Guest Editorial

ITEA Journal 2013; 34: 6–10

Perspectives on Developmental Test and Evaluation

Frank Kendall
Under Secretary of Defense for Acquisition, Technology, and Logistics
The Pentagon, Washington, D.C.

During my first tour in the Pentagon in Acquisition, Technology, and Logistics (AT&L) from 1986 to 1994, I was responsible initially for strategic defense systems and then for tactical warfare programs. During this time, I had the opportunity to work with a Developmental Test and Evaluation (DT&E) organization that was very professional and led by an outstanding civil servant, Pete Adolph. Somewhere along the way, as priorities and personalities changed in the Office of the Secretary of Defense (OSD), the DT&E organization atrophied and all but disappeared. For the last few years, under the auspices of the Weapons System Acquisition Reform Act, we have been strengthening the DT&E organization within OSD. Ed Greer, who retired from public service recently, has rebuilt the DT&E organization to the point that it is now performing a role much closer to the one I remember from the 80s and 90s. As Defense Acquisition Executive, I rely heavily on the DT&E office and staff for sound advice on the adequacy of the test programs being proposed for major programs and on the implications of developmental test results for investment decisions, particularly for entry into low rate production. Developmental testing is a core activity in our acquisition programs, however, not just an OSD oversight function. In this article I discuss the role DT&E plays in our programs, some important principles I believe should be applied to developmental testing, and some common problems I have encountered that relate to the effectiveness of DT&E.

Role of developmental testing

The purpose of developmental testing is simple: to provide data to program leadership so that good decisions can be made as early as possible. I have a sign outside my office displaying a quote from W. Edwards Deming: "In God we trust, all others must bring data." It is our developmental testers who "bring the data" needed to make sound decisions during product development. Programs are organized in various ways, but whatever the specific organizational model, testing is the source of the crucial information that provides feedback to program management, chief engineers, lead system engineers, integrated product teams, and military users on whether their designs meet requirements or not. The spectrum of testing types and venues that is captured in compliance matrices for system specifications runs the gamut of laboratory testing and field testing. All of these sources of information can be valuable, but integrating them into a test program and an overall program plan and schedule that meet the needs of developmental testers' customers requires a high degree of professionalism and a deep understanding of how test results can influence design and program decisions. In my experience, a well-structured test plan makes all the difference in whether a program is efficiently executed or not. There are two layers of DT&E organizational roles and relationships; both are important in determining DT&E's contributions to program success.

The first layer of DT&E organization exists within the program office. I have seen several organizational

models for DT&E offices within Department of Defense (DoD) programs, and any of them can work given professional leadership, well-defined lines of authority, and responsibility, and commitment to working together as a team. The DT&E office or organization within a program usually reports to the program manager, to the chief engineer, or to the lead systems engineer. In some cases, the DT&E staff can be matrix staff allocated from centralized functional test organizations, and in other cases, the testing staff can be organic staff members of the program office. Whatever the model, the role of the test organization is to support the program's leadership by providing timely, accurate, and relevant information to enable efficient and effective program decisions.

The second layer of DT&E organization exists within the Service or Military Department at a higher level than the program office. Here too there are various models, and any of them can be successful. Some Services have centralized DT&E support within test organizations that include operational test as well as DT&E. Others have created DT&E organizations at the system command level. These organizations tend to be focused on ensuring the acquisition and evaluation of the specific data needed to support major decisions, such as initiating production or proceeding to Operational Test and Evaluation (OT&E). This layer of DT&E organization, with some degree of independence from the standard acquisition chain of command of program executive officer and program manager, and even in some cases the acquisition executive, can be effective, but it also runs the risk of diluting the authority and accountability of the acquisition chain of command. In my own OSD AT&L organization, I consider the DT&E organization, which we have rebuilt over the last few years, to have a staff function that supports my acquisition decisions and also provides expertise and other support to the Services. When there are differences of opinion between the OSD DT&E organization and the Service acquisition chain of command, I expect them to be brought to my attention for resolution.

Precepts of effective DT&E

The following "precepts" are based on my own experience and are generated largely from a program or engineering management perspective. They are in no particular order and are intended merely as food for thought by anyone involved in DT&E or any customers or stakeholders in the DT&E functional area.

1. Contribute to program efficiency and effective execution: DT&E is a support function that enables sound design and program decisions, and DT&E leadership should be an integral part of the program planning team. DT&E should be part of program planning from the outset. Much of product development can be thought of as risk management, where design and technical risks are addressed and resolved in an iterative process over time. The way DT&E is structured to contribute to this process can make all the difference in the efficiency (think waste avoidance) with which a product is developed. DT&E leadership should be fully integrated into the program management and system engineering functions. Formal "design of experiments" techniques are being used widely now to ensure that tests are structured to extract meaningful information as efficiently as possible, and I applaud this development. Testing isn't free, however, and we need to balance the desire for thorough testing against the resources in time and money required to conduct the testing. This can only be accomplished through a cooperative effort that fully involves DT&E professionals in the program planning process.

2. Provide relevant information as early as possible: Once a program enters Engineering and Manufacturing Development (EMD), the commitment to design for production unleashes a marching army of interdependent engineers that needs to keep moving in a tight formation through the development process. Any serious design problems that surface late in the development process can stop this marching army in its tracks at great expense while the problem is addressed and resolved. The later a problem is identified and the solution determined, the greater the redesign burden and cost. To avoid this problem, information on the performance of the design in key areas needs to be made available as early as it can be provided and from the most reliable source of information available. As good as our design tools have become, there is still no substitute for physical testing, particularly for our more complex and novel designs. For key program technical risks, the early use of prototypes (full or subsystem level) and developmental testing during technology demonstration risk-reduction activities prior to the commitment to EMD can make all the difference between a successful EMD and one that experiences massive overruns. Again, DT&E isn't free, and like any program, it needs to be conducted as efficiently as possible, but the real benefit of an effectively structured test program is in the cost avoidance it can provide by discovering problems as early as possible.

3. Integrate DT&E planning across the product life cycle: DT&E is not just about production representative prototype testing in a controlled environment prior to the decision to proceed to OT&E. It encompasses the total program of testing, including, for example,

hardware in the loop testing in system integration laboratories, environmental stress screening at the subsystem level, and software testing in emulators. Whenever data are needed to support risk reduction, design validation, and requirements verification, there is a role for DT&E in collecting those data and evaluating them on a continuum over a program's life. This spans all phases of a program's life cycle. Increasingly, the Department is keeping systems longer and upgrading them in lieu of pursuing new designs. Effective DT&E is as central to these efforts as it is to new product development programs. Well-structured developmental testing should be integral to all phases of a product's life cycle.

4. Focus on support to internal program decisions and verification of compliance with requirements: DT&E does not exist in a vacuum and is not a separate function; its purpose is to support program management and technical leadership as it works to develop and field a product that meets user requirements. Programs move through a series of development activities that must be successfully completed and verified through testing, often as a condition of proceeding to the next phase of the program. Sometimes this is the next software build; sometimes it is a higher level of integration, and sometimes it is a decision to commit to initial production. DT&E also provides an indication of the readiness of a program to proceed to OT&E. For any of these decision points, DT&E provides crucial information to support the decisions, and the adequacy of that information is central to controlling program risk and ensuring contractual compliance. Careful planning and well-defined decision criteria are necessary prerequisites, but the discipline to enforce those criteria is what often sets successful programs apart.

5. Use DT&E to improve the efficiency and validity of OT&E: OT&E is conducted with more independence from the program office and the acquisition chain of command than DT&E and with less involvement by the contractor supplying the product, but the two test regimes should work together to complement each other and avoid unnecessary expenses as much as possible. Under Mike Gilmore's and Ed Greer's professional leadership, there has been a very cooperative relationship between the DT&E and OT&E organizations at the Department level. This relationship should continue and be mirrored at all levels. While the OT&E community works hard to preserve and ensure its independence, I am encouraged by the willingness of that community to use the data that DT&E can provide to augment and complement data provided by OT&E. We will never have the resources to do as much testing as we would like, and achieving statistically meaningful testing is sometimes prohibitively expensive. By working together, the DT&E and OT&E communities in OSD and the Services can achieve more valid results, anchor each other's efforts, and do so at less cost.

How we get into trouble in DT&E—Some of the ways at least

There are times when DT&E doesn't fulfill its purpose, and a program ends up with one type or another of acquisition problem. This can take the form of cost overruns and schedule slips, or worse, a product that simply isn't viable, despite having been approved for development and even initial production and after years of effort and expense. The following paragraphs provide some of the types of problems I have encountered most frequently over the last 40 years.

In the technology demonstration or risk reduction phase, we permit the use of test articles that may not be adequately representative of the actual product design. In these cases, the testing that is conducted may be more intended to sell a product than to reduce that product's risks. Motivated by a specific example I encountered (a program that was up for a Milestone B decision), I recently asked a former deputy director of the Defense Advanced Research Projects Agency (DARPA) to review a number of programs that had been through technology demonstration programs, which included DT&E of competitive prototypes. The results were troubling. In the majority of the cases, the design that was demonstrated had little or no correlation to the design that was going to be developed in EMD. The DT&E that was done in the risk reduction phase was not providing data to reduce the risk of the target design. It was providing data intended to sell the government on the prospective bidder. The lesson I derived from this was that the combined government management team (program management, engineering leadership, and developmental testers) was not insisting on the relevance and validity of the test program. We can't blame industry for trying to win the EMD contract; we have to blame ourselves for not understanding industry's motivation and insisting on meaningful testing that actually addresses the risks in the intended design.

We use ill-defined user requirements that have not been translated into testable technical specifications. As a result, we cannot plan the time and resources for appropriate testing in the early stages of a development program, and we cannot hold the contractor responsible for not meeting our expectations. The government generally has to define its requirements and ensure that they are converted into testable requirements that our contractors can demonstrate they have satisfied in

DT&E. If we fail in this responsibility and provide vaguely defined requirements to industry, we have no one but ourselves to blame when our expectations are not met. The largest program I ever worked on had extensive user requirements that were never properly defined to the prime contractor or converted by that contractor or the government into quantifiable and testable technical requirements. When the program eventually died of its own weight, years after it had started and after billions of dollars of cost, the prime contractor and the customer were still debating over how to interpret the requirements and how to test for compliance.

We have to resist the tendency to assume DT&E efficiencies that exceed previous experience in response to financial pressure. Most programs come under financial pressure at some point; often before the program even enters EMD. It is far too easy to assume away the need for an adequate number of test articles, or an adequate amount of test time in order to meet a budget number or a schedule that has been dictated for some reason. Usually in my experience, program leadership, including the DT&E leadership, accepts the constraints that have been provided and gambles on unprecedented test performance and efficiencies. The usual result is increased inefficiency, not the opposite. We don't want to over schedule or buy unneeded test assets, but my experience is that the far more common errors are unwarranted optimism and acceptance of excessive risk rather than excessive conservatism or risk aversion. If we have solid reasons to conclude that we can improve the efficiency of DT&E (and we should always be looking for sources of efficiencies), then we should take those efficiencies into account in our planning, but hope is still not a method.

We sometimes fail to conduct adequate DT&E prior to the decision to start production. About a year ago, I called a particular decision to enter production on an aircraft program without flight testing "acquisition malpractice." If a product enters production before the design is stable, the resulting waste in cost increases and schedule slips can be dramatic, and the program is much more likely to be canceled. I stress solid, well-defined DT&E results as an important prerequisite for this decision because the pressure to enter production can be overwhelming, and doing so prematurely has major consequences. The Service often feels that it will "lose the money" that has been requested a year or more earlier from the Congress if the production contract is not awarded. Industry wants to make the sale, and the user is anxious to get the new product. The decision to enter production is all but irreversible, and to make this commitment for a new design without the knowledge obtained from adequate DT&E entails high risk. That said, there is a balance to be struck. A well-structured DT&E program will provide confidence in the stability of the design as early as possible. Some degree of concurrency between development (including DT&E) and production is usually appropriate. The degree of concurrency that is acceptable depends on several factors, but in every case there should be a well-defined basis rooted in data provided by DT&E to support this critical decision.

We assume untested design fixes to problems discovered in DT&E will be successful, in order to preserve schedule. It is always a judgment call, but in general, design changes have to be verified through DT&E just as much as the original design needs to be verified. Where I have seen this most often is when we are about to initiate or have already initiated low rate production. I recently slowed the rate of production of DoD's biggest program so that we could test design fixes adequately prior to increasing the rate of production. I seriously considered stopping production completely, but made the judgment call to continue at a low rate while the test program verified the design fixes. The cost of stopping and restarting would have been very high, so I limited our exposure but didn't take it to zero. We don't want to be in this position if we can avoid it.

We sometimes over-focus on DT&E as preparation for OT&E. No one wants to fail operational testing, and one of the things we can learn from the last stages of DT&E is whether or not a program has a high probability of a successful OT. This doesn't mean, however, that we should do two rounds of OT&E with the first being called DT&E. In general OT&E is not intended to be a place to discover unanticipated problems, but we shouldn't be so risk averse that we add what amounts to an extra phase of testing out of concern for failing operational test. DT&E should be focused on verifying that the contractor has met the requirements. We should do an effective job of linking those requirements and the DT&E that verifies compliance to the operational performance that we intend to demonstrate in OT&E. If we have done this effectively, the last stage of DT&E shouldn't have to be a full dress rehearsal for OT&E.

The bottom line

Developmental testers are critical professionals who make a major contribution to DoD's programs. They bring a unique body of knowledge to the table that is essential to effective program planning and execution. Again, it is largely the DT&E community that "brings [the] data" the sign outside my door emphasizes. Working with program and engineering leadership as

key members of the management team, developmental testers provide the information that makes program success possible and much more probable. ❑

Frank Kendall is Under Secretary of Defense for Acquisition, Technology, and Logistics (AT&L). Previously, he was the Principal Deputy Under Secretary and also the Acting Under Secretary. He was vice president of engineering for the Raytheon Company and is an attorney who has been active in the field of human rights, working primarily on a pro bono basis. Mr. Kendall is a former member of the U.S. Army Science Board and the Defense Intelligence Agency Science and Technology Advisory Board, and he has been a consultant to the Defense Science Board and a Senior Advisor to the Center for Strategic and International Studies. He also spent 10 years on active duty with the Army; serving in Germany, teaching engineering at West Point, and holding research and development positions. Mr. Kendall is a distinguished graduate of the U.S. Military Academy at West Point and holds a master's degree in aerospace engineering from the California Institute of Technology, a master of business administration degree from the C.W. Post Center of Long Island University, and a Juris Doctor degree from Georgetown University Law Center. Over the course of his career as a public servant, Mr. Kendall was awarded the following federal civilian awards: Defense Distinguished Civilian Service Medal, Secretary of Defense Meritorious Civilian Service Medal, Presidential Rank Award of Distinguished Executive (Senior Executive Service), Presidential Rank Award of Meritorious Executive (Senior Executive Service), and Army Commander's Award for Civilian Service. He also holds the following military awards (U.S. Army): Meritorious Service Medal with oak leaf cluster, Army Commendation Medal, and National Defense Service Medal. E-mail: Frank.Kendall@osd.mil

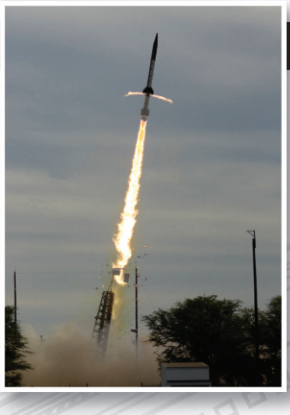
The R&D "Use" Of Propulsion Vehicles
WORKSHOP

April 2-4, 2013
Las Cruces, New Mexico

Hosted by the White Sands Chapter of the International Test and Evaluation Association

ITEA

Connect with ITEA to Learn, Share, and Advance!



The White Sands Chapter of ITEA is hosting the inaugural workshop on Propulsion T&E. The theme and focus of this workshop is to bring the community of "Rocketeers" together to examine the past, present, and future propulsion T&E techniques and procedures. The workshop focuses on booster fuel options, hypersonic, telemetry instrumentation data reduction and data analysis, guidance and navigation, and payload requirements. This workshop will facilitate crosstalk among the Services, Government Agencies, Industry, Academia, and T&E practitioners and provide us with the experience base and lessons learned to move forward.

TOPICS

Booster Fuels | Hypersonics | Instrumentation
Guidance & Navigation | Payloads | Other

REGISTER TODAY www.itea.org