"Yet Another Dose of Acquisition Reform for the T&E Community — What Does the Past Tell Us About the Future?"

Thomas Christie, Former DoD Director, Operational Test and Evaluation

Keynote Address to the 2009 Annual International Test and Evaluation Association Symposium,   Baltimore, MD   September 29, 2009

Let me start by thanking Bert Johnson for this opportunity to come before you today and express a few of my thoughts regarding what has become a long lasting, almost perennial, penchant for undertaking some sort of acquisition reform on the part of our national security apparatus –in both the executive and legislative branches of our government – what it has done for us in the past, and where we might be headed in the future.

I received a call from Bert several weeks ago asking me to pinch hit for Ash Carter, Undersecretary of Defense for Acquisition, Technology and Logistics. Though quite hesitant at first, I agreed to do so. My hesitancy reflected disappointment that Dr. Carter had backed out of coming before this gathering of the T&E community. Hopefully, he understands how absolutely critical this community is to restoring our broken acquisition process and putting it back on the on the path to success after years of poor performance. To do that, Dr. Carter will need to subscribe fully -- not just to the letter of the law, so to speak, embodied in new reform initiatives, but also to their intent -- by giving the T&E community a prominent place at the table and by heeding its advice, its independent evaluations, in his program reviews and decision-making. If he doesn't, then I fear we are in for more of the same.

Clearly we need to do business differently than in the past – Defense Science Board task force reports over the past decade as well as the recent statutory reform legislation have highlighted in no uncertain terms the critical role that well-planned and executed testing and evaluation must play if we are to be successful in developing, producing and deploying effective, suitable and survivable weapon systems to our combat forces.

You will hear more about the latest DSB review of T&E and Development Test and Evaluation in particular when the Chairman of that 2008 Task Force, Pete Adolph, speaks to you tomorrow morning. Pete was eminently qualified to head this task force having served as the director the OSD DT&E office until the mid-1990s. While I will not steal Pete's thunder, suffice it to say that his report played a key role in the latest attempt on the part of Congress at Acquisition Reform, embodied in its Weapon

System Reform Through Enhancing Technical Knowledge and Oversight Act of 2009.

After more than four decades of supposedly well-structured defense planning and programming, as well as numerous studies aimed at reforming its multi-billion dollar acquisition system, one can hardly avoid the conclusion that the decision process governing our defense establishment is broken and our forces are not getting systems with the performance they need in a timely and effective manner. The evidence supporting the need for drastic action abounds.

Despite the largest defense budgets in real terms in more than sixty years, we have a smaller military force structure than at any time during that period, one that is equipped to a great extent with worn-out, aging equipment. To highlight just one example of this critical problem – the fighter-attack forces of our Air Force are roughly half the size they were in the mid-80s and shrinking still; the average age of the fleet is twice what it was or should be and growing; and the prospect of producing enough aircraft to reverse these truly disturbing trends seems to be growing dimmer day by day.

Granted, the employment of our forces in Iraq and Afghanistan has contributed to the wear and tear on our combat and support equipment. The bill for repairing and replacing that equipment – no doubt in the hundreds of billions -- is mostly yet to be faced. And, more to the point, this only exacerbates the already severe modernization problems faced by all three services. Those problems have been on the horizon for decades and would have plagued our forces even if the war on terror had not evolved as it has since 2001.

A fundamental source of DOD's problems is the historically long pattern of unrealistically high defense budget projections into the outyears, combined with equally unrealistic low estimates of the costs, schedules and technical risks associated with new programs. As we have seen, time and again, this make-believe world of defense programming and budgeting, however, inevitably runs head-on into the real world of weapons development and procurement, a world where hardly a week passes without some new horror making the headlines. New systems critical to the Service modernization plans for their aging forces encounter huge cost overruns, serious technical challenges and lengthy schedule slips that call into question the affordability and realism of plans for growing and/or sustaining present force levels and structures.

I believe present Pentagon leadership is painfully aware of the need for drastic action, particularly in the world of acquisition. The actions taken on several major acquisition programs earlier this year by Secretary Gates is an indication of a painful recognition that we must do things differently if we are to field and maintain the world's best combat forces.

Likewise, Congressional initiatives, particularly this year's Weapon System Acquisition Reform Act, are further evidence of the crying need to do business differently in the future.

But, haven't we seen and heard all this before? Over the past 40 years or so, I've observed and participated in numerous Defense and Acquisition Reform exercises, Defense Science Board task forces, working groups, IPTs -- you name it. The Blue Ribbon Commission of 1970; the Carlucci Initiatives of 1981; the Packard Commission of 1986; the Defense Acquisition Performance Assessment or DAPA of 2006; and now we have probably the most far-reaching reform legislation in this year's authorization bill since Goldwater-Nichols in the mid-1980s. To use an overused cliché, I've been there, done that -- time and time again.

Just recall a few of the initiatives and buzz-words of the more recent past: Capabilities-Based Force Planning; Simulation-Based Acquisition; Commercial-Off The Shelf Procurement or COTS; Cost As the Independent Variable or CAIV; Spiral Development; Total System Performance Responsibility or TSPR of the 1990s. What have these much-ballyhooed initiatives done for us? What's our record since the 1990s? This year's GAO report on the defense acquisition system answers these questions. The 96 acquisition programs they evaluated showed an average of more than 40% cost growth in RDT&E alone, 25% cost growth in total acquisition costs and an average schedule slip of two years.

DAPA's statement of the problem it was established to address says it all :

"The current situation is characterized by massively accelerated cost growth in major defense programs, lack of confidence by senior leaders, and no appreciable improvement in the acquisition system despite many attempts in the past."

Some of these reviews over the years have been chartered with the explicit goal of "streamlining" the so-called burdensome acquisition process. Too often, unfortunately, this streamlining mania has translated into casting the blame for the growing costs and development times onto oversight activities, to include, of course, the testing community – both developmental and operational – as well as programming and cost analysts. The thinking has been "if we could only get these testers, these analysts off our backs, we could get new systems into the field on cost, on time, on performance.

This key misconception should be put to bed right up front -- a finding borne out in spades during a DSB review of acquisition carried out in the 1990s. *While oversight*

by government agencies and their reporting requirements can indeed be burdensome, they clearly are not the causes of the continuing miserable record of program stretch-outs and cost growth. This is true independent of whether those agencies and their reporting requirements are internal to DoD, such as the Defense Contract Management Agency (DCMA), Independent Cost Analysis groups, Operational Test and Evaluation organizations; or those external, such as the congressional committees and their staffs and the Government Accountability Organization (GAO).

Instead, that review, covering some 100 major defense acquisition programs, concluded that failure to identify and admit to technical issues and solutions, as well as real costs, before entry into what was known as Full-Scale Engineering Development (FSED) – now referred to as Engineering and Manufacturing Development (EMD) -- was the overwhelming cause for subsequent schedule delays, often in terms of years, and the resulting cost growths. To the extent oversight played any role in these delays, it was the discovery and reporting of test failures – primarily in development testing --during EMD that often necessitated additional time and dollars for system redesign, testing and retesting of fixes, as well as costly retrofits of those fixes.

Despite the overwhelming evidence that oversight per se was not the cause of continuing cost and schedule growth, Pentagon leadership in the mid-1990s implemented the strategy known as Total System Performance Responsibility (TSPR) for many key acquisition programs. This strategy, in essence, relieved development contractors of many reporting requirements, including costs and technical progress. In essence, it built a firewall around the contractor, preventing government sponsors from properly overseeing the expenditure of the taxpayers' dollars.

Accompanying this new approach to acquisition, the Department in essence savaged its in-house technical capability and its ability to gain insight into the technical and cost issues associated with the new programs. For all intents and purposes, the OSD Development Test and Evaluation office went away and the Service DT&E organizations and capabilities were stripped of much of their ability to plan and conduct or oversee meaningful government testing.

During my time as the DOT&E in the first half of this decade, I saw the results of this misguided policy. Time and again I sat in program review meetings, including numerous DABs, where I was struck by the lack of credible information concerning the status or the results of development testing to date. In case after case, Pentagon decision-makers acquiesced in programs entering EMD and even low-rate initial production before technical problems were identified, much less solved; before credible independent cost assessments were accomplished and included in program budget projections; before critical technologies were shown to be sufficiently mature;

and even before the more risky requirements were demonstrated in testing. The overwhelming evidence clearly points to a problem with the DoD acquisition system itself that cannot be written off to poor management of individual programs.

As an aside, I should admit that just having an informed voice for testing – whether developmental or operational – or, for that matter, independent cost estimates, didn't necessarily lead to hard-nosed decisions. Speaking from my own experience as the DOT&E from 2001 to early 2005, my office was responsible for producing roughly 30 Beyond Low-Rate Initial Production, or BLRIP, reports to the Secretary of Defense and Congress. By law, these reports are a prerequisite for any full-rate production decision. These reports assessed over half of these systems to be either not operationally effective or not operationally suitable, or both. In not one case was one of these programs stopped as a result of the information available in the reports or presented at the production DAB. To be fair, many of these production decisions were made with accompanying direction to conduct follow-on testing to verify the fixes designed to address the problems uncovered in operational testing.

Even though DoD's acquisition policies and directives have most often reflected many of the more substantive findings and recommendations from the many reviews of the process over the years, too often the managers of this process have lacked sufficient information, if not the will, to be able to carry through and implement them in program decisions.

Over the years, this process has been further warped by the ever-optimistic projections of available funding both in the near-term as well as into the outyears, in essence relieving the decision-maker of any need to make the hard choices.

Part and parcel of the effort to sell new programs – to get the camel's nose under the tent, so to speak – is the so-called "buy-in" syndrome, whereby costs, schedule and technical risks are often grossly understated at the outset. These low-ball estimates mesh right in with the optimistic overall budget top line projections into the outyears, especially the procurement accounts.

The root of the problem is well known: Review after review has found that we should *"fly and know if it works and how much it will cost before buying."* Building and testing prototype systems and subsystems – hopefully on a competitive basis -- before proceeding with EMD has been a perennial recommendation. In that same vein, these reviews have called for up-front funding of robust efforts to demonstrate technology maturity through testing as a prerequisite for program approval.

DoD's acquisition policy and directives have incorporated these recommendations. Unfortunately, the rising operating and support (O&S) costs of the existing forces, and the fact that there are more acquisition programs being pursued than DoD can

possibly afford in the long term, have combined to intensify the competition between programs for dollars. This, in turn, has led to decision-makers sanctioning low-balled program costs and overly optimistic schedules at the outset of major programs, most often at the expense of building and testing prototypes and critical early-on technology risk reduction efforts.

Having obtained approval to enter EMD with unrealistic costs and schedules based on rosy, if not surreal, technical risk assessments, programs inevitably encounter problems early-on. These problems, in turn, set off the spiral of schedule stretches and ballooning costs that have come to plague the vast majority of DoD acquisition programs. Unfortunately, too often, program managers attempt to limit the damage by trying to maintain the schedule at the expense of critical test events and design fixes for obvious deficiencies. The net result is a schedule-based strategy, rather than the event-based program strategy that the myriad of DoD acquisition directives stress.

The past several years, particularly after U.S. forces entered combat in Afghanistan and Iraq, the pressure has intensified even more to keep programs on schedule, even to accelerate the process, in order to get equipment in the hands of troops sooner than later. As a result, some systems with serious reliability and maintenance problems found in development and operational testing have been waived through the decision process into production and deployment.

It appears that too often programs fail to carry out adequate testing; and in those cases where they do, they often fail to take the necessary corrective actions based on that testing before proceeding with full production and deployment. The Defense Science Board Task Force on Development Test and Evaluation headed by Pete Adolph reported that, in the ten-year period, 1997 through 2006, roughly 70 percent of Army systems had failed to meet their specific reliability requirements in operational testing. And of those that failed, the vast majority failed to reach even half their reliability criteria. The Task Force found that similar problems existed with the programs of the other services. These problems with attaining realistic reliability, availability and maintainability goals resulted in increased logistics burdens on our operating forces and a *de facto* reduction in force effectiveness attributable to low equipment availability.

What is disturbing about these failures is that these most of these programs should not have been cleared to enter OT&E in the first place. They clearly had not completed development testing successfully – they had either failed to meet effectiveness or suitability requirements in DT&E or, in some cases, had truncated planned DT&E in order to stay on schedule or to stay within costs.

The MV-22 is a good example of a major program that encountered technical and cost problems in EMD, yet attempted to hold to a schedule that provided little, if any, slack to address those problems. Clearly, after nearly 20 years in development at the

time, the urgency of replacing the aging CH-46s drove decisions to severely reduce development testing in EMD in order to save dollars and stay on schedule.

The official investigation into the tragic accident that occurred in April 2000 drives home this point. That accident involved the crash of an MV-22 during an operational test mission and resulted in the deaths of 19 Marines. The official report from the investigation states that there were three test events flown as part of the MV-22 EMD Flight Control System Development and Flying Qualities Demonstration (FCSDFQD) Test Plan investigating the phenomenon known as power settling.

As the report notes, the original Test Plan called for 103 test conditions to be flown. In an effort to recover costs and schedule, the conditions to be tested were reduced to 49, focusing on aft center of gravity conditions that were thought to be most critical. Of the 49 conditions, 33 were actually flight-tested. Thus, in order to save dollars and make up for schedule slips, the important testing was severely curtailed – roughly one-third of the planned test events were actually flown – and particularly critical test points were not flown at all.

This series of events, culminating in the April 2000 accident and another crash in December of that year, brought the program to a halt, nearly resulting in termination. In the end, the MV-22 program recovered, executed the full range of technical testing that should have been done previously, went through its second OPEVAL, and was approved for full production in 2005, nearly 25 years after the decision to initiate the program.

Unfortunately, there are many other examples of major programs curtailing DT&E and rushing into OT&E before development is complete; of programs failing to demonstrate the required levels of maturity for critical technologies; of programs that entered EMD before they were ready, resulting in major schedule delays -- accompanied by inevitable cost growth.

While we can all bemoan the fact that the past has witnessed an acquisition system and process gone awry, we need to look to the future and hope that the past is not a prologue to further horror stories in the days to come. This community, this test and evaluation community – both development and operational testers, both government and industry testers – has an extremely critical role in making our acquisition system work as it should. It stands ready to do so – hopefully, it will be permitted to exercise that role.

This critical role for T&E has been emphasized repeatedly by DSB Task Forces, by the National Academy of Sciences, by the Government Accountability Office and, most recently, by this year's Acquisition Reform Act.

Let me briefly mention some of the findings from these reports.

A DSB Task Force on T&E in the late 1990s concluded that *The fundamental purpose for testing is to discover and to learn. The testers as well as the developers and users are the learners in this process. In order to build mutual trust and confidence, and to focus on system operating characteristics that can be tested in a way that makes sense, the expertise of the testers should be sought by the users and developers as the system requirements are being formulated. This means that the testing community should be a part of the requirements development process.*

The Pentagon's Section 231 report responding to a congressional requirement in the FY 2007 Authorization bill to evaluate the department's T&E practices and their effectiveness stated among other things:

• *All T&E should concentrate on user needs and measuring improvements to mission capability and operational support.*

• *All testing and evaluation programs should experiment, learn about the strengths and weaknesses of system and its components, and help incorporate results into system enhancement initiatives.*

• *Testing should continue through the entire life-cycle of a weapon system.*

• *Evaluation should be in the mission context and in terms of the test results' operational significance.*

• *Evaluations should be based on a comparison against current mission capabilities so that the improvements can be measured.*

• *Evaluations should include all available information, whether obtained through DT&E or OT&E or operational experience, whether through testing conducted by the government or by industry.*

While I stated earlier that you would be getting more detail tomorrow from Pete Adolph reporting on his 2008 DSB review of DT&E, let me steal a little of his thunder with a couple of their key results:

• *Changes over the past 15 years have had a significant negative impact on DoD's ability to successfully execute increasingly complex acquisition programs. A major contributor to this problem includes massive workforce reductions in acquisition and*

test personnel compounded by the high retirement rate of the most experienced technical and managerial personnel in both government and industry without an adequate replacement pipeline. As a result, a significant amount of DT is currently performed without sufficient government involvement or oversight.

• As a minimum, government test organizations should develop and retain a cadre of experienced T&E personnel to

> ➢ *Participate in the translation of operational requirements into contract specifications, and in the source selection process, to include RFP preparation*

> ➢ *Participate in DT&E planning, including TEMP preparation and approval*

> ➢ *Participate in technical review processes*

> ➢ *Participate in test conduct, data analysis, evaluation and reporting.*

When the new Congress convened this past January, it was clear that there would be a push to enact yet more acquisition reform legislation. Early on, it appeared that the major thrusts of this evolving legislation would aim primarily at arresting the rampant cost growth in many of the major acquisition programs, at strengthening such supposed brakes on cost growth as the Nunn-McCurdy act. Unfortunately, to many of us, there didn't appear to be sufficient emphasis on addressing critical T&E issues, particularly in the world of DT&E, which so many reviews had highlighted in the past.

In particular, as I mentioned before, failure to identify and admit to technical issues and solutions, as well as real costs, before entry into what was known as EMD -- was the overwhelming cause for subsequent schedule delays, often in terms of years, and definitely the root cause for most of the resulting cost growths in RDT&E. Programs entered EMD and even low-rate initial production before technical problems were identified, much less solved; before planned DT&E had been executed, demonstrating in testing that the more risky requirements were attainable and that critical technologies were sufficiently mature. There was convincing evidence that the lack of a credible OSD DT&E oversight capability, accompanied by drastic cuts in the Service DT&E organizations, had led directly to these fiascoes.

Accordingly, there was an early initiative to bring to the attention of the Armed Services Committees the need for action to restore the T&E capability reductions of the late 1990s and the early years of this decade. These overtures were successful in

that early draft versions of the legislation began to incorporate such actions.

Then, when the Senate and House versions of the reform went to the conference committee, Pete Adolph, Jack Krings, the first DOT&E, and myself signed a memo to Senators Levin and McCain requesting a restoration of a robust and independent DoD DT&E capability. I quote from that letter:

We wish to express our strong endorsement of the Levin-McCain bill approved by the Senate Armed Services Committee that reestablishes the Director, Developmental Test and Evaluation (DDT&E) office as a direct report to the Under Secretary of Defense for Acquisition, Technology and Logistics. This office is the key to providing the requisite unfiltered assessments of program progress through developmental testing and evaluation (DT&E), as well as independent evaluations of technology readiness levels. The absence of independent assessments of developmental readiness and progress is a root cause of the systemic DoD acquisition problems.

Unfortunately, since the late 1990s, OSD has not accomplished meaningful or effective oversight of DT&E. As a result, the visibility and transparency required to perform effective program reviews at the OSD level no longer exists.

The final legislation, known as Public Law 111-23 and signed into law in May of this year, requires several changes in the Pentagon's acquisition organization and process. Of particular interest here is the much–needed restoration of the OSD DT&E office as well as direction to restore DT and other technical capabilities in the Services. The statute requires the Secretary to select officials to serve as the Director for Developmental Test & Evaluation and Director for Systems Engineering, with responsibilities for issuing joint guidance relating to the integration of developmental test and systems engineering. These Directors are responsible for leading the developmental test and systems engineering workforces within DoD. This legislation also requires the military departments and defense agencies to develop and implement plans to ensure they have the appropriate resources for developmental testing and systems engineering, and the Directors are required to assess these plans and report their assessments directly to Congress.

All of that sounds great on the surface, but as I mentioned earlier, we've been down these roads before. I am deeply concerned that, four months after the legislation, that Director for DT&E has yet to be named and the needed steps to staff not just that directorate, but those of the Service RDT&E organizations, with the talent they so sorely need, appears to be dragging, if moving at all.

Recently, Dr. Carter was quoted as saying he is encouraging Defense Department managers to come forward with problems as soon as possible and seek fixes. *"If they bring a problem to you, you will help them solve it; if they hide it from you, you won't forgive them,"* he says.

He went on to say, *"There is another kind of acquisition reform, which is what to do when programs are not performing or are not needed anymore. We have a responsibility to acquisition at the back end as well. When we find programs that are troubled, that aren't performing in cost, schedule or performance or that are no longer needed, then we need to stop doing them."*

Dr. Carter, I implore you to recognize that you have a community that stands ready to help you in implementing the real reforms that are so critical to the success of our acquisition system. Get out in field and meet them where they are on your test ranges and other test facilities – both government and industry. Give this community a seat at your table and listen to their advice. They will not shy away from bringing forward to you very early in the game system problems as well as proposals for their solution. They will not hide these problems if you give them a voice.

This community I speak of -- this T&E community, in both government and industry, both development and operational testers – has served the department very well over the years, despite its ups and downs. The success of our operational forces over the years reflects their dedication to developing and deploying systems proven effective and suitable on our ranges and in our facilities.

There is a new world dawning now as you, Dr. Carter and your fellow Service acquisition leaders undertake to right this defense acquisition ship – so to speak – and put it back on a course to success. I am confident that this T&E community can rise to the challenges ahead and, together with you and the rest of the acquisition community, see to it that our soldiers, sailors and airmen are equipped with the best equipment our nation can provide.