept of operations, the PM and warfighter communities agree on the definition of success. T&E tests to this definition.

"Sharply limit the number and weight of stakeholders that have veto power over TEMP development, implementation, and modification"

This recommendation is not specific enough to be useful – but it should be noted that the Services generally have 15-20 signatures on a TEMP while OSD has only one or two. For example, the recent TEMP for the Global Broadcast Service (GBS) required 22 signatures from the Services, including PMs, Program Executive Officers, requirements authors, and Operational Test Agency (OTA) commanders, but only one signature from OSD – mine.[7]

"Authorities should come from the Service Secretary to the Service Chief and designated subordinates"

This is true today. Service Acquisition Executives work for the Service Secretaries. They make decisions to buy systems. Service Chiefs and subordinates make decisions to field systems – with or without serious deficiencies.

[7] Test and Evaluation Master Plan for Global Broadcast Service, 21 March 2016

"Today's drive toward "zero risk" is an impossible standard and should not be used by the T&E community"

This claim is silly and false. Zero-risk is not used by the T&E community and never has been. However, if the DBB briefing is confusing T&E with requirements, I completely agree that we should move towards stating and quantifying "tolerable risks" in the requirements documents, both in terms of allowable performance degradations and our ability to know them. For many programs, I have engaged early with requirements writers to determine whether 99.99 percent probability-based requirements are necessary, because both contracting for such systems and testing them would be extremely expensive. For example, 99.99percent reliability requirement is extremely expensive to test if we want to ensure we are not actually 99.98 percent. Whereas requiring 90 percent within +/- 5percent is achievable in a relatively short test.

"Adopt "principled compromise" between absolute zero-defect test doctrine and situational operating requirements"

The claim that there is a zero-defect test doctrine is demonstrably false. Operational testing characterizes system performance across its intended operating environment. "Zero-defect" is not a test planning goal, nor could it ever be possible to test to "zero-defects." The test doctrine does not determine the number of allowable defects. The requirements do. The test doctrine then determines how many tests we need to do to prove out that requirement with acceptable statistical risk. Operational testing uses sound test science principles to design efficient and adequate testing, meant to determine the actual performance of systems in as close to operational environments as possible. Implementation of rigorous test design and analysis provides defensible, factual information to scale test designs and to provide credible test results on system performance to the Congress and civilian and military leaders so that they could make informed decisions regarding acquisition and employment of those systems.

"Tolerable" risk should be stated and quantified"

It is. Design of Experiments (DOE) provides a defensible and efficient methodology for not only determining test adequacy but also ensuring that we obtain the maximum value from scarce test resources. DOE has been proven to elicit the maximum information possible from constrained resources, provide the ability to combine information across multiple independent test events, and produce defensible rationale for test adequacy and quantification of risk as a function of test size. One clear advantage of statistical approaches to evaluating test adequacy is that they provide a means to quantify how much information can be derived from each test point.

"Apply more consistent analysis of common data in the T&E process"

Our reports demonstrate clearly that we use to the fullest all the data that are available on operational performance under realistic conditions, to do otherwise would be stupid. The Services and DOT&E provide independent analyses and evaluation. My office is working towards developing analysis tools that can be used by all Services, but ultimately the evaluations and supporting analyses between oversight and the Services should remain independent.

"Services are best positioned to determine "good enough" when program length results in new threats that didn't exist when requirements set"

Actually, Service acquisition professionals are not best positioned to determine "good enough," especially considering that the current acquisition system incentivizes them to stay on cost and schedule regardless of the performance of their system. Before Congress created the office of DOT&E, the Services provided inaccurate reports of system performance, took shortcuts in testing because of budget pressures, and conducted inadequate testing without realistic combat conditions and operationally realistic threats. We would do well not to forget the lessons we learned in that time. Whenever a program's requirements were established, it must operate in combat conditions present today and in the future. Operational testing against new threats will inform Commanders in which environments the system can operate. Moreover, in instances when the threats against which we should test are contentious, as in the case of the Joint Strike Fighter, DOT&E has taken those issues to the Secretary of Defense for decisions. In all cases the current Director has done this, the Secretary has decided to adopt the Director's recommendation.

Early Testing

The final process recommendation deals specifically with costs, so it seems appropriate to mention here that the report ignores recent "shift left" strides in T&E involvement. T&E needs to be involved earlier in process to ensure appropriate resources and test conditions are part of the Requests for Proposal (RFPs) sent to industry.

DOT&E has long advocated for earlier realistic testing and problem discovery so that acquisition decision makers can make timely decisions. To help avoid expensive programs continuing in development while not delivering military utility, DOT&E now requires operational assessments (OAs) for all programs prior to the Milestone C production decision, when problem discoveries may highlight significant mission shortfalls and problems are cheaper to fix.

Early testing (both developmental test events and OAs) should inform the development process and enable the early identification of major problems. More than just providing an early opportunity for problem detection, OAs provide a chance to build knowledge on how the system will perform once placed in an operational environment. The use of design of experiments, even in early testing, allows efficient test designs that cover the operational envelope. Knowledge gained from OAs can help refine the resources necessary for the Initial Operational Test and Evaluation (IOT&E), such as the most significant factors affecting operational performance, potentially reducing the scope for the IOT&E. In ideal cases, the use of sequential test design from early testing including OAs through IOT&E can provide even more efficient use of test budgets by combining information across test phases.

Industry (Slides 13-14)

The briefing slides appear to show a fundamental misunderstanding of the different incentives between commercial industry and defense when it comes to development and acquisition of new products. In the commercial sector, companies are motivated by the threat of losing customers and failing to make profits; therefore, they are motivated to find and fix any problems before selling their products. The Defense Department, however, is the sole customer of highly complex systems, often bought in low quantities (compared to the commercial sector). The government program managers of these systems are incentivized to hide problems in order to keep their program within the resourced cost and schedule and to avoid informing leadership in the Department and the Congress of problems that might lead to funding reductions.

"Industry in general is further along in the application of software and simulation for T&E"

The claim is that "industry in general" is ahead of government in using modeling and simulation (M&S) for T&E – models reduce cost and time to market and are validated and supplemented by historical live testing data and select physical tests. However, commercial and national defense products face very different challenges in applying historical data to new systems. In defense products, historical data may not apply to new systems. Some phenomena, including especially new advances in threat systems, cannot be simulated without having collected live data first.

"Successful companies … include leadership from business and engineering organizations who together make risk tradeoff decisions, with business leader generally holding the "final vote""

As I have previously stated, in no phase of defense T&E is a tester required or asked to make acquisition decisions. DOT&E finds and presents information to the Services, which then make decisions regarding fielding. Weakening the independence of OT&E and LFT&E would not make this process any quicker or less contentious, but it would risk decreasing the quality of the performance information taken into consideration at the time of a fielding decision.

In the KC-46A example described above, the independent action of the Air Force test agencies led the contractor program management to halt unsafe testing and redesign the system. Fortunately, testing discovered the system deficiency early, and the only casualty was an already-unrealistic schedule. Boeing, a premier defense contractor (with a similarly reputable commercial industry) is reported to be losing billions of dollars over their mismanagement of the KC-46 test program.[8]

Finally, as noted above, incentives in industry---and for its leadership--- to find and fix problems are essentially the opposite of the incentives facing the Department's leadership.

[8] http://www.defensenews.com/story/defense/air-space/air-force/2016/07/21/boeing-kc46-cost-overrun-penalty-air-force-boom/87409004/

Infrastructure (Slide 12)

The independence of DOT&E allows us to require adequate and realistic operational testing and to advocate for resources to improve our T&E capabilities. I have observed that some of the most important capabilities or tests that we have prescribed have been met with substantial resistance from the Services, sometimes requiring adjudication by the Deputy Secretary of Defense. In light of the remarkable resistance from the Services to prioritize adequate testing and test assets in their acquisition programs, it is even more apparent that the independence of this office is critical to the success of finding problems before systems are used in combat.

An alarming trend I have seen during my tenure is that our threats are increasing their capabilities faster than our test infrastructure can adapt to them. Through the yearly budget review process, I have advocated for resources to improve test range infrastructure to support rigorous testing of modern combat systems. Most notably, in 2012, I convinced the Department to invest nearly $500M in the Electronic Warfare Infrastructure Improvement Program (EWIIP) to upgrade open-air test ranges, anechoic chambers, and reprogramming laboratories in order to understand the performance of the F-35 Joint Strike Fighter and other advanced air platforms against near-peer threat integrated air defense systems. The open-air test and training ranges owned and operated by both the Air Force and Navy lack advanced threat systems that are being used in combat by our adversaries today, are proliferating, or are undergoing significant upgrades; yet both services strongly resisted incorporating these modern threats that we proposed until directed to do so by the Deputy Secretary.

Other significant T&E shortfalls that I have highlighted include: Fifth Generation Aerial Target, Self Defense Test Ship, Multi-stage Supersonic Target and the Warrior Injury Assessment Mannequin for assessing force protection of ground combat vehicles to underbody blast events.

The conclusion that the infrastructure is not fully postured for the future is perhaps the sole, valuable contribution of this briefing. A follow-on study should focus on assessing the state of the art of our T&E capabilities (to include M&S) in the context of existing and anticipated theaters of operations: what are the T&E shortfalls and what does the Department need to do to close those capabilities gaps --- what would it take, how long would it take, how much would it cost and what would it buy us. The Air Force Science Advisory Board is conducting such a study now.[9] A business plan for that would also likely highlight any T&E management shortfalls that could then be more specifically addressed.

J. M. Gilmore

J. Michael Gilmore
Director

[9] Air Force Science Advisory Board, "Adapting Air Force Test and Evaluation to Emerging System Needs," Proposed 2017 study.