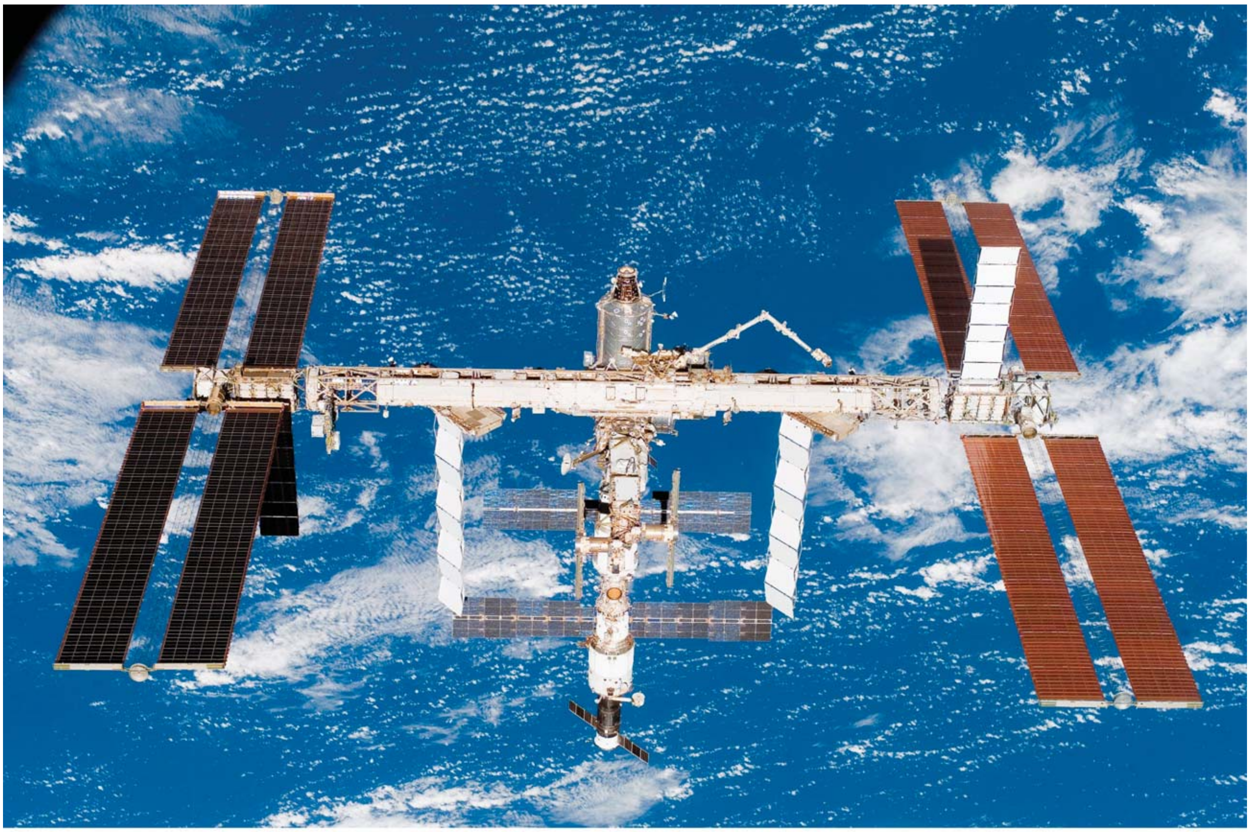
The

ITEA Journal

March 2008
Volume 29, Number 1



S118E09416

The Evolving T&E Infrastructure

Published quarterly by the International Test and Evaluation Association

CONTENTS

ITEA BEST PAPERS

TECHNICAL ARTICLES

DEPARTMENTS

ON THE COVER: Disruptive technology arises not as a linear evolution of the existing state of the art but as a non-intuitive leap in thinking and capability—the ultimate manifestation of innovation—and its impact is often far beyond that expected. Wireless devices, robotics, digital technology and the Internet, biotechnology, hydrogen fuel cells, nanotechnology, and nuclear weapons are all examples of past (or potentially future) disruptive technology. Some of them have brought about entirely new industries and changed society. Cultivating innovation, preparing for and responding to breakthrough ideas and capabilities, closing gaps between the research and test communities, and more rapidly transferring technology and capability from science and technology programs to T&E are challenges and opportunities for the test community. (Photo of unmanned robotic system undergoing testing in the deep snows of Fort Greely, Alaska, at the Cold Regions Test Center courtesy of Yuma Proving Ground Public Affairs; hydrogen fuel cell photo courtesy of Matt Stiveson, National Renewable Energy Laboratory, Golden, Colorado; scanning tunneling microscope image courtesy of J. A. Stroscio, A. Davies, D. T. Pierce, and R. J. Celotta, National Institute of Standards and Technology. Cover design courtesy of Headquarters, U. S. Army DTC, Aberdeen Proving Ground, MD.)

■ ITEA Headquarters: 4400 Fair Lakes Court, Suite 104, Fairfax, Virginia 22033-3899; Tel: (703) 631-6220; Fax: (703) 631-6221, E-mail: *itea@itea.org;* Web site: *http://www.itea.org.*
■ ITEA is a not-for-profit international association founded in 1980 to further the development and exchange of technical information in the field of test and evaluation.
■ *The ITEA Journal* (ISSN 1054-0229) is published quarterly by the International Test and Evaluation Association at 4400 Lakes Court, Suite 104, Fairfax, Virginia 22033-3899. Single issue cover price for *The ITEA Journal* is $20. ITEA membership dues are $45 for individuals, $25 for full-time students, and $745 for corporations. Annual dues include a one-year subscription to *The ITEA Journal*. The annual subscription rate for libraries and other organizations providing timely reference material to groups is $60. All overseas mail (air mail or AOA) requires an additional $20. *The ITEA Journal* serves its readers as a forum for the presentation and discussion of issues related to test and evaluation. All articles reflect the individual views of the authors and not official points of view adopted by ITEA or the organizations with which the authors are affiliated.

POSTMASTER: Send address changes to: ITEA, 4400 Fair Lakes Court, Suite 104, Fairfax, Virginia 22033-3899.

ITEA Journal 2008; 29: 3–5

2003 National Defense Authorization Act—A U.S. Army Developmental Test Command Perspective

James B. Johnson
U.S. Army Developmental Test Command, Aberdeen Proving Ground, Maryland

The U.S. Army Developmental Test Command (DTC), a subordinate command activity of the Army Test and Evaluation Command (ATEC), is charged to plan, conduct, and report developmental and production tests (including Title 10, United States Code, Live Fire; virtual; simulated; and other tests), across the full spectrum of environments, in addition to verifying the safety of Army systems. To fulfill this mission, DTC maintains and operates seven test centers throughout the continental United States. Seven sites within these test centers are designated as Major Range and Test Facilities Base (MRTFB) activities.[1]


James B. Johnson

The Last Decade. During the 1990s, DTC's institutional operating budgets reflected the downward profile experienced by most activities within the Department of Defense (DoD) following the First Gulf War. Submissions for the Army's Program Objective Memorandums (POMs) reflected DTC's concern that, as stewards of the bulk of the Army's MRTFB activities, it would violate the intent of *DoD Directive 3200-11, Major Range and Test Facility Base* that specifies activities "shall be sized, operated, and maintained primarily for DoD T&E (Test and Evaluation) support missions." Specifically, institutional funding was insufficient to adequately support the infrastructure and operating costs required to provide the testing services our DoD customers required. To compensate for this funding shortfall, DTC ranges were forced to pass on the critical portion of this institutional funding shortfall to those customers. At the time, DTC also was "making ends meet" with the help of our customers; Program and Product Managers (PMs) were exceptionally forthcoming in making capital investments in DTC range instrumentation and facilities in cases where required capabilities did not exist and DTC was unable to fund these investments.

Public Law 107-314. In December 2002, Congress enacted the *Bob Stump National Defense Authorization Act for Fiscal Year 2003 (2003 NDAA)*. Among its provisions, the Act sought to ensure the "institutional funding of test and evaluation facilities" by FY06. The provisions of the Act directed that

- Both institutional and overhead costs of facilities or resources within the Major Range and Test Facility Base shall be fully funded through the major investment accounts of the military departments, the DoD Central Test and Evaluation Investment Program (CTEIP) account, and other appropriate accounts, and
- Charges to DoD users of MRTFB activities are limited to not more than the direct costs of use.

The provisions of the Act and the details of its implementation were the subject of much discussion in the Army Test and Evaluation community both from the standpoint of legality—what practices were and were not permitted—as well as fiscal policy and funding. In particular, DTC and its seven MRTFB activities were deeply involved in the discussions and decisions given the huge potential the law had for impacting test operations at the affected sites. As policy emerged, it became clear that the former practices of passing on institutional costs to customers as well as garnering capital investments from the customer base were no longer viable means of replacing institutional shortfalls. Foremost, a clear understanding throughout the command of what costs appropriately could be passed on to a customer, as well as what investments by a customer were legal, was needed. To answer these questions, ATEC obtained the assistance of the Deputy Assistant Secretary of the Army for Cost and Economics (DASA-CE) in developing common definitions for direct versus indirect costs to ensure compliance and consistency. These definitions became of particular importance not only in considering appropriate customer charges, but also in addressing the issue of support to multiple customers as it relates to investments on DTC's ranges.

DTC's FY06-11 POM submission reflected those unfunded requirements essential for compliance with the 2003 NDAA. Recognizing the Army's responsibility to meet the provisions of the law, the Assistant Secretary of the Army for Acquisition, Logistics, and Technology (ASA[ALT]) responded by realigning internal funds from the Army program executive officer or project manager organizations into the DTC operating budget. In addition, ASA(ALT) served as the Army advocate for realigning funds from other DoD agencies that use DTC's facilities. Beginning, as required, in FY06, DTC began operating under the requirements of the 2003 NDAA.

Post-2003 NDAA Operations. One of the most marked changes after the law took effect was the receipt of increased institutional funding. This increased funding provided DTC numerous advantages. Among these was greater flexibility, and as a result, increased economy in letting contracts, because funding now was available "up front." It also permitted the routine maintenance that in the past had often been neglected because necessary funding first had to be obtained by reimbursement. Another advantage was that it provided DTC's ranges with a remedy for an issue often viewed as unfair by test customers, but that was forced on us by reimbursable funding practices—specifically, if a major item of test instrumentation broke down during a customer test, the costs of repair or replacement historically had been passed on to that customer. Yet another positive change affected by the law's introduction was a decline in test costs; DTC's customers have, on average, paid markedly less for services since the beginning of FY06 when the law went into effect.

While DTC's experiences operating under the 2003 NDAA for the past two full fiscal years have generally been positive, two key challenges that affect test operations have emerged. The two challenges are: (1) if workload exceeds institutional and overhead budget guidance, and (2) if customers wish to invest in DTC range instrumentation and facilities.

Budget guidance at the start of each fiscal year fully funds the institutional and overhead costs of test center operations and supports execution of a *specific level* of direct test effort. Issues ensue when that level of effort is exceeded. To fully understand, it must be realized that all direct labor effort carries some indirect components, for example, mandatory training. As such, any increase to the DTC workforce—not uncommon in reacting to the Army's operational needs—results in an unanticipated and, hence, unfunded burden on the institutional/overhead operating accounts. Since, to comply with the law, these overhead labor costs cannot be charged to a DoD test customer, test centers must either delay test needs to the start of the next fiscal year or identify some other budget requirement that can be reduced or deferred to absorb the unanticipated institutional labor bill. Because few critical test needs can be delayed when supporting an Army at war, we have been forced to defer the much needed technology improvements for our ranges until later. If this trend continues, our test centers will become hollow—less capable of providing state-of-the-art and timely test services. In short, no mechanism exists within the process to recoup year-of-execution increases to overhead costs *that cannot be anticipated or passed on to the customer.*

As previously mentioned, prior to the 2003 NDAA, test customers made investments to DTC range instrumentation and facilities when needed capabilities did not exist. Such investments currently are not permitted under the law if the resulting capability creates for the test center an asset that subsequently may be used to support multiple test customers. While not unknown, it is, however, rare that a capability/facility would be used exclusively to support a single test customer. In addition, should the investment prove to be acceptable (i.e., supports only a single customer); DTC then must consider the downstream costs of accepting that capability into the test range inventory. Since, for the most part, it will require out-year maintenance, sustainment, and revitalization costs; it becomes yet another asset competing for scarce institutional funding.

As of this writing, the latter concern soon may be remedied. Language has been included in the current draft of the *DoD Financial Management Regulation, Volume IIA, Chapter 12, Major Range and Test Facilities*, to allow

> *"by mutual agreement, investments in new or existing T&E facilities … in whole or in part, by one or more DoD customers of an MRTFB … ." Such agreements will "… delineate responsibilities for funding, staffing, operating, and maintaining the facility and must be approved by all parties … ."*

As we proceed through the current decade and into the next, anticipated funding profiles for DTC operating and investment budgets reflect a pattern similar to that of the 1990s. Heeding the historical profile, this alerts DTC to the challenges in complying with the provisions of the 2003 NDAA, while steadfastly maintaining its mission of providing world class testing capabilities for our DoD test customers. ❑

James B. Johnson was appointed to the Senior Executive Service and assumed his present position as executive director at the U.S. Army Developmental Test Command, Aberdeen Proving Ground, Maryland in 2007. He has management responsibility for the Command's test and technology mission and all associated resources. He is responsible for planning, executing, and reporting 1,700 tests supporting more than 300 weapons programs annually, with a total budget of $2 billion and a workforce of more than 8,000 employees and for ensuring operational readiness of the Army's developmental test range infrastructure. Prior to this position, Johnson served as director of the U.S. Army Redstone Technical Test Center (RTTC), Developmental Test Command.

He previously served multiple assignments in the Missile Defense Agency (MDA), Ground-Based Midcourse Defense Joint Program Office in Huntsville, Alabama and was director of Test Operations; deputy product manager of the Test, Training and Exercise Capability (TTEC) Product Office; and chief of the Test Products Division in TTEC.

Prior to his MDA service, Johnson served at RTTC as team leader of the Radar Systems Group. He began his government career with the Aviation and Missile Research, Development and Engineering Center (AMRDEC) located at Redstone Arsenal, Alabama.

Johnson graduated with a bachelor's degree in electrical engineering from the University of Alabama in Huntsville (UAH) and a master's degree in systems engineering, also from UAH. He also holds a master of strategic studies degree from the Army War College. Mr. Johnson is a graduate of the Army Management Staff College and the Advanced Program Management Course. Johnson's awards include the Army Achievement Medal for Civilian Service, the Edward H. Gamble Award, the TECOM Professional Certificate, the AMRDEC Extraordinary Performance Award, and a U.S. patent. E-mail: James.B.Johnson@us.army.mil

Endnotes

[1]Of the Army's nine Major Range and Test Facilities Base (MRTFB) activities, seven are Developmental Test Command (DTC) components: White Sands Test Center; Yuma Test Center; Cold Regions Test Center; Tropic Regions Test Center; West Desert Test Center; Aberdeen Test Center; and the Electronic Proving Ground.

ITEA Journal 2008; 29: 7–8

A Dramatic Approach to High-Power Fiber Lasers

Burke Nelson Ph.D. and Sami Shakir Ph.D.
Northrop Grumman Information Technology, Albuquerque, New Mexico

The expansion of the market share of fiber lasers for industrial applications is fueled by some of the significant advantages of fiber lasers. These advantages include high efficiency, compactness, good beam quality, and low cost of maintenance and operation. For military applications, in which power levels of 100 kW or more are necessary, fiber lasers are beginning to show promise as credible candidates. From a military standpoint, lasers form a class of weapons called Directed Energy Weapons. Laser beams travel at the speed of light, which makes their target effects instantaneous. The power level attainable with fiber lasers is continually increasing. We have a current contract with the U.S. Navy where we will passively phase two 1-kW Nufern amplifiers. We will be ready for demonstrations on a test range within a few years. We do not foresee any insurmountable barriers to high power, the time to a prototype demonstration being primarily dependent on available funding.

What's a fiber laser?

A fiber laser utilizes very thin glass fibers doped with special materials to convert poor-quality laser light generated by diode lasers to a high-quality laser light beam. The fiber waveguide, which is composed of a glass core region doped with a rare-earth material such as ytterbium (Yb) surrounded by a regular glass cladding region, confines the laser beam to the fiber core region where laser gain and amplification take place *(Figure 1)*. Depending on the type of doping material used, the laser operating wavelength can range from 0.8 µm to 2.3 µm.

Why fiber lasers?

Fiber lasers are driven by electrical power. Modern aircraft, naval vessels, and ground combat vehicles incorporate significant amounts of electrical power. Using that electrical power, the only logistical supply required is fuel. While there are other electrically driven lasers, fiber lasers have significantly higher efficiencies than bulk solid-state lasers. Fiber lasers offer better waste-heat management and are relatively immune to the deleterious effect of heat on the beam quality of the laser. (Beam quality is a measure of how well a laser beam can be focused at a target.) In contrast, the major impediment to bulk solid-state lasers is the sensitivity of the bulk laser materials to temperature gradients, which causes the beam quality to deteriorate. Fiber lasers also have the advantage that the system can be monolithic in the sense that the laser beam is confined within the flexible fibers of the system and requires no alignment or free-space bulky optics such as lenses and mirrors. This is a significant advantage in harsh military environments

Passive phasing's demonstrated success

Passive phasing is a coherent phasing process in which an array of high-power fiber amplifiers is locked and phased automatically by the system itself. A small fraction of the output beams of the array are sampled and fed back into a single-mode feedback fiber in a ring configuration,

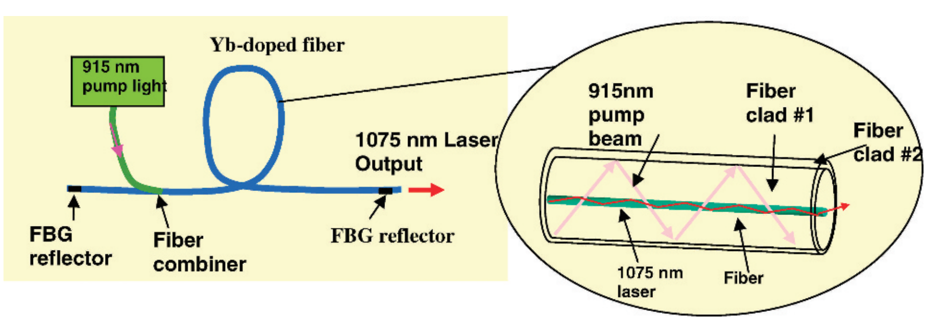


Figure 1. Fiber laser. The fiber core that guides the laser beam (red) is pumped by diode-pump beams (pink). A portion of the laser beam is reflected back by a reflector (Fiber Bragg Reflector, FBG) to form a fiber laser. The doped core forms the gain medium, while the fiber and the end reflectors (FBG) form the resonator. Unlike conventional lasers, the fiber and end reflectors form an all-fiber monolithic laser, which is a significant advantage for fiber lasers.

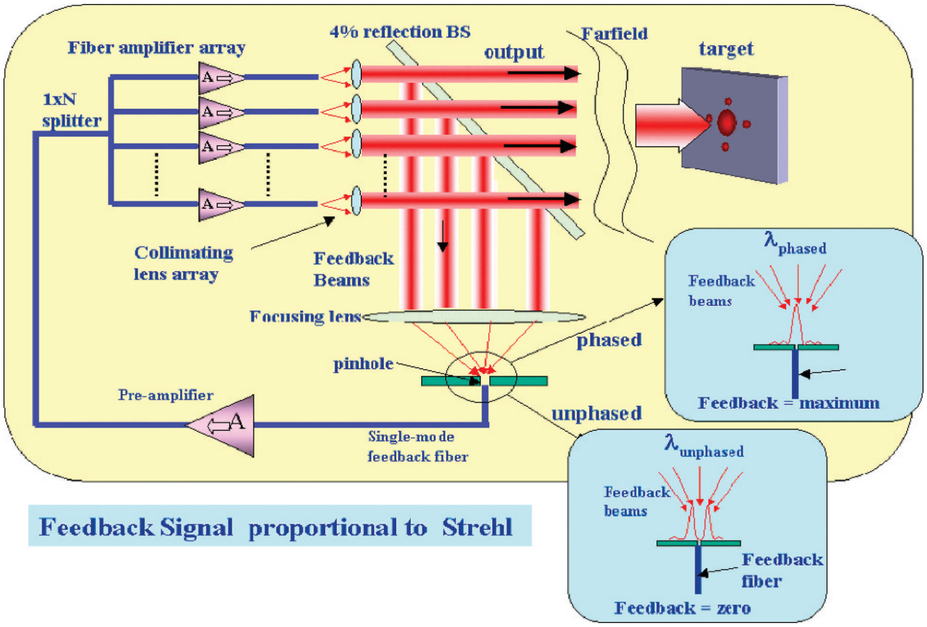


Figure 2. The patented passive feedback drives the output to the highest intensity level

as shown in *Figure 2*. This concept is covered by U.S. Patent 7,130,113. The feedback signal is split equally, and each portion serves as the input signal for one of the amplifiers in the system. Since the system wavelength is not fixed, the system runs at a wavelength that has the highest feedback signal. By design, the feedback signal is highest when the beams are in phase. Therefore, this passive phasing approach locks the fiber amplifiers to the same wavelengths and also causes the output beams to have the same phase. Frequency locking and phasing are necessary requirements for effective coherent combining of laser beams.

The effect of passive phasing on the performance of the system is shown in a dramatic way in *Figure 3*, which compares the intensity profile when the feedback loop is blocked (i.e., no passive phasing) to the case when it is restored.

Having successfully phased 16 fiber lasers, and with our near-term plans to phase two 1-kW amplifiers, it is time to plan prototype demonstrations of high-power, high-intensity lasers at a range test facility. ❑

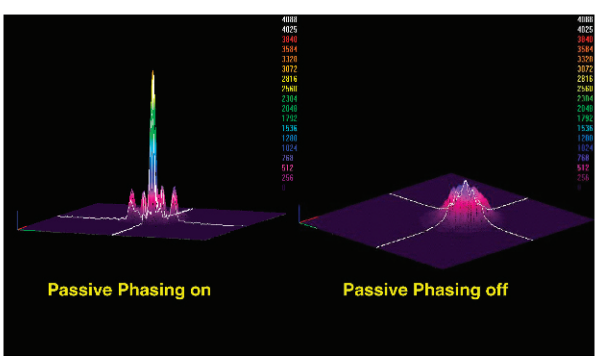


Figure 3. An array of 16 fibers demonstrates the dramatic increase in central intensity provided by passive phasing

Burke Nelson is the deputy program manager of the Airborne Laser Advisory and Assistance Services contract. He is also program manager of a U.S. Navy contract, jointly with Northrop Grumman Electronic Systems, to develop fiber lasers for weapon system applications. His areas of expertise include chemical and solid-state lasers, as well as large-optics applications. Dr. Nelson has been with Northrop Grumman in Albuquerque, New Mexico, for over 22 years. His previous assignments included associate director of Research and Director of Engineering, PerkinElmer, Inc., and executive director of the American Society of Mechanical Engineers (ASME). He was also a congressional fellow for ASME. He has been awarded five U.S. patents, including, most recently, a patent on passive phasing of fiber lasers. He holds a bachelor of science degree in mechanical engineering from Michigan State University, a master of science in engineering degree in aeronautics and astronautics from the University of Washington, and a Ph.D. in materials science from Drexel University. E-mail: burke.nelson@ngc.com

Dr. Sami Shakir is a senior scientist with Northrop Grumman Information Technology (NGIT). He obtained his Ph.D. in optical sciences from the Optical Sciences Center in Tucson, Arizona, in 1980. Before joining Northrop Grumman, he was a professor at the University of New Mexico until 1986, when he joined R&D Associates, which was later acquired by Northrop Grumman. His interests are high-power solid-state and fiber lasers and beam propagation and control. He is the inventor of the Northrop Grumman passive phasing approach. E-mail: sami.shakir@ngc.com

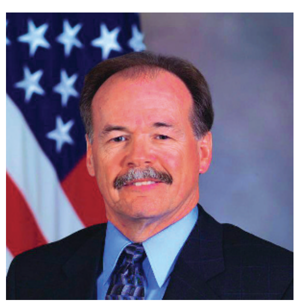
ITEA Journal 2008; 29: 9–10

Assuring the Future—How We Gained Access to Additional Radio Spectrum for Flight Testing

John B. Foulkes, Ph.D.
Test Resource Management Center,
Office of the Under Secretary of Defense, (Acquisition, Technology, and Logistics), Arlington, Virginia

In the mid-1980s, spectrum used by the Department of Defense (DoD) test community began to be reallocated for the rapidly growing consumer electronics market. By the mid-1990s, the DoD had lost access to 30 percent of the spectrum used to carry data during testing of aerospace vehicles. At the same time, advances in computer technology and onboard sensor electronics resulted in an exponentially increasing demand for additional spectrum to provide useful test results in real time during test events. The DoD joined with the National Aeronautics and Space Administration (NASA) and the aircraft manufacturers' industry group, the Aerospace and Flight Test Radio Coordinating Council (AFTRCC), to address this shared problem. Operating as the ad hoc Range Spectrum Requirements Working Group (RSRWG), the three partners developed a plan that ultimately led them to the International Telecommunications Union (ITU), an international treaty organization based in Geneva, Switzerland. This challenge required the three partners to use nontraditional approaches to address this issue. This article discusses the working relationships and approaches used to ensure that we successfully addressed the issues of spectrum encroachment.

The Range Spectrum Requirements Working Group (RSRWG) planning process began in 1996 with the development of a three-pronged approach: (a) defend against further losses of spectrum; (b) develop new technologies to more effectively use the spectrum; and (c) devise new approaches and processes, including gaining access to additional spectrum. While most of the plan called for straightforward use of internal Department of Defense (DoD) and interagency processes in Washington, gaining access to additional spectrum presented unique challenges for the test community. This feat would require buy-in of the senior leadership in each of the organizations represented, both government and private industry. The RSRWG spearheaded the effort to implement this plan, which involved foreign governments and ultimately resulted in the consortium's participation in the International Telecommunications Union's (ITU's) international radio frequency spectrum regulatory arm, the World Radiocommunication Conference (WRC). Changes in radio frequency (RF) allocations are tantamount to revisions to an international treaty.



Dr. John B. Foulkes

The RSRWG plan called for a globally harmonized RF band or set of bands to allow interoperability of test assets, reduced equipment cost through commonality, global testing, and increased protection against RF spectrum encroachment. The National Aeronautics and Space Administration (NASA) had earlier submitted a proposal to the ITU to consider "spectrum for wideband telemetry in the 3 to 30 Gigahertz (GHz) region" at some future WRC. Such undertakings invoke all the machinery of state, and the process takes years. The RSRWG partners began by giving as many educational briefings on the requirement for additional spectrum as needed to all of the stakeholders in spectrum management to include the Department of Commerce, the Department of State, and the Federal Communications Commission. Furthermore, the NASA and DoD representatives were responsible for communicating the new requirements to their respective spectrum management offices. For the DoD, that required getting approval from the Director of Spectrum Policy within the Office of the Assistant Secretary of Defense (ASD) for Command, Control Communications, and Intelligence (C3I) (now ASD for Networks and Information

Integration or NII) and the spectrum management agencies of the military services.

The road to WRC-2007

The road to WRC is a very long process. The first step is developing a case to justify placing a proposal on the agenda for a future WRC. A proposal first goes to a WRC as a recommendation that it be placed on the agenda for consideration at the following WRC. Typically the WRC convenes every three to four years, and changes are agreed to only if a consensus is reached among the 191 member nations. Therefore, it was necessary to ensure that a sufficient number of nations would support the U.S. telemetry agenda item to ensure consensus. Accordingly, the RSRWG had to develop a plan that allowed it to educate foreign nations. The RSRWG, while representing a powerful U.S.-based coalition, needed to become an international force. As a result, the group teamed with the International Foundation for Telemetering (IFT) to work together to garner the international support necessary to make the essential additional spectrum allocations a reality. Together, the RSRWG and the IFT worked to charter the International Consortium for Telemetry Spectrum (ICTS). The ICTS membership encompasses representatives from most major aircraft manufacturers and military and flight test establishments throughout the world. The goal of the ICTS was to facilitate the development of a set of internationally agreed-upon technical recommendations and implementation alternatives. The sharing of information within the ICTS provided the foundation for the ICTS members to convince their national spectrum managers that the telemetry spectrum proposal for WRC was important to their nation's interest. As a result of RSRWG and ICTS efforts, the United States succeeded in getting the telemetry spectrum proposal approved at WRC-2003 as Agenda Item 1.5 for WRC-2007.

A grassroots effort was required to communicate the details of Agenda Item 1.5 directly to as many member nations as possible prior to the WRC. This was accomplished by informational briefings at regional level forums. The ITU has divided the world into three regions. Groups of nations within these regions have formed regional organizations. These regional organizations have official standing within the ITU and submit a single consolidated set of positions for their respective organizations. Building support within these regional organizations was one of the keys to success at WRC-2007.

RSRWG members knew that even with extensive grassroots efforts, more work was required to ensure the success of Agenda Item 1.5. Approximately 3,000 delegates from more than 150 countries attend the WRC. Many of these delegates arrive with little knowledge about the agenda items of other delegates. To secure the highest probability of success for Agenda Item 1.5, the ICTS developed an information booth to educate delegates on the initiative.

Because ITU is an international treaty organization, each country's delegation at WRC is led by an ambassador. Another key aspect of the RSRWG's outreach efforts was to brief the leader of the U.S. delegation. A few weeks before the WRC, representatives of the U.S. DoD test ranges led by the TRMC, along with representatives of the commercial aircraft manufacturers, met with the U.S. WRC Ambassador in Washington, D.C., to brief him on WRC Agenda Item 1.5. The Ambassador immediately grasped the significance of the item and remained an effective advocate throughout the duration of the WRC.

WRC-2007—October 22 to November 16, 2007

The years of preparation by the RSRWG partners came to fruition in the four-week period beginning on October 22, 2007, in Geneva. Three representatives of the partnership were members of the U.S. delegation. ICTS colleagues from Germany and France were members of their respective nation's delegations.

Agenda Item 1.5 took 23 of the 26 days of WRC-2007 to make it through the process. In the end, the international telemetering community gained access to substantial amounts of additional bandwidth, including the first globally harmonized band. The band of 5091–5150 MHz is now authorized for aeronautical telemetry in every country in the world. Many regions of the world have access to substantially more bandwidth in addition to the global band. The United States and Canada have the ability to access up to 1.4 GHz of additional spectrum for telemetering applications. Although this may seem like an overly generous amount, a majority of this spectrum is already in use by incumbent users. However, the RF bands approved by the WRC will provide all users with greater flexibility to work within these allocations together with minimal impact on each other.

The success of WRC-2007 Agenda Item 1.5 is clearly due to the strong partnership that led to effective channels of communications and the willingness of senior leadership across government and industry to make a long-term commitment to pursuing a common goal. The 10-year quest for telemetry spectrum ended in a resounding success. ❑

JOHN B. FOULKES is director, Test Resource Management Center, Office of the Under Secretary of Defense (Acquisition, Technology and Logistics), Arlington, Virginia.

ITEA Journal 2008; 29: 14–16

The Teardrop That Fell From the Sky: Paul Jaray and Automotive Aerodynamics

Prof. Guillaume de Syon
Albright College, Reading, Pennsylvania

"What's the difference between an airship and the Chrysler Airflow?" This question is taken neither from a quiz show, nor the opening line of a joke. The answer, according to engineer Paul Jaray, would be "not a lot, aerodynamically." Paul Jaray's name still appears in many popular works on automobile design yet is often ignored otherwise. Jaray, however, represents a successful cross-over not just in the realm of engineering (from airplanes to ground vehicles), but also in that of aesthetics as well.

Jaray's background (he was born in 1889, one of five children in a Viennese Jewish family) did not predispose him to scientific study. His father was a salesman, but Paul's artistic and mechanical inclinations, as well as the dynamism of Central European higher education before World War I, enabled integrated Jewish families like his to send their children into technical pursuits and other professions. Jaray trained as an engineer in Vienna and chose aerodynamics as his specialty. His first contact with aviation came in 1909, when he witnessed a Blériot flight.

By 1912, the anticipation of war in Europe prompted the fledgling aviation industry to hire new hands. The outbreak of fighting found Jaray employed by the Zeppelin Works in Friedrichshafen, Germany.

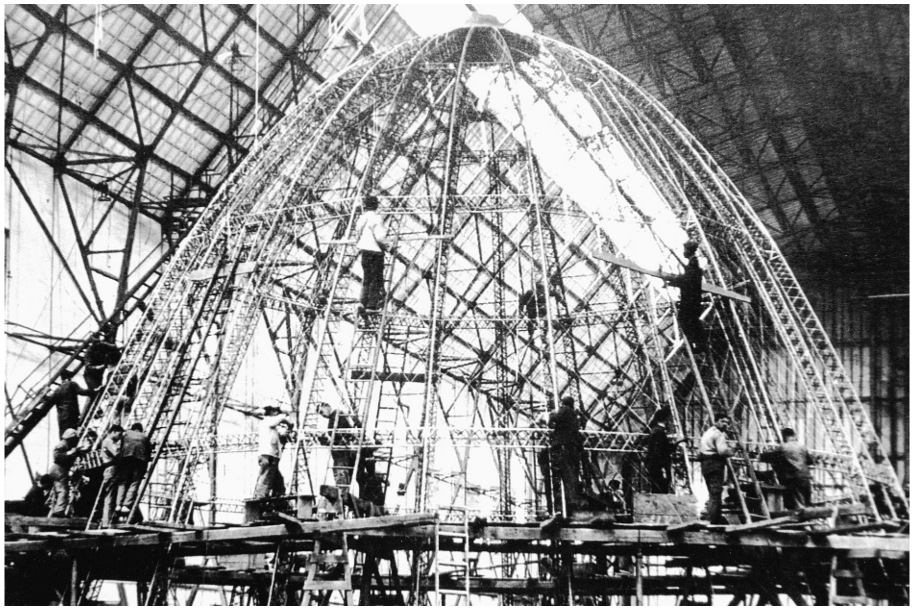
Jaray's aerodynamic insights began to assert themselves in World War I. Unfortunately, the Zeppelin Company—where the paternalism and conservatism of its founder was felt in all design realms—did not welcome innovation. Some of Zeppelin's lieutenants did display business sense and even social conscience (by providing subsidized housing for workers, for example), but the brilliant technical minds at Zeppelin, like Claude Dornier, chafed at the constraints. Not that opportunities did not present themselves. War conditions allowed for the transfer of airship patents from Schütte-Lanz, a Zeppelin competitor, to Zeppelin as a means of improving the quality of airships delivered to the German navy. Nonetheless, the shape of airships remained faithful to the pencil-like design, thwarting improvements in speed.

Undeterred, Jaray began studying the best airship form based on a combination of diameter, cross-section, volume, and stress points (Pfeiffer 1935). The systematic study of shapes led him to conclude that teardrop cross-sections (in which increases in capacity were achieved through thickening fuselages rather than lengthening them) offered the best solution. Pencil- and tube-shapes had to go. It would be several years before this concept was accepted.

Presenting the calculations was one thing, but only actual flight tests could confirm Jaray's assertions. The early results on a 1915 airship model proved disappointing. They yielded a speed lower than projected, and Jaray was ridiculed by pilots for his miscalculations. But after reexamining his work, Jaray realized that the problem was simple: his instructions had not been followed. The propellers, which needed to be changed, were the same as those from the earlier airships. Insisting on a redesign and supervising it himself, Jaray convinced his superiors to retest. His calculations proved to be off again, but this time in his favor. The speed of the airship exceeded 95 miles per hour, better than Jaray's predictions. This delayed success helps explain why Jaray did not get the wind tunnel he had been demanding until 1916, which became essential to the design of two subsequent Zeppelin aircraft built for the military (Kleinheins 1994). By 1917, Jaray had been promoted to supervising engineer, and he oversaw several improved designs in the naval airships completed by the end of the war, *Figures 1 and 2*. He was also instrumental in the design of postwar transport airships, and his work there influenced the shape of all machines including the Hindenburg long after he had left aeronautical design.

With the end of World War I, Jaray turned his attention to streamlining ground vehicles. Many aviation companies like Zeppelin, no longer sustained by military contracts, pursued auto making and industrial products after the conflict. Until this time, car manufacturing (with the notable exception of the Ford Model T) had involved quasi-artistic craftsman-

Figure 1. LZ120, one of Jaray's projects, under construction in Friedrichshafen, Germany (Astra, reproduced with permission)



ship (based on individual orders for a chassis, a body, and so on). But new materials—sheet steel and aluminum—changed the equation of auto manufacturing and design. Jaray wedded these materials to his aeronautical concepts starting in 1919 and introduced a streamlined car body in 1921. Jaray also took out a patent on his own streamlined vehicle, although he remained on the Zeppelin payroll.

Jaray had little success with his initial efforts because any aerodynamic solution, as he discovered himself, could not simply rely on redesigning the car's body. Aesthetics also mattered, as did comfort and practicality. His first model—constructed in 1921 but identified as a 1922 Ley T-6 by Zeppelin—looked ridiculous because of its high stance and narrow cross-section. Still, compared with a standard Ley model, it was more fuel efficient and could negotiate climbs better (Curcio 2000). Jaray had taken as a departure point a split airship cross-section, adapted to the requirements of the automobile. But there were serious complications. The chassis had proven a problem during the whole design phase. It was not streamlined, and considerable modifications were necessary to achieve a proper flow between the vehicle and the ground. That said, wind tunnel testing yielded an impressive drag coefficient of .28.

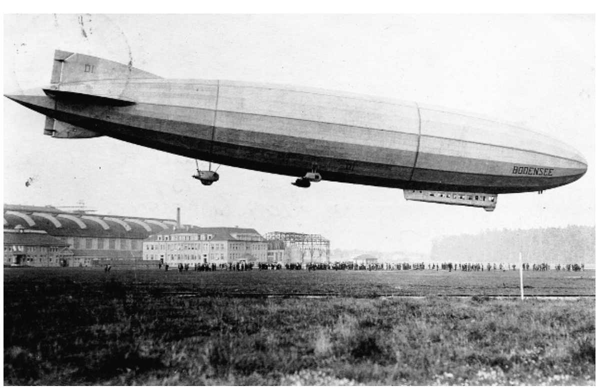


Figure 2. A 1920 postcard showing one of Jaray's last Zeppelin projects, the LZ120 transport: the teardrop-shape is emblematic of Jaray's work (Astra, reproduced with permission)

Jaray finally left Zeppelin in 1923, and although he was out of a job, he was not out of ideas. He quickly formed a company in Switzerland (where he moved) to promote his automotive designs. He also set about obtaining a patent for the most promising market, the United States. The timing seemed right. The streamline shapes derived from wartime aerodynamics had by then jumped into the realm of culture, resulting in designs for commercial goods that combined the quest for speed with an aesthetic that incorporated functionality. Popular culture had assimilated the teardrop shape advanced by Jaray. Indeed, the interwar years saw an unimaginable number of attempts at incorporating the teardrop into designs ranging from cars to kitchen utensils. In the United States, as historians of the "Airstream" trailer have noted, such eminent minds as Raymond Loewy and Glenn Curtiss believed, somewhat naively, that the teardrop held the solution to all problems of drag (Burckhart and Hunt 2000).

Despite such popularity, the only firm willing to transform Jaray's ideas into production automobiles was the Czechoslovak Republic's Tatra Company. (Others had built test vehicles but never brought them to market.) In 1934, Tatra introduced a V-8 model that included pontoon panels, integrated headlamps, and a tapered rear end. Its drag coefficient (.36) made it one of the most aerodynamic mass-produced cars ever conceived (de Noblet 1993). Yet, the combination of political and economic turmoil of the 1930s, as well as the Tatra's cost, made it a short-lived experiment. Still, it inspired others to emulate it.

As for the American audience, although Jaray had established a company to market his hard-earned 1927 patent on aerodynamic streamlining (it took almost five years to obtain), he was dismayed to discover that Chrysler brought out its Airflow model in 1934 without any acknowledgment of his work. By then, Jaray was far from alone in advocating such approaches to automobile design. Norman Bell Gedes, for one, was particularly famous for his own models. But Jaray's work was among the best known in automotive circles and publications. His lawsuit in 1935 against Chrysler was momentous, but not very profitable. Chrysler agreed to pay Jaray's company just $5,000 in damages.

In Germany, Austrian-born Ferdinand Porsche also realized when examining pictures of the Chrysler Airflow, the Peugeot 402, and of course the Tatra that Jaray was on the right track. What was missing, however, was affordability and practicality. The Devil's pact that Porsche made with Adolf Hitler resulted in the Volkswagen Bug, a car familiar now, but one that looked radically different from typical small cars of that era.

Jaray led a quiet life after World War II, focusing on consulting for the automobile industry. He died in 1974. As one automobile historian summarized his legacy, it was not that he was responsible for specific details on aerodynamics and streamlining; many others, like him, offered these. Jaray's main contribution was his insistence that all production automobiles incorporate aerodynamics into their designs *(Figure 3)* (Sloniger 1975). His work also represents a symbolic link, not just between aeronautics and ground vehicles, but also between formal engineering practices and culturally oriented design. ❑

Guillaume de Syon teaches history at Albright College in Reading, Pennsylvania. He is the author of Zeppelin! Germany and the Airship, 1900–1939, recently issued in paperback by the Johns Hopkins University Press. E-mail: gdesyon@fandm.edu

Figure 3. Jaray's automobile for one, as demonstrated in Switzerland (Corbis, reproduced with permission)

References

Burckhart, B. and Hunt, D. 2000. *Airstream: The History of the Land Air Yacht.* Chronicle Books: San Francisco. p. 60.

Curcio, V. 2000. *Chrysler: The Life and Times of an Automotive Genius.* Oxford University Press: New York. p. 527.

Juchet Gaston. 1993. "Car Design." *Industrial Design*. Edited by Jocelyn de Noblet.

Kleinheins, P. 1994. *LZ 120 "Bodensee" and LZ 121 "Nordstern".* Zeppelin Museum: Friedrichshafen. pp. 26–27.

Pfeiffer, E. A. 1935. *Fahren und Fliegen. Ein Buch für alle von Auto, Flugzeug, Zeppelin.* Franckh'sche Verlagshandlung: Stuttgart. pp. 82–83.

Sloniger, J. 1975. "The Slippery Shapes of Paul Jaray." *Automobile Quarterly*. XIII. p. 3.

ITEA Best Paper
ITEA Journal 2008; 29: 17–21

P-8A ''Poseidon'' Collaborative Simulation and Stimulation for Electromagnetic Environmental Effects Test & Evaluation

Paul Achtellik
Naval Air Systems Command,
Integrated Battlespace Simulation and Test Department,
Integrated Combat Environments Division,
Electromagnetic Compatibility Branch, Patuxent River, Maryland

Over the past several decades, technological advances have provided the Naval Air Systems Command (NAVAIR) with exciting opportunities while creating significant challenges to those who design, test, and operate the complex mission systems found on today's war-fighting aircraft. The responses to these challenges are well underway and began with innovative planning and cost-wise construction of various next-generation facilities, conceptual planning of integrated and extensible network infrastructures, and the insistence on collaborative engineering across all phases of the acquisition life cycle. Today, the challenge continues and, in many aspects, has become even more difficult, stretching our fiscal, technological, and personnel resources to their limits. This article addresses one of the more difficult aspects of today's challenges: Conducting Ground-Based Full Spectrum Test & Evaluation on Next-Generation Systems.

Key words: advanced test facilities; complex operational systems; electromagnetic compatibility; electromagnetic environmental effects (E^3); network-centric warfare; realistic mission environments; simulator/stimulator testing labs.

The Naval Air Systems Command (NAVAIR) has many robust, state-of-the-art test and evaluation (T&E) facilities that evaluate entire systems before significant decisions are made to deliver some of the world's most advanced weapons systems into the hands of our sailors and marines. Advanced Installed Systems Test Facilities, managed and operated by the Integrated Battlespace Simulation and Test (IBST) Department, provide realistic ground-based test environments during various phases of systems development to identify and reduce risks prior to more costly and rigorous flight-test phases. A multitude of potential risks associated with overall system performance, personnel safety, and intra-system electromagnetic compatibility are identified during all phases of system development in a scientifically-controlled environment through the use of advanced simulation and stimulation techniques. Test results provide critical data to developers and program managers well before important program milestone decisions, and provide insight into how our next-generation systems will function in joint and coalition mission threads and future battle space environments. Facilities, such as the Air Combat Environment Test and Evaluation Facility (ACETEF), the Surface/Aviation Interoperability Laboratory (SAIL), the Integrated Battlespace Arena, and a variety of advanced electromagnetic environmental effects (E^3) facilities were purposely designed to facilitate the immersion of installed systems in an environment that can repeatedly replicate realistic mission environments and provide detailed data to evaluate potential system effectiveness during actual missions. Simulators and stimulators are designed to provide realistic Electronic Warfare (EW) threat environments, authentic Global Positioning System (GPS) satellite signals, friendly and hostile communications and data link signals, and accurate electromagnetic environmental effects in a

This article was selected as the ITEA Best Paper—First Place at the Annual ITEA International Symposium, November 12–15, 2007, Lihue, Hawaii.

scripted, realistic and cohesive test event that replicates any level of detail desired. Central to these test events are models, such as the Joint Integrated Mission Model (JIMM) and the Next Generation Threat System, that "set the stage" and drive computer simulations and facility stimulator hardware, sharing necessary data through standardized interfaces. From single-threaded, focused test vignettes to fully integrated wartime scenarios, the laboratories and facilities in NAVAIR are capable of emulating a wide range of realistic environments in a live, virtual, or constructive manner. As Major Range Test Facility Base (MRTFB) unique national assets, these advanced capabilities support NAVAIR and U.S. Navy testing, but are also available to support all joint-service programs.

A central component of advanced ground test capability, the Advanced Systems Integration Laboratory (ASIL) is a radio frequency (RF) - shielded anechoic chamber measuring 180′ × 180′ × 60′ with over one-hundred-thousand square feet of RF-absorbing material. This chamber provides "the stage" for some of the most advanced test laboratories, distributed simulation and stimulation hardware and software, and fully integrated aircraft and facility instrumentation components. The resulting simulated environment is capable of providing test articles with virtual, scripted mission scenarios that provide flight-like realism to test the complex suite of communication, navigation, identification, and mission systems.

As the complexity of tomorrow's systems increases, so does the requirement for research, development, and test and evaluation facilities to provide matching levels of complexity to produce realistic testing environments. Advanced weapons systems, such as the P-8A "Poseidon" and the F-35 "Lightning II" boast unparalleled intra-system workings and will demand integrated testing methodologies never before imagined. To illustrate the challenges and their potential solutions, we look at the early stages of test planning for the P-8A "Poseidon," focusing on the new complexity required for what was once straightforward E^3 testing.

Advanced electromagnetic compatibility testing for next-generation multi-mission maritime aircraft

The P-8A "Poseidon" Multi-Mission Maritime Aircraft will become the newest addition to the U.S. Navy's airborne surveillance and reconnaissance arsenal, bringing unparalleled capabilities and complexities to the future of naval aviation (*Figure 1*). A cornerstone of the Navy's ongoing transformation in naval war-fighting doctrine, the P-8A brings forward-looking operational concepts of jointness, interoperability, and full-spectrum dominance of sea-, air-, space-, and information-domains to its primary mission.

Figure 1. P-8A "Poseidon"

Keys to achieving full spectrum dominance are information superiority and operations, through the application of network-centric warfare. Information, information-processing, and communications networks provide the core of every military activity, and sharing this information seamlessly through robust communication networks that provide common operational and tactical pictures to naval commanders is crucial to the Navy's effectiveness in supporting national interests. The P-8A will be a major airborne asset providing intelligence, surveillance, and reconnaissance information; information processing; and communications in network-centric warfare.

Testing advanced systems

The challenges of testing such a complex collection of systems and subsystems are daunting, considering the interdependencies and interrelationships of each of the aircraft's mission systems. These challenges are combined with rigorous intrasystem electromagnetic compatibility (EMC) compliance requirements (Military Standard MIL-STD-464A 2002) and will demand a great deal of collaboration and coordination among and across organizational boundaries, facilities, and test phases. This level of integrated testing is the reason NAVAIR needs such advanced T&E facilities, and while the facility's architecture can provide critical tools, the collaboration of the facility's workforce becomes equally critical to meaningful testing.

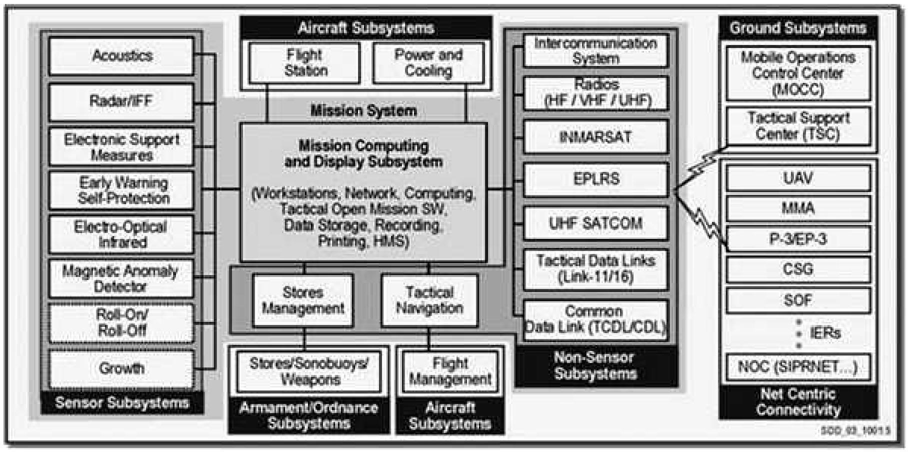


Figure 2. Mission computing and display system

The P-8A's operational environment will be a complex and adaptive blend of sensors, shooters, Command and Control assets and data links; in essence, a collection of nested systems and subsystems operating in unison. To properly test the effectiveness of such advanced weapons systems, the entire aircraft must be stimulated as it would be in an actual mission environment. Stimulating only a few mission systems leaves the remainder of the aircraft's integrated systems in a static state and represents unrealistic mission profiles. Stimulating only a portion of the mission systems also allows little chance of identifying adverse electromagnetic interactions.

As an example, the Mission Computing and Display System (MCDS) (*Figure 2*) requires a blended GPS and Air Data Inertial Reference Unit's (ADIRU) input for proper operation. The GPS/ADIRU can be energized, but without stimulating these systems with valid signals and data, the GPS will "search the skies" and be unable to calculate a position. The GPS receivers must have valid satellite and positional data that agrees with the latitude and longitude entered into the ADIRU; anything less will result in immediate ambiguities within the overall P-8A navigation system with unforeseeable complications for the MCDS and mission systems.

Various stimulators and simulators are required to exercise systems like those on the P-8A. In facilities like the ACETEF, many advanced electronic combat stimulation capabilities are co-located with the chambers and test assets, while others can be remotely networked to support testing. For example, the SAIL has remote connections via fiber optics to provide acoustic and RF ship data links to aircraft under test. Other Joint Service capabilities can be linked and utilized as needed. The table below lists some of the facility's current capabilities (*Table 1*).

IBST simulators, stimulators, and laboratories are integrated into a single virtual dynamic environment using JIMM. JIMM becomes the executive run-time controller for the integrated assets and provides controlled parallel simulation events, using advanced multi-threading processes, to maintain a fully repeatable ordering of events to all interfaced stimulators. The aircraft data bus is instrumented and the data flow

Table 1. IBST simulation/stimulation laboratories supporting the P-8A test

Simulator/stimulator/lab	Purpose
Automated Identification Friend or Foe (IFF) Test Set	Simulates the SIF modes 1, 2, 3, C, and 4; two operating modes; interrogation mode; and transponder mode.
Multiple Link Test and Training Tool	Full network simulation of Link 11 and 16 data links has the capability to simulate any combination of tactical digital information links simultaneously.
Strategic Data Link System (SDLS)	A multi-channel UHF satellite communications (SATCOM)/line-of-sight radio system.
GPS Test Equipment (GPS/SPIRENT)	Simulates a constellation of up to 12 satellites in both L1 and L2; the system under test can be placed anywhere and at any time.
Advanced Multiple Environment Simulator (AMES) III	A dynamic RF threat simulator capable of generating complex radar threat environments.
Infrared Sensor Stimulator	Designed to support the design, development, integration, and testing of infrared electro-optical sensor systems.
Joint Communications Simulator (JCS)	Produces motion, range, and direction of arrival for hundreds of independent high fidelity CNI emitters.
Surface/Aviation Interoperability Lab (SAIL)	Provides tactical common data link and multiple sonobuoy signals.

is time-tagged and captured to provide before-and-after comparison of data processed by the P-8A. This continual real-time feedback allows for detailed post-test analysis of obvious and not-so-obvious adverse intrasystem EMC interactions. In this manner, an intrasystem EMC test of the P-8A can be efficiently conducted while the mission systems and subsystems are artificially immersed in "virtual flight" with relevance to anticipated operational missions.

In order to achieve flight-like realism and mission relevance, JIMM is programmed to run pre-scripted warfare "scenarios" which, for the purpose of this article, refers to the textual depiction of P-8A crew actions, system functions, external activities or stimuli, and all preconditions in the course of accomplishing a whole or partial mission. Scenarios are based on actual Operational Situations and Tactical Situations (TACSIT) as defined in the P-8A Scenario Development Strategy, 2006 (Scenario Development Strategy 13126/A1J1B/PMA-290/SE/1053 2006); the same Operational Situations and TACSITs used for systems integration and crew training in the P-8A's Systems Integration Laboratory (SIL). In practice, the missions conducted inside an actual P-8A aircraft in the ASIL will mimic previous missions that have been rehearsed in the P-8A SIL.

To illustrate how a portion of the intrasystem EMC testing will be performed in relation to these scenarios, TACSIT 5-4, a hypothetical search and rescue mission will be utilized. But, before this search and rescue mission is conducted, EMC engineers will create an appropriate "communications plan" within TACSIT 5-4 to satisfy one of the more critical facets of these tests; to evaluate RF interference between all P-8A transmitters and receivers. EMC engineers use a standalone internally developed software tool called Prediction of Intra-system EMC to help predict where RF interference will be at its worst. This is a mathematical analysis and prediction program that is used in advance of testing to predetermine most likely RF interference combinations.

The Prediction of Intra-system EMC program makes the assumption that all receivers and transmitters are potential victims and sources of interference against one another and properly lists all frequency combinations where interference is likely. These predetermined "worst case" frequency combinations are written into TACSIT 5-4 as part of the detailed communication plan. This mission scenario involves take off, climb out, transit to an operating area, coordination of rescue efforts with Navy surface assets, and electronic surveillance measures to keep track of unfriendly forces. Mission system avionics use involves line-of-sight communications with encryption, various data link operations, identification friend or foe, shipboard automatic information system, geo-locating targets with the electro-optical/infrared turret, inverse synthetic-aperture radar, and electronic surveillance measures. This four-hour mission scenario is flown over hostile littoral waters and concludes with the P-8A returning home safely.

Intrasystem EMC tests in the ASIL will be concentrated on the integrated P-8A mission systems. Since a single source/victim test matrix listing the individual mission system components would be too difficult to manage, tests will be parsed into smaller more manageable matrices using a layered approach to test the whole mission system. Equipment such as line-of-sight communications, satellite communications, identification friend or foe, radar, navigation, sensors, MCDS, weapons systems, etc. will be logically grouped into smaller matrices with a goal of (x) number of victim/source tests per hour or per scenario run. Each scenario-driven test event is intended to allow for a manageable, but thorough evaluation of a small number of systems and subsystems rather than risk the potential chaos of doing too much at one time. In this manner, individual system-versus-system will be scrutinized for adverse EMC, while building up to and ultimately achieving 100 percent-versus-100 percent operation of the whole aircraft and mission systems suite. We find it critical that EMC test engineers and scenario developers collaborate continually to ensure mission scenarios match EMC test requirements. For all P-8A tests, attempts will be made to use pre-existing TACSIT scenarios. These scenarios or vignettes can be modified in accordance with the P-8A Scenario Development Strategy 13126/A1J1B/PMA-290/SE/1053 (2006) to satisfy the EMC test requirements.

From an intra-system EMC perspective, all receivers and transmitters can be evaluated in this manner, along with search and rescue mission systems and subsystems. The hypothetical TACSIT 5-4 scenario includes elements critical to the intra-system EMC evaluation which are modifiable, yet can run as many times as necessary until one of the previously mentioned victim/source test matrices is complete. Minor changes to the detailed communication plan in the TACSIT will blend the software tools of the EMC engineering discipline with the modeling and simulation tools of IBST. This allows a thorough EMC evaluation of the integrated P-8A systems and subsystems with relevance to the aircraft's intended mission. Advanced EMC cannot neglect the air/surface integration challenges nor ignore crucial joint interoperability issues. As programs evolve and plan for joint interoperability and net-ready Key Performance Parameter, E^3 and mission system performance testing will evolve as well.

Conclusion

Creating operationally relevant test scenarios in a controlled environment is necessary to accomplish effective and affordable testing on the extremely complex weapons systems of tomorrow. The P-8A may be one of the first "next generation" systems to undergo testing in such an environment, but will be followed by a surge of advanced programs in an increasingly difficult and demanding T&E world. The MRTFB, T&E communities, and NAVAIR have taken proactive steps by creating the framework for full-system collaborative and cooperative testing and are poised to take these concepts further as integrated systems advance. For programs like the P-8A, we are learning to leverage simulation expertise, tools, and facilities across test phases. Collaboration between E^3/EMC test engineers and flight/ground test engineers reduces cost by sharing simulation and stimulation assets and using common test methodologies. Significant schedule improvements can also be realized by conducting tests concurrently. These types of advanced ground tests have proven to reduce risk for programs and platforms undergoing developmental and operational tests. The ability to transition from ground-to-flight test with the confidence that all systems work as expected, that interoperability in stressing missions is assured, and that mission crews have fully rehearsed missions is key to efficient and cost-effective execution. With the facilities, laboratories, and simulators in place, the next challenge is to continue to strengthen working relationships and collaboration between Systems Engineering, Modeling & Simulation, Analysis, Training, and T&E communities, as well as strengthening interfaces with the commercial developers of tomorrow's weapons systems. The path to the future of a usable Joint Mission Environment for all phases of testing begins with small steps and innovative thought. For programs like P-8A and others, the process has begun and collaborative facilities and infrastructure are critical to future success. ❑

PAUL ACHTELLIK served in the U.S. Navy from 1968–1980 as an aviation electrician prior to his present involvement with electromagnetic compatibility (EMC) T&E. He's now a senior member of the EMC Branch within the Integrated Battle Space Simulation and Test Department at the Naval Air Warfare Center Aircraft Division, Patuxent River, MD. The EMC branch is responsible for conducting electromagnetic environmental effects (E3) tests on U.S. Navy aircraft, other DOD aircraft and similar full scale integrated systems. With 27 years of "hands on" experience, he's participated in over 150 E3 tests on a wide variety of aircraft and has covered all aspects of E3 T&E including; EMC, EMV, P-Static, Lightning, ESD, EMP, and EMI evaluations. Paul is Narte Certified in E3 and considered a "subject matter expert" in his field. He currently serves as E3 project lead and manager for all P-3 type aircraft, Unmanned Aerial Systems and the Navy's new P-8A aircraft. The complexity of hardware, software, and C4ISR intra/interoperability within these new aircraft systems cannot be understated. Advanced systems of systems are data driven and require complex inputs to determine if they are working correctly. Paul's E3 test methodology has evolved to meet this challenge, where, the future of testing requires a shift to "Operationally Relevant" test environments to accomplish E3 T&E effectively and affordably. E-mail: paul.achtellik@navy.mil

References

Military Standard MIL-STD-464A. 2002. Electromagnetic Environmental Effects Requirements for Systems, of December 19, 2002.

Scenario Development Strategy 13126/A1J1B/PMA-290/SE/1053/. 2006. P-8A Scenario Development Strategy of August 22, 2006.

ITEA Best Paper

ITEA Journal 2008; 29: 23–28

Design of the Ballistic Missile Defense System Hardware-in-the-Loop

James Buford, John Pate, and Bernard Vatz
Systems Simulation & Development Directorate,
Aviation & Missile Command, Huntsville, Alabama

Test and evaluation (T&E) of geographically dispersed integrated systems are severely constrained by cost, range safety restrictions, and ability to test while in an operational state. The Missile Defense Agency has embarked on a hardware-in-the-loop (HWIL) framework development that has the capability to characterize the performance of the Ballistic Missile Defense System by integrating the operational software in a distributed laboratory architecture. The HWIL framework is also intended to test the operational assets in their fielded configuration and location. As more advanced radar discrimination algorithms are developed, testing these algorithms and determining the impact on the system performance becomes increasingly more difficult. The ability to stimulate radar signal processors with synthetic signatures has also advanced over the last few years, thus enabling greater opportunity for testing. The integration of separate defense programs, and thus independently developed HWILs, has been a concern for the agency. The development of the Ballistic Missile Defense System HWIL will provide the agency with a unified architecture across all Missile Defense Agency programs, allowing consistent threat and environmental effects across all systems.

Key words: accreditation; advanced test facilities; complex operational systems; integrated network; realistic mission environments; simulator/stimulator testing labs; verification & validation.

Using the Ballistic Missile Defense System (BMDS) as an example, this article articulates the Missile Defense Agency's (MDA) hardware-in-the-loop (HWIL) framework design and development for testing the BMDS. This framework will allow MDA to establish a degree of confidence in the expected performance of a very complex operational system that cannot be evaluated by conventional tests. The inherent difficulty in executing an operational test in the conventional sense presents the Operational Test and Missile Defense Agencies with challenges to field such a complex system.

This article examines the benefits and challenges of implementing a distributed HWIL framework and articulates areas that are critical in design, implementation, and execution of the BMDS HWIL. In addition, the framework test and control functions, communication architecture, and interface requirements are discussed. Topics include

- BMDS components
- BMDS HWIL fidelity requirements
- Challenges of distributed simulation execution, including data latency, data rates, and synchronization
- Management and coordination of complex test requirements
- Common threat and environment for stimulation of simulation elements
- Methods for HWIL verification, validation, and accreditation.

The ballistic missile defense system

The BMDS Program is designed to provide protection against limited ballistic missile attacks targeted at the United States. The MDA mission is to develop, test, and field this missile defense system. Using complementary interceptors; land-, sea-, air-, and space-based sensors; and battle management

This article was selected as the ITEA Best Paper—Second Place at the Annual ITEA International Symposium, November 12–15, 2007, Lihue, Hawaii.

command and control systems, the planned missile defense system will be able to engage all classes and ranges of ballistic missile threats. All ballistic missiles share a fundamental characteristic—they follow a trajectory, which includes three phases: boost, midcourse, and terminal. By fielding a layered defense system and attacking the missile in all phases of flight, MDA can exploit opportunities to increase the effectiveness of missile defenses and complicate an aggressor's plans. The MDA has connected several test ranges to form the BMDS Test Bed, which will add realism to ground- and sea-based midcourse testing by allowing multiple engagements and different trajectories and adding additional intercept areas. The BMDS Test Bed also includes boost and terminal segment tests, which will demonstrate the viability of the layered missile defense concept.

The potential boost-phase defense elements are high-power Air-Borne Lasers and kinetic energy systems. The primary elements in the midcourse phase are the Aegis Ballistic Missile Defense and the Ground-based Midcourse Defense (GMD). The terminal elements are the Theater High Altitude Area Defense (THAAD) and the Patriot Advanced Capability 3 (PAC-3). Other elements include the experimental Space Tracking and Surveillance System along with its strategic and theater mission controller, the Command & Control Battle Manager and Communication system, and other agency experimental and operational sensors.

The test and evaluation challenge

Classical test and evaluation (T&E) of a new weapon system entails repeated live "firings" by forces that would be employing the system against the expected threats in an environment similar, if not identical, to the expected battle space. Although the BMDS Test Bed provides for more realistic operational testing and capability assessments, only a limited number of flight tests will be conducted. In support of system assessment activities, the T&E community will use flight test, digital simulation, and HWIL simulation data.

The BMDS HWIL framework provides a means to test the BMDS operational software in a controlled laboratory environment. The HWIL framework is also intended to test the operational assets at their fielded sites and host country. As new advanced radar algorithms are developed, the need to inject threat stimuli directly into the signal processor hardware increases. As much as possible, the architecture incorporates the component operational processing hardware and software that will be used in the field, implementing the "Test What You Fly, Fly What You Test" paradigm.

As the BMDS Block upgrades are developed, the impact on system-level performance must be determined. The HWIL framework will allow MDA management to evaluate the upgrades before fielding.

The MDA is requiring the BMDS HWIL to support BMDS system-level performance-based assessments and support BMDS system-level concurrent test training and operations functions. The HWIL framework will allow simultaneous execution of engagement sequence groups; testing both theater and strategic assets. MDA can use test data to assess interoperability of MDA elements, demonstrate the Command & Control Battle Manager and Communication system capability to control and manage BMDS communication networks, sensor management, and display situational awareness to the warfighter.

The Operational Test Agency also uses this test data to characterize BMDS operational capability, which includes threat detection, tracking, discrimination, engage, intercept, and destroy. Other objectives include characterization of information exchange capabilities among BMDS elements. The warfighter additionally wants to verify courses of action, tactics, techniques, and procedures.

Benefits to HWIL testing

With the complexity of the BMDS, integrating multiple systems into a joint fighting force is a challenge. Each element is a completely different acquisition and each has somewhat different requirements. Being separate, each element does not know exactly what dependencies and needs it requires for interoperability with the other elements. Independent testing and verification of the elements does not necessarily fully verify the BMDS or fully assess the system capabilities. If, for instance, the boost-phase elements cannot destroy the threat, their tracking data could be used to enable the midcourse battle-manager to use earlier and more accurate data to cue the midcourse element radars. The benefits of the BMDS HWIL are to help in flight test planning, interoperability, and performance assessment.

Flight test planning includes development of flight test concept of operations, timeline analysis for the mission director, determination of when to filter or include range radar track reports, evaluation of the exclusion of test range assets, pre-mission testing, verification of element interfaces, predicting the probability of mission success, and testing of off-nominal excursions.

The BMDS HWIL may also be instrumental in the design and development of the BMDS Battle Manager, which will have to interface with all element battle management systems. Areas of interest include message translation, message traffic analysis, situational awareness, allocation of interceptors, track correlation, search cueing, drop track reasoning, estimates of sensor covariance, and hand-over strategies between sensors of different elements during different engagement phases.

The most critical benefit is determining system capability and testing of block upgrades. The results of HWIL testing can be used to demonstrate and verify that system requirements are met. Analysis efforts include system capability assessment, kill vehicle and sensor acquisition, tracking and discrimination, and system battle-space evaluation.

HWIL description

This article provides a construct for implementation of a BMDS HWIL and is defined to include as much as possible the tactical hardware and software. HWIL facilities consist of space-based and radar sensors, interceptors, and battle management and communications. Obviously the radar antenna and the interceptor booster cannot be implemented in their entirety. Typically, the radar HWIL consists of the data processors and, in some instances, the signal processors. The interceptor HWIL usually consists of the data processors, which execute the guidance software and the software utilized to process the seeker imagery and determine the interceptor's acquisition, tracking, and discrimination performance. Typically the Battle Manager is represented by the actual tactical hardware and software, with the communication interfaces and simulated delays and timing.

The BMDS HWIL will integrate laboratory facilities in locations across the United States and integrate the fielded operational assets, including those in other countries and at sea. The BMDS HWIL will contain a network to transmit simulation truth data to the elements; a tactical communication network is also available to exercise and evaluate the real communication between elements. The simulation network uses the simulation protocol messages, while the tactical network uses satellite and fiber-optic links, with a variety of tactical message types.

The development of the BMDS HWIL framework will provide the agency with a unified architecture across all MDA programs, allowing consistent hardware, environment, and threat stimulation. Commonality is needed in order to reduce risk. The benefits to achieving commonality in the target generator include:

- Ensuring confidence and control of target data—"Single Source of Models."
- Ensuring consistent target representation across multiple elements—"ALL right or ALL wrong."
- Minimizing the difference in performance between elements—"Level Playing Field."
- Reducing development/modification cost and schedule—"One Time Fix."
- Reducing cost and schedule for element project offices (provides elements with HW/SW to drive stand-alone element testing/verification).
- Reducing target & environmental model verification & validation (V&V) cost and schedule.
- Maximizing reuse of target development efforts and code.
- Reducing risk of interpretation.
- Maximizing configuration control.
- Providing linkage and heritage between elements.

Depending on whether the test is for interoperability or performance verification significantly drives the fidelity and commonality of the target generator.

HWIL framework. The fidelity of the simulation representations can vary across different programs; however, the BMDS system engineer and integrator must determine the fidelity of the configuration needed based on the requirements and intended use of the simulation output data.

The element representations should at a minimum have the operational software integrated into the simulation or hosted on the actual tactical data processor hardware. In addition, the signal processor could be added, along with the missile HWIL, and in-band injection of scenes to the sensor.

The basic BMDS HWIL architecture will consist of the test, execution, and control (TEC) module, the Test Interface Unit, and the element HWIL representations.

Test, execution, & control (TEC). The importance of the TEC module is to establish the connectivity and determine the particular test cases and setup required. The TEC module must synchronize all participants' simulation time and provide the necessary initialization and start commands to each representation. The TEC module also provides updated interceptor state information from each element to the other elements participating in the exercise.

The TEC conducts three major functions: pre-mission, mission, and post-mission execution. In general, the BMDS HWIL pre-mission TEC provides single point control in defining test cases and providing the capability to specify test simulation start time (past, present, future).

During the actual test event execution, the BMDS HWIL mission TEC provides displays that summarize BMDS HWIL framework and element health and status, situational awareness of BMDS elements under test (element positions, sensor coverage, and threat), and framework and system events for monitoring. BMDS HWIL mission TEC also provides the capability to monitor and display run-time test integrity metrics to include framework and tactical message traffic, message latency, and loss.

After completion of the test case, the BMDS HWIL post-mission TEC provides the capability to import raw and/or processed data to a centralized database management system. This data will be provided to the MDA and Operational Test Agency (OTA) communities for analysis.

Test interface unit. Another critical piece of any HWIL is the target generator module. The test interface unit comprises modules to generate threat trajectories and dynamics, radar signatures, threat plume intensities, and interceptor signatures. In conjunction, common environmental libraries are utilized to induce effects to the signatures. The environmental effects include ionosphere, earth limb, refraction, attenuation due to standard atmosphere, and rain. Other celestial objects modeled include satellites, the sun, and the moon. Interceptor debris is also modeled. The resultant signatures are then provided to the component representations.

As more advanced radar discrimination algorithms are developed, testing these algorithms and determining the impact on the system performance has become increasingly more difficult. The ability to stimulate radar signal processors with synthetic signatures has also advanced over the last few years, thus enabling greater opportunity for testing. The test interface unit will have the ability to drive both the data processor and the signal processor to minimize the cost impacts of replacing all element representations.

Having a distributed, HWIL simulation architecture only amplifies the need for adequate timing analysis. Bandwidth often limits the data rates between facilities and elements. The HWIL system architectural engineer must determine the data rates at each level of the simulation from the TEC, to the target generator, to the element interface, and even the rates associated with tactical communications between the elements. A test interface unit will be co-located with each component to minimize data latency. Each component will have to have an element-specific interface to incorporate the different radar waveforms and integration rates needed.

MDA test events

The MDA has embarked on a test campaign for each year and block upgrade. The campaign consists of laboratory testing and operational asset testing.

Ground Test–Integrated (GTI) will be a distributed laboratory system-level test, utilizing MDA element HWIL facilities. The purpose of the test is to demonstrate the performance capability of the BMDS. The GTI will provide data for element and system-level assessments by executing a variety of scenarios and conditions, and evaluating sequences of events from the BMDS kill chain (e.g., detection, tracking, engagement, etc.).

Ground Test–Distributed (GTD) will be a distributed fielded system-level test. Each BMDS element has incorporated into the tactical operational software the ability to execute simulated tests, similar to the HWIL laboratories. The major difference between the GTD and GTI is that the GTD will exercise the tactical communication links from the actual fielded locations. In general, the test cases in the GTD are a subset of the GTI. The GTD is a progression of the GTI testing. GTD are intended to double check that the performance of the operational assets replicate the performance evaluated during the GTI test campaign.

The concurrent test training and operations concept will capitalize on the GTD architecture to allow the warfighter the opportunity to train and test on the operational assets, while maintaining operational capability to defend the nation. This concept will increase the requirements on both the HWIL framework and the operational system. However, the benefits to the warfighter to train while on station will significantly increase troop efficiency. The crews will be able to evaluate their tactics, techniques, and procedures and the command structure communications.

Evaluation

The test requirements process is a large and complex job. The challenge of writing good test requirements can be lessened if the flow down process is used to define overall objectives and operational scenarios. These will flow down to the system requirements, which will flow down to the subsystem requirements, and so on down to the test requirements. Simultaneously while developing a flow down process for the requirements, each requirement must be verifiable and able to fit into specifications. Good test requirements will be very specific and reflect the functionality of the components and, in turn, the system.

The primary objective of any evaluation activity is to determine if the test objectives and requirements have been met. This requires that any observed or potential system performance shortfalls be identified. A comprehensive set of system performance measurements, applied on a per-run basis is used to verify that system performance is maintained within established margins.

These margins define the limits of system performance relative to ensuring successful test implementation.

During each test case run, the critical mission timeline and the expected results for key system events will be documented on the test case run log for each test case. As the test case run is completed, the test director will indicate on the log sheet if the key system events occurred as predicted and if the expected results were obtained. All test case anomalies will be recorded on the test case run log and will be provided to the personnel performing the analysis. After the test case runs are completed, a post-test analysis will be performed. The analysis determines if the mission objectives were met and what the system performance margins are relative to the requirements. In the event of an anomaly, further analysis will be performed on the test case to determine the root cause of the problem and to provide a resolution. A daily assessment report summarizes the information collected during the post-test data analysis activities.

At the completion of the test, the evaluation team will produce a test evaluation report. The contents of this report will include a comprehensive evaluation and analysis of all test objectives and test requirements along with the system level assessment. The results will be made available to the BMDS systems engineer who, in turn, directs future development to improve performance and capability.

HWIL integration and accreditation process

There are four phases in an HWIL integration and accreditation process. The first phase is the delivery of element representations and their stand-alone, checkout testing. During this phase, it is the responsibility of the element integrated process team to deliver V&V data certifying that the model is a valid representation of the element within specified limitations and usage constraints. The second phase is the integration of the element representations into the BMDS HWIL framework, in accordance with jointly defined integration plans. Both the framework and element representations verify the interface control documents have been met.

The third phase includes two distinct activities: (a) element-to-element integration buildup, and (b) test readiness. The integration buildup part of this phase includes testing each element with system Battle Manager and then testing with all elements scheduled to participate in the HWIL configuration. After integration buildup, test readiness activities are conducted including regression testing, dry run execution, and finally lock-down of the HWIL configuration baseline.

All anomalies found during integration, regression, and engineering tests will be documented in Test Incident Reports (TIR). Each TIR will be isolated to an operator, framework, or element issue. The TIR is a management process used for documenting, disposition, and tracking test incidents for future development throughout the testing life cycle.

The output of phase 3 is a signed certification letter from each participating element stating their respective element has been successfully integrated into the HWIL in compliance with the Interface Control Document and can support the test objectives and test requirements. Collectively, the MDA and BMDS elements are executing an ongoing suite of V&V activities to establish the credibility of the element test articles. Each element program manager is responsible for reviewing the V&V data and the integration testing results, after which caveats and limitations are generated. This recommendation is to be delivered to the accreditation agent at the Preliminary Test Readiness Review (PTRR).

The fourth phase is the accreditation of the integrated HWIL test configuration. During this phase, the accreditation agent produces an acceptability assessment and accreditation recommendation, which is provided to the MDA directors of systems engineering and test and evaluation. The directors evaluate the accreditation recommendation and determine if the configuration is ready for test. A signed accreditation letter is then prepared and presented at the Test Readiness Review, which allows the formal start of test execution.

Inherent in this proposed accreditation paradigm is the execution with due diligence of commonly accepted modeling and simulation (M&S) V&V practices.

Verification and validation (V & V)

Verification is the evidence of compliance with requirements for a system (i.e., "Did I build it right?"). Simulation verification is confirmation that all data inputs, logic, calculations, and engineering representations within the simulation accurately portray the intended characteristics and interactions. Validation is the evidence of the system successfully achieving its intended purpose, or function (i.e., "Did I build the right thing?") Validation confirms that a simulation reflects real world expectations and is generally accomplished by comparing simulation results to actual flight test results or other external data. V & V should be implemented in the initial stages of the HWIL development and followed throughout its life cycle.

Failure to plan for proper V&V activities can lead to costly design and schedule ramifications. A clear process for the flow-down of accreditation needs into V&V data products and findings is required. The specific V&V activities identified for execution and the resultant V&V

documentation is explicitly identified in a formal V&V plan. All V&V activities should be selected for execution with the goal of satisfying the fundamental data needed to support an accreditation decision.

Caveats and limitations

A key feature of any accreditation decision is the identification of the caveats and limitations associated with the simulation configuration. Caveats caution analysts on the proper use of the test data, while limitations identify capability shortfalls in the test configuration. These caveats and limitations are linked to the specific test objectives and test requirements of a given test.

Accreditation

In accordance with MDA policy, all core M&S will be accredited to support acquisition decisions. M&S are abstractions and may not duplicate all actual, observed phenomena; however, they can provide reasonable approximations. Based on V&V activities and integration testing, an assessment is performed to determine the extent to which the HWIL configuration can meet specified test objectives and requirements. Accreditation is the official determination that the test resource provides credible data that can be applied to meet the intended uses within the stated caveats and limitations.

Summary and conclusion

This article articulates how fundamental test objectives can be met for a very complex system of systems, which cannot be evaluated fully through conventional developmental or operational tests. It examines the benefits and challenges of implementing a distributed HWIL to support such assessments using the BMDS as an instance. Areas that are critical in design, implementation, and execution of the BMDS HWIL are addressed. Based on V&V activities and integration testing, an accreditation assessment is performed to determine the extent to which the HWIL configuration can meet specified test objectives and requirements and to establish a degree of confidence in the expected performance. ❑

JIM BUFORD is an electronics engineer in the Aviation and Missile Research, Development, & Engineering Center (AMRDEC) of the U.S. Army Aviation & Missile Command at Redstone Arsenal, Alabama. He is the chief of the Missile Defense Functional Area and the BMDS HWIL system integrator working under the guidance and direction of Missile Defense Agency director of modeling and simulation (MDA/DES), specifically HWIL branch (DESH) or the MDA Model and Simulation Center of Excellence – Huntsville (M&S CoE-HSV). Buford's primary role is as the technical and programmatic lead for the BMDS HWIL Test Framework. He has a bachelor of science degree in electrical engineering with a double minor in mathematics and computer engineering from the University of South Alabama and has completed an extensive array of graduate and professional development courses at the University of Alabama in Huntsville. E-mail: jim.buford@us.army.mil

JOHN PATE is an aerospace engineer in the AMRDEC of the U.S. Army Aviation & Missile Command at Redstone Arsenal, Alabama. He is assigned to the System Engineering and Integration Directorate, Ground-Based Midcourse Defense (GMD) Joint Program Office in Huntsville, Alabama and serves as the government's verification, validation and accreditation lead for GMD HWIL simulations. Pate has a bachelor of science degree in aerospace engineering from Auburn University. E-mail: john.pate@mda.mil

BERNARD VATZ is the BMDS HWIL system integration lead, working under the guidance and direction of MDA/DES, specifically DESH or the MDA M&S CoE-HSV. The missile defense system exerciser (MDSE) provides the test framework for MDA's BMDS ground tests, exercises, training, and continuous test training and operations (CTTO) venues. Mr. Vatz leads the HWIL testing development and integration of defense system components. He is an electronics engineer at the AMRDEC of the U.S. Army Aviation & Missile Command at Redstone Arsenal, Alabama. Vatz has a master of science degree in engineering, systems engineering, from the University of Alabama in Huntsville and a bachelor of science degree in electrical engineering from the University of South Alabama. E-mail: bernie.vatz@us.army.mil

ITEA Journal 2008; 29: 29–36

Innovative Technologies and Techniques for *In-Situ* Test and Evaluation of Small Caliber Munitions

Andre Lovas
Georgia Tech Research Institute
T. Gordon Brown and Thomas Harkins
Army Research Laboratory

The Georgia Tech Research Institute and the Army Research Laboratory have collaborated in the Defense Advanced Research Projects Agency–sponsored SCORPION program exploring the application of microadaptive flow control techniques to small caliber munitions. This article discusses innovative techniques and technologies created in pursuit of the development, test, and evaluation of this new control technology. Tools developed include the use of g-hardened sensors, processing and actuator control electronics in 25 mm and 40 mm munitions. Inertial measurement units meeting all survival, packaging, and power requirements were designed and implemented using low-cost commercial off-the-shelf sensors including micro-electromechanical systems accelerometers and rate sensors and solid state magnetometers. Using resources integrated on the processor, flight data were recorded and stored for post-flight retrieval. An innovative projectile soft capture system allowed the projectiles to be safely recovered and reused multiple times. Data analysis techniques were extended to evaluate the in-flight performance of the microadaptive flow control technology. Further, the data served as a diagnostic tool to compare system flight performance with ground-based tests.

Key words: dynamic engagement test environment; guidance and control; integrated electronics; Maneuverable munitions; microadaptive flow control technology; SCORPION Program; spinning projectiles.

The Future Force Concept for the U.S. Army clearly outlines a strategy for operational scenarios that feature combined-arms operations in a multi-threat, dynamic engagement environment. Precision small to medium caliber munitions are integral and necessary elements of this strategy. To meet this vision, innovative techniques and technologies are needed for both the realization and test and evaluation of small, spinning, guided projectiles.

With support and direction from the Defense Advanced Research Projects Agency (DARPA), the Georgia Institute of Technology and the U.S. Army Research Laboratory (ARL) have teamed on the SCORPION (Self CORrecting Projectile for Infantry OperatioN) program to explore and develop the applicability of Microadaptive Flow Control (MAFC) technology for aerodynamic steering of spinning projectiles.

The SCORPION program was a multi-phase effort that comprised an initial technology feasibility phase, a technology demonstration phase, and a follow-on extension. The objectives of the feasibility and demonstration phases were accomplished through the successful integration of MAFC into a 40 mm infantry grenade surrogate, while providing sufficient divert control authority and adequate guidance and control to correct for projectile delivery errors and achieve required target impact accuracies. The work in the follow-on phase explored advanced microgenerator actuator technology and application of adequate MAFC-based divert capability in a high subsonic velocity 25 mm projectile (McMichael 2004). Program objectives included

- Develop g-hardened gas generator actuators and fabrication technology;

The Defense Advanced Research Projects Agency sponsored this work under grant No. DAAD19-00-1-0518. This article was selected as the ITEA Best Paper—Third Place at the Annual ITEA International Symposium, November 12–15, 2007, Lihue, Hawaii.

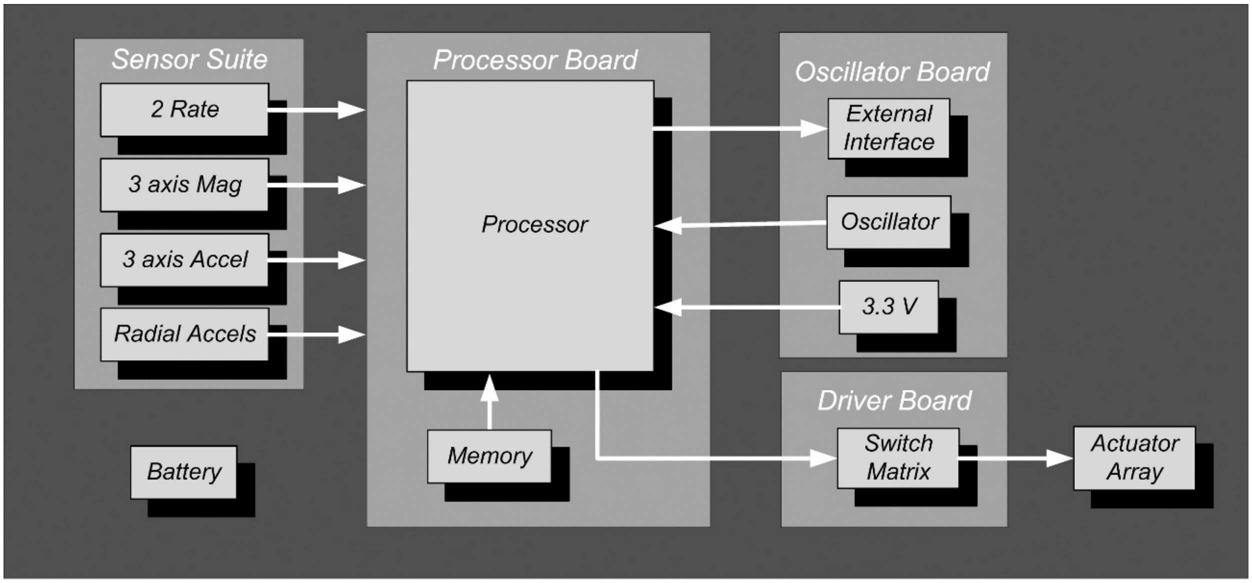


Figure 1. System block diagram showing sensor suite, processing, oscillator, and driver board

- Design, build, integrate, and test power, processor, and driver electronics for gas generator actuator systems;
- Research the nonlinear aerodynamics associated with the application of MAFC gas-generator actuators to high subsonic spinning projectiles;
- Integrate actuators and electronics into the flight control system;
- Miniaturize and g-harden the driver and flight control system for launch in a surrogate 25 mm projectile;
- Perform an open loop divert validation flight experiment of a gas generator actuator system using 40 mm projectile at Mach 0.25; and
- Perform an open loop divert validation flight experiment of a gas generator system using a 25 mm projectile at Mach 0.6 to 0.8.

This article summarizes the latest work concerning the open loop divert flight experiment of the high subsonic 25 mm projectile.

Integrated system description

While the program focus was on the development of MAFC technology, significant progress was made in the tools, techniques, and integration of technology for the guidance and control of small-caliber projectiles. Using a combination of commercial off-the-shelf components and components originally developed within the Hardened Subminiature Sensors Systems program for use in ARL's diagnostic fuze, an on-board inertial measurement system was designed and assembled (Lyons 2004).

The block diagram, *Figure 1*, shows the integrated electronics on board the 25 mm projectile. These electronics are hardened to withstand the in-bore acceleration forces experienced during gun launch.

Inertial sensor suite

The sensor suite contains two axes of rate sensors, three axes of accelerometers, and three axes of magnetometers oriented parallel to SCORPION's principal axes, and two additional radially oriented accelerometers. Outputs from these sensors combined with timing information from the oscillator were used by the processor to initiate commanded maneuvers. Sensor outputs were also stored in the processor for post-flight analysis and diagnostics. The processor and oscillator boards are shown in *Figure 2*. In *Figure 3*, the oscillator board, processor board, and the board-mounted sensor suite are combined (bottom to top) in a stack that functionally includes all the components of the inertial sensor suite (ISS) and the command guidance.

With the addition of batteries and a driver board, the electronics assembly is complete. This assembly along with the 25 mm SCORPION main body is shown in *Figure 4* with the driver board, batteries, inertial sensor

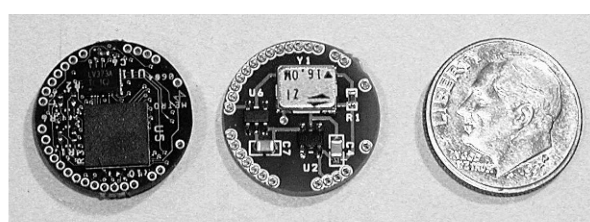


Figure 2. 25 mm SCORPION processor board and oscillator board

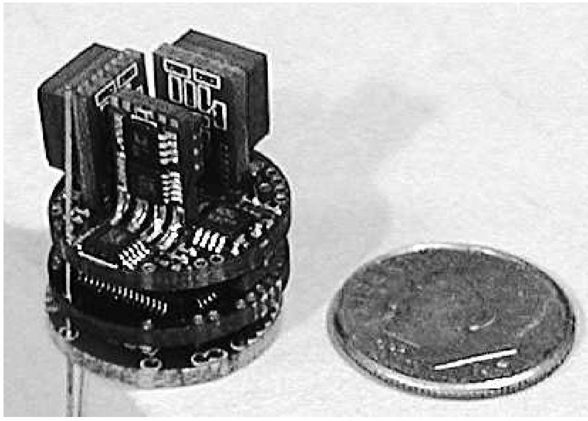


Figure 3. 25 mm SCORPION inertial sensor suite

boards, processor board, and oscillator/daughter board used for interface connection (from left to right, respectively). The diameter of each board is 17.5 mm, and the volume of the electronics package is 0.79 cu in.

On-board processing

Prior to launch, the on-board electronics are functionally checked and programmed with the initial flight conditions. The processor senses launch using the longitudinal accelerometer and starts the flight timing and data recording. The radial magnetometer is used to measure roll position and rate. For the open loop divert tests, the processor controls the start of the actuator firing sequence, the timing between firings, and the orientation of firing.

Calibration of inertial sensors was performed at various stages during the assembly process. However, careful attention to measuring the scale factor and bias was made before final assembly. Sensors were individually tested and aligned to assure that performance met requirements for bias and scale factor before integration with the electronics assembly. By using a methodical procedure of assembly and test from the component to the board level to the unit level, the need for corrective rework was reduced in the final assembly. Checkout and calibration of the integrated electronics included spin, magnetic, rate, and acceleration performance tests. Data from calibration performed after final assembly and potting were used to convert the inertial sensor outputs to engineering units.

Data acquisition system

An on-board data recording capability was developed and integrated into the SCORPION design. The data recorder stored 8064 records of data at programmable sampling rates from 1 kHz to 6 kHz. In typical conditions, the 4 kHz sample rate was used giving full coverage over the duration of flight lasting one to two seconds. The data system recorded 11 analog channels and four additional vehicle state channels. The data record had a 256 sample prelaunch record with the balance of recording data during flight.

Projectile design

The design of the 25 mm SCORPION was established with safety, reliability, and functionality in mind. The projectile is composed of two sections: the electronic control module and the actuator module. To meet the functional and safety criteria, the actuator module was separated from the rest of the assembly. This design allows for the separation of any potentially hazardous material, such as propellant, from the control electronics until just prior to firing. The electronics module is a potted cylindrical section housing the power, driver, IMU, processor, and connector boards, and a removable ogive (wind-shield) allowing access to the connector for communication, programming, and down-loading of data. The propul-

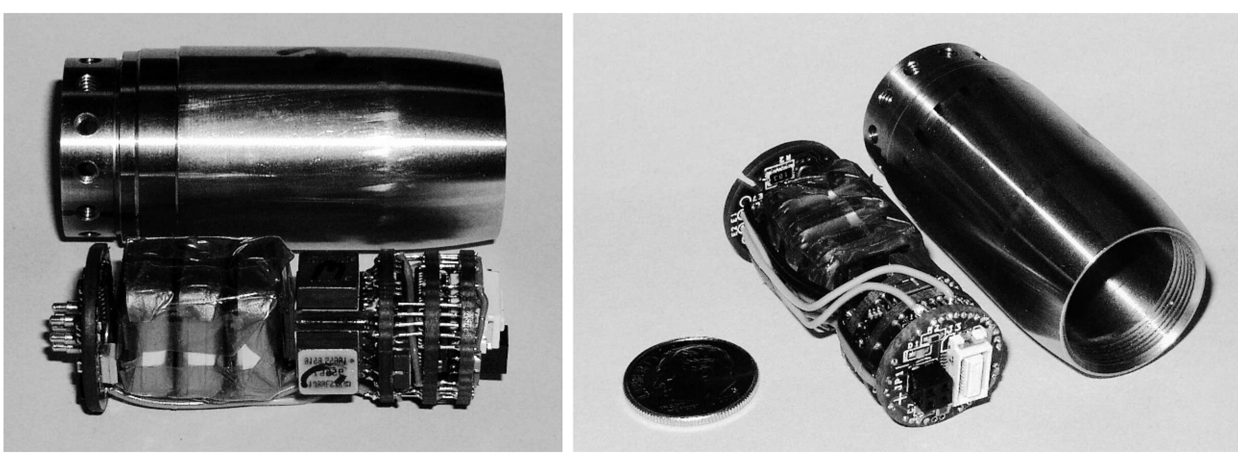


Figure 4. 25 mm SCORPION hardware and electronics assembly

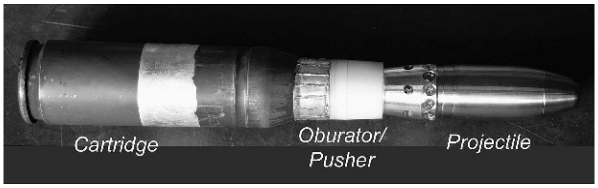


Figure 5. 25 mm SCORPION assembly with cartridge case, obturator/pusher, and projectile

sion system consists of a cartridge case housing the propellant, and an obturator/pusher assembly to seal the high pressure combustion gases in bore while transmitting torque for spin stabilization and distributing axial force to accelerate the projectile within the gun tube as shown in *Figure 5*.

Flight experiments

Initial flight experiments were conducted at ARL using a 25-mm barrel, shown in *Figure 6*, for interior ballistic design and soft recovery design. The primary objective for these tests was to establish an understanding of the propellant and cartridge case design required to launch the projectile at 0.8 Mach. This phase of experimentation was very successful at establishing a charge weight needed to meet the velocity requirements. Another goal was to establish a method of soft recovery for the 25-mm projectile. None of the techniques used in the past to recover small caliber projectiles was suitable for these tests because of the large standoff distances and other safety concerns. However, the idea of using layers of draped Kevlar to nondestructively absorb the kinetic energy of the bullet was explored and tested. This capture method proved successful as both the projectile and the pusher were slowed and captured. One of the captured projectiles, shown in *Figure 7*, was recovered after sustaining a launch acceleration of 25,000 g's.

Figure 6. 25-mm barrel used for interior ballistic design and obturator efficiency evaluation

Figure 7. Recovered projectile after successful in-flight silicon chip bridge initiation

Spark shadowgraphs taken during a test flight trajectory are shown in *Figure 8*. The initial yaw of the projectile at launch is approximately two to three degrees. Yet, after the maneuver, the resulting angle of attack is approximately 17.5 degrees. This result closely matches predictions from modeled trajectory simulations of approximately 18 degrees computed before flight testing.

Data from two of the sensor channels recorded onboard the projectile during a representative flight experiment are shown in *Figure 9*. These data begin just prior to launch and continue until shortly after impact. Thus, data from the launch event and the entire free flight motion of the projectile before, during, and after maneuver are included. The commanded divert was a single initiation at a timed delay from the launch. The launch was internally detected through comparison to an on-board accelerometer. Depicted are the two of the three axes of magnetic field measurement. Also recorded are angular rate in both the pitch and yaw directions, accelerations in all three orthogonal directions, and outputs from an additional pair of accelerometers used to estimate the projectile spin rate. From this raw data, post-processing could be accomplished.

Post-flight processing

Formulations of projectile flight dynamics; guidance, navigation, and control; and strap-down sensor locations, orientations, and outputs are most often done in a so-called "projectile-fixed" or "body-fixed" coordinate system. This system is right-handed Cartesian

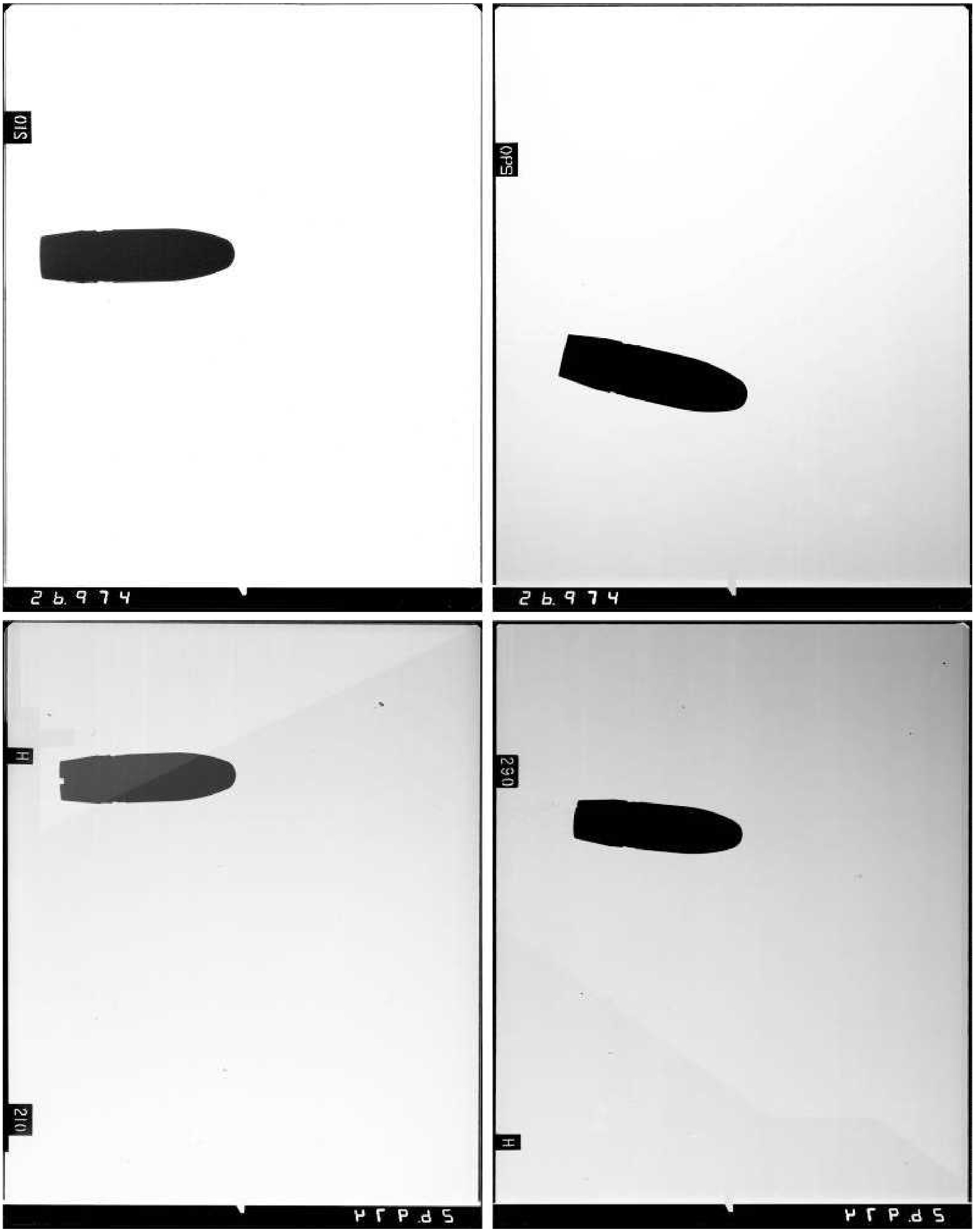


Figure 8. Orthogonal spark shadowgraphs depicting angle of attack before and after maneuver

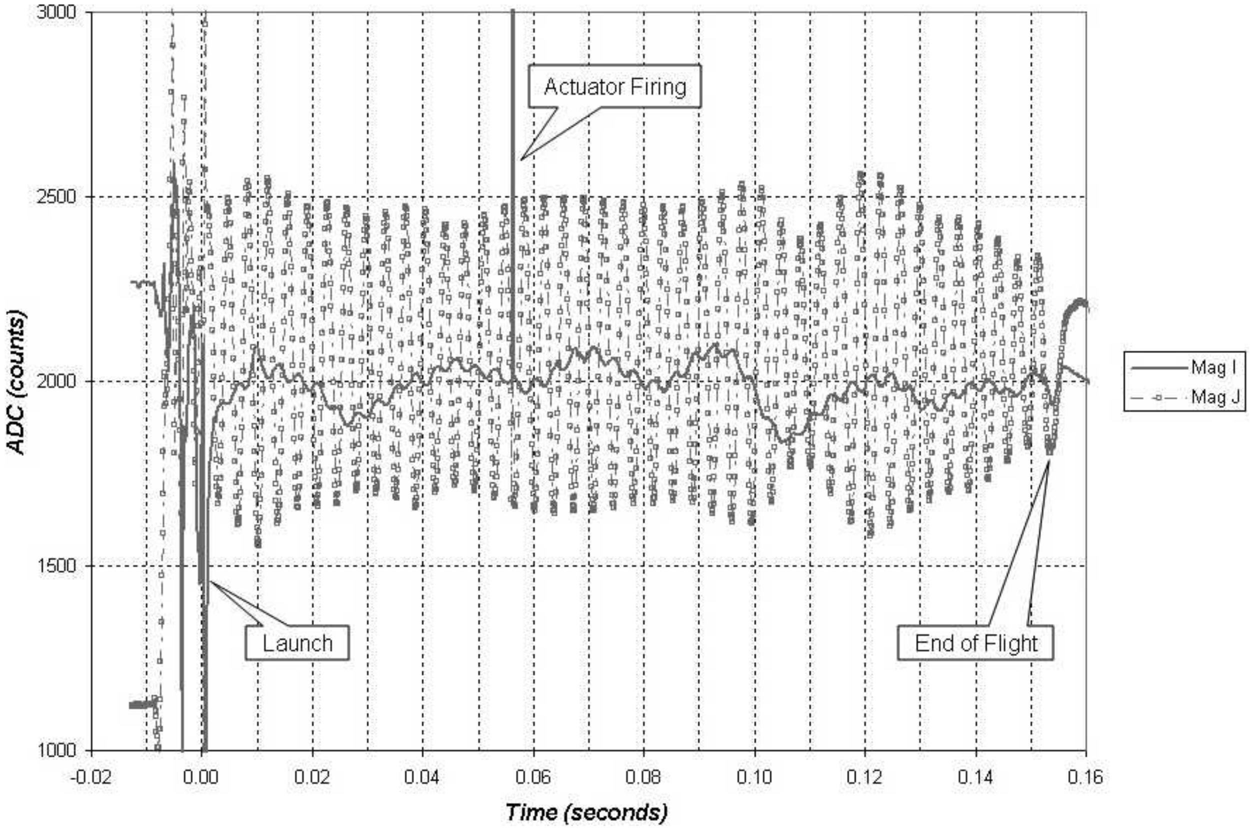


Figure 9. Recorded flight history of actuator initiation at approximately 0.06 seconds

with its origin at the center of gravity (cg) of the flight body. The body-fixed (I,J,K) coordinate system has its I axis lying along the projectile axis of symmetry, i.e., the spin axis (with positive in the direction of travel at launch). The J and K axes are then oriented so as to complete the right-handed orthogonal system *(Figure 10)*.

Among the many varieties of magnetic sensors, "vector" magnetometers are devices whose outputs are proportional to the magnetic field strength along the sensor's axis(es). SCORPION is equipped with a tri-axial vector magnetometer oriented with the sensor axes parallel to the projectile's principal axes. The

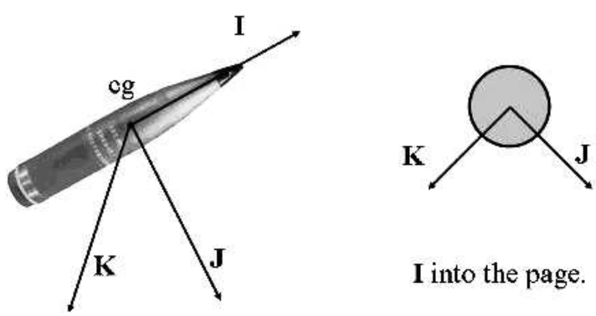


Figure 10. Body-fixed coordinate system

projections of the earth's magnetic field onto each of the sensor axes are given by the following equations:

$$M_I = \cos(\theta)\cos(\psi)M_n + \cos(\theta)\sin(\psi)M_e - \sin(\theta)M_v \tag{1}$$

$$\begin{aligned} M_J = {} & [\sin(\theta)\sin(\phi)\cos(\psi) - \cos(\phi)\sin(\psi)]M_n \\ & + [\sin(\theta)\sin(\phi)\sin(\psi) + \cos(\phi)\cos(\psi)]M_e \\ & + \cos(\theta)\sin(\phi)M_v \end{aligned} \tag{2}$$

$$\begin{aligned} M_K = {} & [\sin(\theta)\cos(\phi)\cos(\psi) + \sin(\phi)\sin(\psi)]M_n \\ & + [\sin(\theta)\cos(\phi)\sin(\psi) - \sin(\phi)\cos(\psi)]M_e \\ & = \cos(\theta)\cos(\phi)M_v \end{aligned} \tag{3}$$

where $\overrightarrow{M}_N = (M_N, M_e, M_v)$ is the magnetic field vector in a north, east, down earth–fixed navigation system, and (θ,ψ,ϕ) is the Eulerian projectile orientation vector in elevation, azimuth, and roll, respectively.

Because SCORPION's spin rate is large with respect to the yawing rates, the output from a magnetometer axis oriented parallel to the K body axis, designated Mag_K, is a sinusoid whose frequency varies with the projectile spin rate. For spin-stabilized and rolling

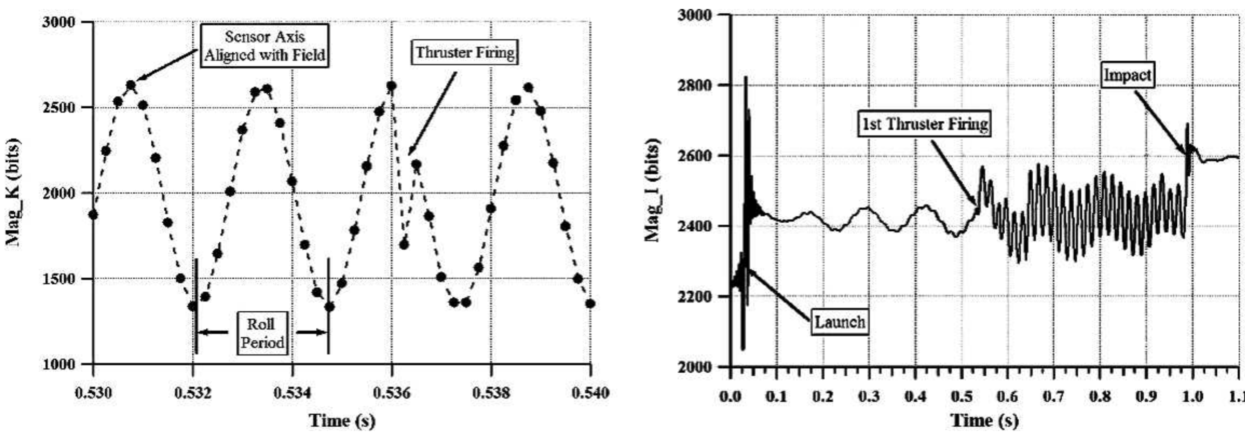


Figure 11. Representative magnetometer outputs: (left panel) radial magnetometer data—Mag_K, (right panel) axial magnetometer data—Mag_I

projectiles, the roll orientation must be known in order to properly execute desired maneuvers. With knowledge of the magnetic field, and knowledge of projectile elevation (θ) and azimuth (ψ), the roll angles at which Mag_K crosses the field (ϕ_M) correspond to the Mag_K extrema within a period. Ergo:

$$\phi_M = \tan^{-1}\left(\frac{\sin(\psi)M_n - \cos(\psi)M_e}{\sin(\theta)\cos(\psi)M_n + \sin(\theta)\sin(\psi)M_e + \cos(\theta)M_v}\right) \quad (4)$$

Evaluating Equation 3 at the principal value solution for ϕ_M shows whether Mag_K is at a maximum or minimum. Projectile roll orientation (ϕ) is estimated by computing ϕ_M at the times of each local maximum and minimum and then interpolating at intermediate times. Having thus produced a projectile roll angle history, the roll orientations at times of interest during flight can be computed. The output from an axis oriented parallel to the I body axis, Mag_I, varies directly with the angle between the spin axis and $\overrightarrow{M}_N$. This is called the magnetic aspect angle (σ_M). Time histories of σ_M provide information on projectile stability, yawing motion, damping characteristics, and maneuverability. An example of magnetometer data from a SCORPION experiment, annotated to highlight identifiable events during flight, is shown in *Figure 11*.

Post-flight application of these techniques to the magnetometer data yields critical information on maneuver mechanism performance and airframe response (see *Figure 11*). For this experiment, a 25 mm SCORPION projectile was programmed to execute a three-thruster divert to the right when looking downrange. After establishing the roll orientation of the Mag_K axis

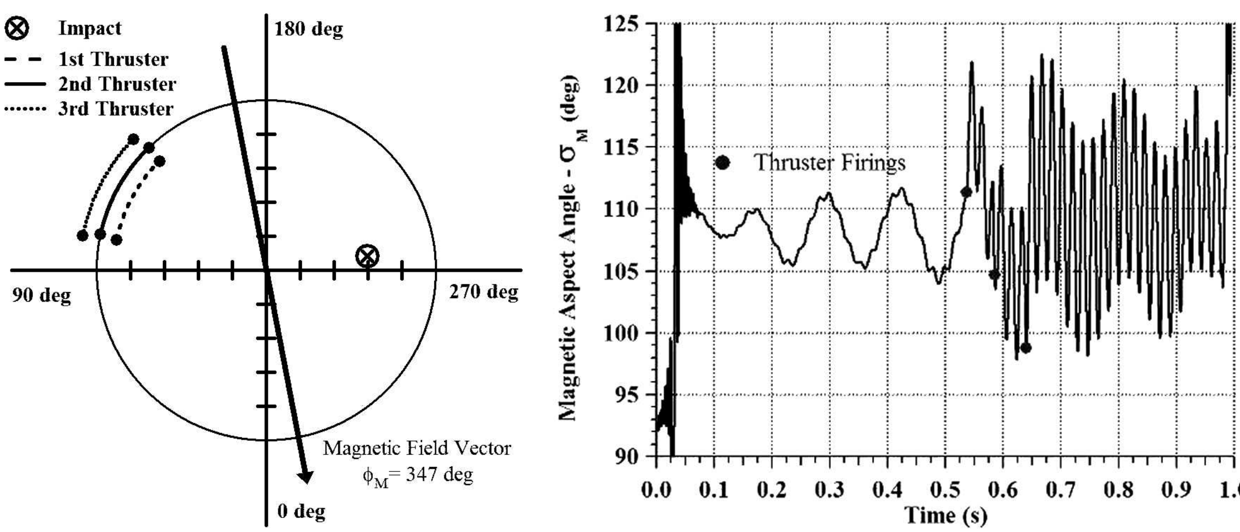


Figure 12. Performance measures derived from magnetometer data: (left panel) thruster orientations and projectile impact, (right panel) magnetic aspect angle history

at the thruster firing times, the roll orientation of the thruster nozzle when firing was computed from the known relative orientations of the thruster nozzles and magnetometer axes. The thruster orientations at their respective firing times are seen in *Figure 12 left panel* to be at about the 10 o'clock position. These orientations are plotted as arc segments to indicate the resolution of these roll angle measurements resulting from the combination of projectile spin rates and magnetometer sampling rates. These arcs indicate the performance of the on-board guidance, navigation, and control in executing the commanded maneuver. Also included in the figure is the projectile impact location (to the right) with respect to the mean impact point without maneuver. The associated magnetic aspect angle history, *Figure 12 right panel*, demonstrates that the yawing motion and maneuver, resulting from an individual thruster firing, depend on both the thruster orientation and the projectile yawing rates at the time of thruster firing. Understanding these interactions is crucial to designing an effective SCORPION guidance law in a tactical round.

Conclusions

In researching the feasibility of small caliber maneuvering munitions, a new diagnostic capability was developed. An integrated system design was required to provide a 17.5 mm data recorder with inertial sensor suite. This system has proven to survive in excess 25,000 g's in other applications. Its capability provides numerous opportunities for furthering the effort of guided small and medium caliber munitions. ❑

ANDRE J. LOVAS has a master of science degree in electrical engineering and is a senior research engineer at the Georgia Tech Research Institute. He is currently pursuing research on embedded processing and sensor integration in small caliber munitions. His research interests include the use of very high speed integrated circuit hardware description language (VHDL) to implement embedded digital hardware using field programmable gate arrays (FPGAs). He has developed an embedded microcontroller with real-time software control of imager operation, auto-exposure control, magnetic sensor processing, system timing, and 1 Mbps telemetry link. This embedded microcontroller is installed in a mortar-launched digital imaging reconnaissance system. E-mail: andre.lovas@gtri.gatech.edu

T. GORDON BROWN has a master of science degree in mechanical engineering and is the ballistics team leader in the Advanced Munitions Concepts Branch, Weapons and Materials Research Directorate, ARL. Recent research efforts have focused on design and development of high-G qualified miniature inertial sensor suites utilizing microelectromechanical systems (MEMS) components for medium-caliber military applications. Projectile design aspects covered by Brown and his team include mechanical and aerodynamic design. E-mail: brown@arl.army.mil

THOMAS HARKINS has both a masters of science and a masters of arts degree and is a research mathematician in the Advanced Munitions Concepts Branch, Weapons and Materials Research Directorate, ARL. Recent research efforts have included the design, modeling, implementation, and analysis of low-cost sensor systems for use in military ordnance. He holds five patents related to sensor-equipped projectile technologies. E-mail: tharkins@arl.army.mil

References

Lyons, D. 2004. "Embedded Instrumentation Technologies for Munitions." 2004 NDIA Armaments Conference.

McMichael, J., Lovas, A., Plostins, P., J. Sahu, Brown, G., and Glezer, A. 2004. "Microadaptive Flow Control Applied to a Spinning Projectile." AIAA-2004-2512, January 2004.

ITEA Journal 2008; 29: 37–44

Test and Evaluation: Department of Defense and Private-Sector Resources—Assessing and Resolving the Modernization Paradox

Drexel L. Smith
Wyle Laboratories Inc., El Segundo, California

A critical need exists to manage current test and evaluation assets, as well as to implement plans to maintain seldom-used facilities in parallel with developing new facilities for new and emerging technologies. Test and evaluation (T&E) provides a critically important element to the entire acquisition process, ensuring that any weapon system meets its intended purpose—from the component level to the full-up system. The T&E community consists of five primary sectors: (1) Military bases and laboratories (government), (2) Major prime contractors (industry), (3) Specialized subcontractors (industry), (4) Independent laboratories (industry), and (5) University research facilities (academia). While independent laboratories are often grouped within the "industry" category, they do in fact compose a separate tier consisting of thousands of dedicated firms providing laboratories and independent test services, and they invest heavily in their infrastructures. The federal government also has a significant investment in equipment and infrastructure in the United States for research and development (R&D) and T&E, much of which also is underutilized. As the experts examine the need to modernize, streamline, and better utilize existing facilities, it is important to consider all five components of the T&E community. This article provides a basis for discussion between government—primarily the National Aeronautics and Space Administration and the U.S. Department of Defense—and industry.

Key words: government-owned and contractor-operated; infrastructure; modernization; privatization; streamlining; test facilities.

Across the board, the test and evaluation (T&E) community is facing a paradox: As weapon systems advance in sophistication and complexity, T&E facilities require continuing modernization to keep pace. On the other hand, with U.S. Department of Defense (DoD) outsourcing dwindling, the industry and academic T&E communities find themselves with costly excess capacity in many areas, as well as a lack of capability to support new or emerging technologies. The unfortunate result is a spiraling increase in costs for maintaining enormous T&E infrastructures, thus severely reducing internal capital available for upgrades and modernization.

Nearly a decade ago, U.S. Secretary of Defense William S. Cohen called for streamlining of the Science & Technology, Engineering, and Test & Evaluation Infrastructure in a report to Congress (April 1, 1998). Secretary Cohen realized that significant cost savings could be realized, and test capabilities improved, by implementing a plan for streamlining and privatizing operation and management. However, government laboratories and weapons ranges have continued to grow in size, complexity, and configuration during the ensuing years.

Many T&E facilities exist that should remain owned and operated by the government, such as ranges that combine training and T&E, live-fire areas, and sea and air ranges. However, many others would benefit from streamlining and privatization if potential pitfalls can be identified and avoided.

Issues for consideration

Some issues to contemplate when developing a plan for streamlining include the following considerations:

- Government needs to recognize that it has been conducting itself in a "conflicted-environment way" when it comes to T&E.
- Government must recognize that the testing industry is more than just the capabilities that reside within the enterprises of its prime and subcontractors.
- The most effective use of existing T&E assets (government and industry) needs to be thoroughly evaluated and defined.
- Government needs to make hard decisions, specifically in terms of its own test facilities.

Conflicted environment: The government operates many T&E laboratories and facilities, much to the detriment of the industry that was created to provide T&E services in the first place. It has created a situation where many of its own facilities compete for business with private industry—which is the antithesis of privatization—and the result stifles private investment due to fear of competition with the government.

Beyond prime contractors

Thousands of independent test laboratories have been created over the years based on meeting DoD's needs. These companies employ a majority of the industry's true testing experts, who in turn provide "third-party" evaluation and qualification for DoD's most critical defense systems. Yet, these same companies are overlooked when DoD endeavors to evaluate the status of T&E and determine its future. These companies are crucial to the overall plan, and their voices need to be heard and acknowledged.

Defining effective use of assets

The DoD, the National Aeronautics and Space Administration (NASA), prime and subcontractors, independent laboratories, and academic institutions all have "labs." These labs are used to provide vital services for preproduction testing as well as long-term research and development (R&D), where they specialize in both R&D and T&E. To ensure the best results from any new streamlining effort, planners must define roles and responsibilities, review available assets and infrastructure, and define the alternatives for new working relationships among government, industry, and academic participants.

Making the hard decisions

New approaches are needed, including creation of new performance objectives and metrics to measure utilization for comparison with critical requirements. In some cases, the requirement is based on utilization and cost; in other situations, a facility is needed to demonstrate a new or emerging technology that may not offer a financial payback. For example, NASA, DoD, and the private sector all have wind tunnels for development and qualification work; however, too many facilities are competing for work in the lower-flow regimes and none that operate efficiently in the high mach numbers. For more than 10 years, industry and government have been in discussions, and thus far have been unable to come together with a workable solution that takes advantage of all facilities or that addresses both current and projected needs.

Optimal solution

Based on decades of experience in the independent testing industry, it is fair to say that the optimal solution is one in which the government drives the existing private testing industry to invest its own money in future technologies and facilities. However, to accomplish this goal, the government needs to reduce its role as a direct competitor in the T&E industry, except where unique situations exist.

Current assets and infrastructure

As mentioned previously, the federal government has a significant investment, measured in tens of billions of dollars, in R&D and T&E equipment and infrastructure and laboratories, throughout the nation. However, industry also has a significant investment in facilities and equipment, especially the "full-service, one-stop" independent test laboratories that maintain facilities to accommodate everything from basic components to full-scale rocket and weapons systems. Since the end of the Cold War, many test-related resources in both the government and private sectors are now underutilized, even with the new set of conflicts and challenges in the Middle East.

Government range and test facilities, covering thousands of square miles, account for much of the government's investment. However, there is also a large mix of facilities operated as federally funded R&D centers, including those operated by industrial firms (Sandia, Oak Ridge, Savannah River); those associated with universities (Ames, Joint Propulsion Laboratory, Lawrence Livermore); and those operated by nonprofit institutions (Aerospace Corp., National Defense Research Institute, Project Air Force). Contractors also operate a large number of T&E facilities and ranges for the government, including the Air Force Flight Test Center at Edwards Air Force Base (AFB), Space and Strategic Defense Command, and the Naval Surface Warfare Center.

Overall, downsizing and Base Realignment and Closure (BRAC) programs have had little impact on apparent laboratory over-capacity, so planners need to rethink how to best utilize all available resources and develop a cooperative, long-term relationship between government and industry. Part of this process will be to more clearly define: (a) the true cost of operating, maintaining, and improving R&D and T&E capabilities; and (b) the most effective interfaces among government, industry, and academia.

As stated at the outset, T&E facilities are critically important to the acquisition process. Yet, as critical as they are, they offer a poor return on investment and return on net assets, especially in light of the continuing investments for modernization to accommodate the latest DoD warfare technologies. With regard to over-capacity, the result for DoD is significantly increased R&D and T&E costs that eventually drive up the overall procurement costs for any new weapon system.

Defining roles and responsibilities

The government (NASA and DoD specifically) is inherently responsible for: (a) defining mission requirements and specifying the needs, (b) establishing procurement and fiscal controls and contracting methodologies, and (c) accepting the fully developed systems.

Traditionally, industry has taken the government's requirements, created the optimal design and performed the complete manufacturing process (design, fabrication, production, and distribution). At the same time, whether intended or not, the government has moved increasingly into R&D, laboratory testing, T&E/operational T&E, and live-fire testing and training. Producing an effective plan for streamlining the process will require revisiting each of these areas. Over the years, many studies have been undertaken to review the nation's research, development, and test and evaluation (RTD&E) needs and capabilities, and some of the results have been implemented while others have been largely ignored.

Independent test labs

With commercial off-the-shelf (COTS) and fixed-price and warranty systems, industry in general has assumed increasing responsibility for the demonstration, validation, and reliability of weapons systems. Often forgotten, however, are the independent laboratories, which are grouped into the industry category but are not part of the typical prime contractor lab structure. With thousands of facilities around the world, they have a separate identity and perform a unique range of valuable roles and functions:

(1) Trained personnel and dedicated facilities in specific areas of expertise;
(2) Costs allocated across a large number of users ("pay as you use");
(3) Testing is core business and focus is on innovation, modernization, and cost effectiveness;
(4) High volume and repetitiveness that allow complex work to become routine;
(5) Unbiased results without conflict of interest; and
(6) Independent quality assessments.

Industry's role is too often viewed simply as "contractor operated," but it is now necessary to look at the bigger issue. This is not just about changing badges of existing staff. Both government and industry have top-quality people, but government and industry (including independent test labs) must have *clear roles and responsibilities*. Both parties must participate in a dialog to review appropriate roles and missions to ensure the nation's continued excellence in T&E.

As part of this process, industry must continue to: (a) maintain a solid infrastructure of laboratories and support facilities; (b) provide third-party demonstration/validation services; (c) provide investments in modernization of facilities, equipment, and manpower; and (d) provide operation and maintenance (O&M) contract labor *at competitive costs*.

Facilities and capabilities in both government and industry are underutilized. The independent lab industry, for instance, has more than 7,000 organizations that list 8734 as their primary SIC Code. Many are narrow-range, special purpose facilities, but a relatively few possess a full range of test capabilities and must—of necessity—compete with the small specialty labs, which have lower operating costs and overhead. At the same time, according to the General Accounting Office, significant T&E excess remains in DoD and other government organizations.

Government and industry relationships

Traditionally, industry provides support service labor at RTD&E facilities, but partnering on facilities can be accomplished much more comprehensively than is being accomplished today. Government and industry already have a long-term partnering relationship that includes general services (cleaning, cafeteria, gardening, maintenance), technical services (computer systems, metrology, O&M test systems, O&M ranges) and facility O&M (joint facility use, multi-investors, expanded partnering). With these as a model, planning must move forward to combine T&E resources to make the industry cost-efficient once again for all concerned.

When it comes to increasing the return on investment for laboratories, a new language must be developed with the right vocabulary:

- *Activity-Based Costing*: Assigns costs based on consumption;
- *Joint Ventures*: Public/private partnerships;
- *Outsourcing*: Government is responsible while another organization completes the work;
- *Service Shedding*: Divestiture when service is no longer provided; and
- *Vouchers*: Government subsidies.

Key elements of the acquisition process in T&E measuring are effectiveness, reliability, and suitability. By definition, T&E requires a significant investment in infrastructure, specialized equipment and skilled manpower, so the critical question becomes: How does DoD ensure, with a high degree of confidence, that T&E data are unbiased, objective, appropriate, reliable, and valid?

In many ways, industry is ahead of the government in implementing capacity reductions, driven by the constant review of assets that do not produce a return. The metric most used is "RONA"—a review of Return on Net Assets including land, buildings, facilities, and equipment (minus liabilities). For instance, a major aerospace contractor has instituted significant consolidations by focusing on its core business, developing or enhancing partnerships with suppliers and service providers for noncore services and goods, and reducing the number of internal laboratory facilities.

Also, many prime and subcontractors are *"surplussing"* their excess test equipment, allowing smaller companies, such as independent labs, to purchase well-maintained and reliable equipment to continue its useful life cycle. Partnering agreements between contractors also have provided cost-effective ways to keep vital facilities available. If one company does not want to lose access to an underused capability, it has the option to outsource the entire facility to a commercial lab. This is a win/win strategy: The first company can negotiate full access on a priority basis, while the second company enhances its capabilities to support other customers.

As another strategy, several major prime contractors are participating in a "laboratory alliance" to minimize excess capacity through the sharing of resources. This method uses a single work authorization under a multilateral services agreement implemented by several companies.

Two obstacles that continue blocking effective partnering are the incomplete move to COTS standards and the lack of good databases. While the move to COTS is long overdue in some circles, no commercial standards have been determined except in a few specific areas. Moreover, failure (fragility) limits have not been established, and there are no reliable comparisons of reliability versus performance limits at the component level.

Both government and industry representatives need to change the way they think about the respective R&D/T&E roles. For example, not every entity needs its own lab or test range, and with today's data transmission technologies, researchers need not be in proximity to the actual test. Examples abound: from flight test models to the eventual exploration of Mars, *a reduction in facilities will allow for more efficient use of those remaining*.

Metrology is another area for consolidation. Each branch of the military, NASA, and Department of Commerce (National Institute of Standards and Technology) maintains extensive calibration standards laboratories, but industry also has extensive capabilities. Unique requirements do exist, and specialty facilities will always be required, but there is an optimum balance between maintaining in-house capability, outsourcing and combining facilities across all branches of the military services and government agencies.

Time for decision and solutions

The time for decision is rapidly approaching, and that decision must resolve the original paradox if T&E is going to move forward: Investments are needed to modernize, but excess capacity needs to be reduced to free up capital.

Competition between the public and private sectors is not the answer. The real solution is to eliminate excess capacity, develop partnerships for joint use, and *provide sustainable funding opportunities to industry for continuing operations*.

A mechanism is required to review all options across the military services, government agencies, industry, and academia. A set of standard reporting formats must be developed to ensure consistency. Roles and mission statements must be drafted in enough detail to determine and eliminate overlaps, because industry does not want to invest in capabilities that already exist or in resources that will force it to compete with government facilities.

Accurate and detailed assessments are needed to determine the cost basis. After many studies, there is still no accurate database with regard to R&D, T&E, and laboratory capabilities across the United States. Facilities often have similar equipment performing different functions. The task at hand is to perform a needs assessment, properly align the cost basis, project future needs and make detailed adjustments. This effort will be unproductive, however, unless the available information covers all government agencies, industry, and academia.

Industry's role in fielding top-quality systems, earlier and at less cost, is this: Continue to provide O&M contract labor at competitive prices; maintain a solid infrastructure of laboratories and support facilities; participate in rethinking how vast resources can be "right-sized"; work to change the paradigm of "every organization needs its own facility"; and support the development of COTS standards.

The solutions are within reach, but the hard decisions still need to be made. A good start is to define the roles and responsibilities of all major stakeholders, which include: DoD, other government agencies, industry (especially independent test labs), and academia. The objective is to discontinue the competitive approach for utilization of T&E facilities, as well as to eliminate and discourage excess capacity throughout the T&E community.

Finally, and most important, the Secretary of Defense needs to make a firm commitment to move forward, by taking concrete action that encourages partnering with the government and that stimulates industry investment via increased opportunities to share in test services contracts.

Following is a discussion of several attempts at implementing a Government-Owned and Contractor-Operated (GOCO) plan for major test facilities—an idea whose time has come once again, and one that needs to be carefully reexamined in light of new and innovative concepts, as well as lessons learned.

GOCO case studies

In recent years, Wyle Laboratories and others have worked with both DoD and NASA to manage and operate government test facilities, including:

(1) McKinley environmental test facility, Eglin AFB
(2) Landing gear/tire/brake test facility, Wright-Patterson AFB
(3) Building 65 structural test facility, Wright-Patterson AFB
(4) Environmental test facility at Naval Air Warfare Center (NAWC)-China Lake
(5) Laboratory consolidation at NASA Kennedy Space Center (KSC)
(6) Hyperbaric chamber at Wright-Patterson for treatment of burn victims
(7) Human centrifuge at Warminster, Pennsylvania
(8) Propulsion test facilities at Air Force Research Laboratory (AFRL)/Arnold AFB
(9) Centrifuge and human effectiveness facility at Brooks AFB

While some of these studies have been successful, others have not fared as well. In either case, the lessons learned will be extremely valuable in determining a future methodology for using a partnership as a means of reducing redundant resources and the associated O&M costs—for both DoD and Wyle.

Figure 1. The McKinley climate laboratory main test chamber at Eglin AFB

McKinley environmental test facility: Eglin AFB

The McKinley Climate Laboratory Main Test Chamber at Eglin AFB is the largest facility of its kind in the world. The environmental trials at Eglin are a major milestone on the way to proving the maturity of the Nimrod design, as well as freezing the production aircraft design by the end of the year (*Figure 1*).

At issue is the fact that this facility is very expensive to maintain and operate. And, as a U.S. Air Force facility, the demand is limited.

Competition is from a variety of sources, and options could include:

(1) A cold weather outdoor test facility in Fairbanks, Alaska, which operates at $-65°F$ for more than four months of the year as ambient conditions. The cost of operation is minimal, but operators are subject to nature for control of test conditions. This may be acceptable for many test programs, but for running a test under specific laboratory conditions, it may be too risky (a similar outdoor facility exists in North Dakota).
(2) Hundreds of environmental conditions test chambers exist in government and private facilities. If testing can be performed at the component level, the need for full-scale testing might reduce the need to maintain the McKinley Climate Laboratory.
(3) Some customers do not want testing to be conducted in a government facility because of concerns about protection of data (a lingering view exists that for any test in a government facility, all data become public). In addition, the

government cannot commit to a fixed cost or specific schedule. Both concerns could be alleviated with a government-owned/contractor-operated type of program. A critical question then becomes: "Could a contractor-operated McKinley Climate Laboratory allow for additional testing to be conducted so that there would be sufficient funds to offset the costs?"

Landing gear, tire, brake test facility: Wright-Patterson AFB

Through its legacy companies, Wyle operated this facility from 1966 to 2005. During that period, the government workforce declined from nearly 20 to just two personnel, and it became a GOCO facility. This required Wyle to assume increasing duties beyond the core test and engineering mission.

With the expanding scope, Wyle developed a greater ability to scale personnel resources up or down, and the company established a cooperative agreement with the government to keep facility utilization high. Wyle brought $1.5 million of external funding into the facility in its last year of operation, making it the location of choice for outside testing by companies such as Goodyear and Michelin.

Wyle offered to take on full responsibility for the facility if the Air Force would provide a commitment of workload to baseline the costs. The result was that O&M was moved from the research section of the laboratory to logistics.

Building 65 structural test facility: Wright-Patterson AFB

As with the McKinley Climate Laboratory, the structural test facility (Building 65) at Wright-Patterson has housed a world-class structural test capability able to accommodate full-scale aircraft. On several occasions, the AFRL expressed an interest in making the facility a GOCO operation, and Wyle has offered to take on the role of contractor—with a commitment to migrate the facility to a Contractor-Owned/Contractor-Operated (COCO) facility.

By including both government and commercial workloads, there should be sufficient demand to maintain such a world-class facility should the Air Force decide to move in that direction.

Environmental test facility: NAWC-China Lake

Wyle Laboratories and NAWC-China Lake developed a working model by which NAWC could perform tests for Wyle, and Wyle would have access to NAWC's environmental test facilities. The long-term vision was for Wyle to establish a commercial laboratory within the NAWC complex, with NAWC providing support in specialty activities such as insensitive munitions and ordnance function tests. A cooperative agreement was developed as a contracting vehicle, and a number of tests were performed by NAWC under its terms.

Unfortunately, Wyle could not determine a sufficient level of business, nor could NAWC commit to a continuing level of environmental testing to justify proceeding with the development of the internal laboratory concept. However, such an agreement is an attractive option that remains open to this day.

Laboratory consolidation: NASA's KSC

As part of the winning proposal to manage the Joint Base Operations Contract, which covered both NASA/KSC and the Air Force/Cape Canaveral Air Station, Wyle offered to review and develop a concept to consolidate the metrology and nondestructive testing laboratories into a single complex that would be operated as a COCO facility. Located outside the complex gate (the research park is near the visitors center), it allows easy access for commercial users throughout Central Florida.

During the first year of the contract, Wyle completed the business model, participated in site selection activities at the proposed KSC Research Park, and developed a business plan and funding program. The process was stopped, however, as a result of land management and environmental sensitivity issues requiring a complete site assessment.

Wyle was able to successfully integrate the Air Force and NASA metrology laboratories into one operation (three locations), which resulted in substantial cost savings; however, total commercial laboratory consolidation remains under discussion.

Although a work in progress, this consolidation is an example of a government/private-sector program that can work, with all stakeholders emerging as winners: (a) The government will be able to reduce spending to maintain expensive facilities; (b) a world-class capability will be available to a wide range of users throughout Central Florida; (c) the throughput of work will be increased, reducing time and costs per item; (d) the contractor will secure a steady workflow for many years; and (e) employees will enjoy a new and dynamic work environment.

Hyperbaric chamber for treatment of burn victims: Wright-Patterson AFB

This unique facility, used by AFRL to study oxygen effects under pressure, has been semi-privatized to allow burn patients to receive oxygen treatments. The program allows for increased utilization of the facility to defray costs and allows AFRL to retain ownership for continued research as funding permits. This is another example of a program in which *everyone wins*.

Human centrifuge: Warminster, Pennsylvania

A Wyle legacy company, Veda, won a competition in 1996 to privatize the NAWC's Aircraft Division Warminster dynamic flight simulator/human centrifuge. The Navy retained ownership of data and removable cockpits (and other proprietary materials) in exchange for issuing Veda a sole source contract to operate the facility.

This turned out to be an *unworkable arrangement* for the following reasons:

(1) Although Veda operated the dynamic flight simulator, it was responsible for all costs, including the facility's rental. So, the "Government Owned" portion of that deal was actually a misnomer (Navy support of some of the incurred facility costs would have helped the privatization effort).

(2) Veda had no real control of the facility or freedom to market it commercially. The Navy controlled who used the facility and what could be done in there through the constraints it built into the sole-source contract on using the Navy's government-furnished equipment and how operations had to be conducted. For example, Veda could not use the government-furnished equipment to market the facility for g-tolerance improvement program training to augment its limited R&D projects because that was viewed as competing against other dedicated Navy g-tolerance improvement program training facilities.

(3) The original business plan involved trying to run the facility as the Navy had, with a marketing strategy based on previous customers providing adequate funding for operation. When these old funding sources dried up, a scramble took place to identify new funding sources and new customers. The Veda business plan assumed the Navy would provide at least $500,000 per year of project work to support the transition to a commercial operation, but it never materialized.

During this period, Veda also tried to privatize the large anechoic chamber at Warminster to support an emerging communications and antenna prototyping and testing business. In that case, there was no commercial competition, but the Navy stopped the deal to avoid competition with the facility that it was replicating in Maryland. Penn State's Applied Research Laboratory (ARL) has been in negotiations with the local township to take over operations and resurrect the facility, but the specific terms of the deal are not known.

One example of a successful privatization is the Inertial Navigation Facility at Warminster by the Penn State ARL. This was successful because a continual flow of funded Navy navigation programs enabled the operation to remain viable without participants having to find new customers and funding sources for near-term survival. This relationship continues 10 years after base closure and has allowed ARL to expand its navigation resources to other government and commercial areas.

Propulsion test facilities: Arnold AFB

Arnold Engineering Development Center (AEDC) and Lockheed Martin Space Systems Company signed a memorandum of agreement in December 2000 for a 10-year alliance for electrical propulsion testing opportunities in the center's Space Environmental Chamber 12V.

The purpose of the agreement is to work together to accomplish product research testing, product development testing, and engineering manufacturing development testing of Lockheed Martin electric propulsion systems at AEDC. When upgrades and checkout are complete, AEDC will provide electrical propulsion testing facilities and capabilities, and Lockheed Martin will provide the integrated component systems for testing.

Centrifuge and human effectiveness facility: Brooks AFB

Under the most recent BRAC program, the human training centrifuge and other human effectiveness test systems (hyperbaric and hypobaric chambers, disorientation simulators, and ejection seat trainers) are to be relocated from San Antonio, Texas, to Dayton, Ohio. From a practical standpoint, the cost of relocation is so substantial that constructing a new centrifuge and other equipment is more cost effective.

At this point, the AFRL is performing cost studies, but it seems clear that a new dynamic flight simulator will be constructed in Dayton to replace the centrifuge at Brooks AFB. This situation creates another opportunity, and Wyle has proposed to take over the system's operation at Brooks AFB to provide a commercial screening and training facility for potential commercial space travelers. This concept will allow the Air Force to have continued access to the Brooks AFB system as a backup until the replacement facility in Dayton is operational.

Lessons learned from the failed attempt at Warminster are being applied to allow for this anticipated operation to become a commercial success—provided there is sufficient demand for the services. As a contingency, the business model has been constructed based on no commitment for continuation by the Air Force.

Summary and conclusions

There are countless examples of situations in which the government can and should make an investment and then own and operate a test facility (the Department of Energy weapons complex is a prime example). And, there are equally many examples in which original equipment manufacturers, prime contractors, academia, and independent laboratories should make the investment to own and operate their own test capabilities.

At one end of the spectrum are material coupon testing and standard analytical chemistry tests, where the U.S. marketplace hosts several thousand privately held, commercial testing laboratories. This type of work can be purchased as a "price sheet" commodity on a by-the-test basis. With this highly competitive and robust industry available, one would need a strong justification to develop a new facility. At the other end of the spectrum are the highly sophisticated and unique facilities such as a high-energy laser facility where only DoD could justify its need.

But, one must consider all the test facility demands that fall somewhere in between. Following are some thoughts on the topic:

(1) Commodity-level testing should be left to the private sector.
(2) Because a prime contractor or original equipment manufacturer is generally in the business of manufacturing a product, an investment in a full test facility or in costly test equipment is viewed as a business cost that must be recovered.
(3) Academia is in the business of sharing knowledge, so access to test facilities on an as-needed basis is more important than ownership.
(4) Only the government has the ability to construct and operate test facilities that are unique and may not provide a financial return.
(5) The independent test laboratory industry fills the gaps.
(6) Not every organization can or should own a test facility because data and information can be transferred and shared in real time with today's technology.
(7) Costly test facilities should be shared to the greatest extent possible to lighten the burden on any one organization and to add value to the test results.
(8) For unique facilities, a forum involving all stakeholders should be established to review these key questions:
    (a) Who is in the best position to establish the requirements?
    (b) Who can best design and construct the facility?
    (c) Who should finance and own the facility and, by what means?
    (d) Who has the required experience and track record to operate the facility?

These considerations and questions are not new, but they need to be revived and revisited as part of a meaningful dialog on resolving the modernization paradox. As is obvious, the government needs to recognize that the testing industry is more than just the capabilities that reside within the enterprises of its prime and subcontractors. It must work with academia and the industrial sectors to determine the most effective use of existing assets (government and industry) and plan for the efficient use of new assets.

It is time for the government to step up and make the difficult decisions, and for stakeholders in the other sectors to demonstrate their commitment as well.

It is recommended the senior leadership within DoD and NASA join together to facilitate further discussions with industry to develop a pathway for determining the best utilization of test facility resources. The term "industry" should not be considered just the major prime contractors, but in this case must include the "independent laboratory" industry as well. ❑

DREXEL L. SMITH is the senior vice president, Corporate Offices, for Wyle Laboratories Inc., El Segundo, California. He has more than 39 years of experience in the test and evaluation of weapons, munitions, propulsion systems, nuclear systems and components, in positions ranging from test engineer to corporate management. Most recently, he was senior vice president and general manager of Wyle's Technical Support Services Business Unit, which operates and maintains laboratories and research facilities for government and industry clients such as the Department of Defense (DoD), National Aeronautics and Space Administration (NASA) and major prime contractors. Programs include support of the General Electric Turbine Engine test facility; Propellant Systems, Non-Destructive Evaluation and Metrology Laboratories at Cape Canaveral/Kennedy Space Center; and research centers at Wright-Patterson Air Force Base. He also directed the Research Instrumentation and Metrology Services contract for NASA Langley. Smith is a long-time supporter and member of ITEA, having served on the Board of Directors (1996–1999) and was re-elected in 2005. He also served on the DoD Defense Science Board, UXO/Land Mine Task Force (1996–1998). E-mail: drexel.smith@wylelabs.com

ITEA Journal 2008; 29: 45–50

Testing and Training 2020: From Stovepipes to Collaborative Enterprises

Jim Sebolka
The Paulus Institute, Washington, D.C.
David Grow
Headquarters, U.S. Army Training Support Systems, Washington, D.C.
Bo Tye
Developmental Test and Evaluation, Department of Defense, Washington, D.C.

This article presents approaches for overcoming the obstacles in the path to integrating Department of Defense (DoD) test and evaluation (T&E) and training communities to better support the modern-day warfighter and to enable new opportunities for shared investment, development, and process improvement. Testing and training, now managed under separate fiscal and managerial constructs, are hindered from establishing shared capabilities by distinctly different goals and funding. Each community must synchronize its priorities and funding with those of the other community to secure joint investments. A two-community perspective of the future path was expressed at the International Test and Evaluation Association (ITEA) Open Forum on Testing and Training. The proposed path to achieving integration of testing and training includes the establishment of a singular management backbone that encourages joint investment to eliminate duplication of effort and thus, systemically bring about cost reduction and enhance effectiveness across both communities. Such a shared backbone would enhance testing, training, interoperability, and warfighting through the increased commonality and realism of warfighting systems prior to fielding.

Key words: acquisition; combat readiness; cost reduction; interoperability; joint investment; testing and training.

Faced with a 30-year history of efforts to integrate defense testing and training, the International Test and Evaluation Association (ITEA) determined to bring forth the leaders of these two communities to address a path forward focused on success and documented findings. The result was the ITEA National Open Forum on Testing and Training hosted by the ITEA George Washington Chapter on October 3–4, 2007. Creative approaches, lessons learned, and community insights were also solicited for input. This article represents the collective findings from that effort, and stands ready to serve as the first touchstone on future efforts.

A major issue for the Department of Defense (DoD) is, "How does the DoD improve readiness and capability while cutting costs for training and test and evaluation (T&E) within the context of an overarching defense enterprise"? A corollary to this is, "How does DoD accomplish cost cutting while improving synergies in the areas of test and training based on principles that will survive from one set of leadership in DoD to another over time"?

It is the purpose of this article to assess where we have been, where we are, and where we should be in the year 2020. The authors aim to present approaches to overcome the issues blocking the path to synergize T&E and training for a greater good than that which can be achieved by each community acting alone and to empower defense leadership with enduring solutions for the defense enterprise.

Challenges

On the surface, the DoD has numerous capabilities and substantial funding to upgrade those assets. However, peeling back the onion reveals the challenges that exist within that structure. Combat readiness and technical evolutions present immediate demands upon the test and training infrastructures. DoD testing and training must therefore remain ready to execute any instruction at any time, recognizing that a lengthy planning, programming, budget, and execution process is the only path to new investment. This is the current sense by those communities. Therefore, each test or training facility within each Service must retain as much of its assets as possible to prepare for the next requirement, even as the costs required for maintenance rise with inflation and equipment age and the financial support for such maintenance diminishes. This paradigm must change. Examples of these challenges for change include:

- The Strategic Missile Defense test capabilities at Kwajalein Atoll have large-scale radar systems which exceed 35 years of age and operate using vacuum tubes no longer in production. Maintenance of these systems mandates the customized manufacture of these tubes or piecemeal replacement technologies at great expense, but at lesser expense than the wholesale replacement of the radar systems. Also, the Roi Namur large-scale radars at Kwajalein serve Army Space Command.
- Directed energy weapon systems add new technological challenges to the DoD. Speed-of-light weaponry requires specialized targets, instrumentation, and ranges to handle the direct effects of the weapon beam, recognizing that any error (even something as simple as a coffee cup in the path of the test beam) can lead to disastrous consequences as the weapon changes trajectory.
- Hypersonic and large footprint weapons have additional challenges to find sufficient airspace and range capacity for testing and training operations.
- Improvised explosive device (IED) defeat mechanisms have mandated the use of high power jammers in an environment encroached heavily by commercial spectrum use.
- Recognition of the individual warfighter as a key element of technology has led to new instrumentation requirements, which forces the addition of more weight and bulk onto overburdened training mission participants. As technology capability grows, so does the fielding to the individual warfighter, adding to the combat load even as the technology diminishes in size and weight.
- Incorporation of new aviation platforms mandates reexamination of airspace usage and monitoring. The F-22 is flying at higher altitudes than most combat aircraft, while unmanned aerial vehicles are flying lower. Thus, the definition of airspace is now requiring more accurate monitoring of simultaneous activities across the airspace.
- IED usage and world population shifts have led to a change in the warfighting spectrum. Urban canyons, multistory buildings, and close-in combat operations now must supplement the conventional force-on-force combat mechanisms in training and testing, without adding to the training or testing time horizons for completion. Peacekeeping and nation building efforts have also mandated new duties for military personnel that had not been part of the original training designs.

While the DoD attempts to keep costs as low as possible for its testing and training operations, such cost savings mandate retaining equipment and facilities that are unaffordable to replace, but also expensive to retain and maintain. Each Service, each community, and each functional capability must provide the resources and staffing to keep these capabilities available to support an ever-changing DoD mission profile. While an enterprise-wide approach would integrate these solutions for cost effective benefits, the current business model places defense test and training ranges under various management structures and financial oversight processes. Stovepiped approaches to these issues thus become institutionalized across the DoD to meet individualized requirements.

While these approaches provide near-term solutions (without waiting for the execution of the full budget process), they serve as a lightning rod for criticism from various analyses of the Department. Congress, the Base Realignment and Closure (BRAC) Commission, the Quadrennial Defense Review, and periodic audits pursue opportunities to save funding and reduce perceived excess capacity. Meanwhile, the Department has historically struggled to secure additional funding and lands to be used for test and training in preparation for the inevitable next war. Today, those efforts are validated as the Department fights the Global War on Terror, also known as "the long war."

In the training world, funding is focused on operations and maintenance (O&M), procurement, and military construction (MILCON) with a lesser amount on R&D. The reverse is true for testing. It is a question of proper balance, which neither community optimizes for the overall defense enterprise. *Table 1* shows distinctions between the testing and training missions and roles.

Today's testing community grew out of an acquisition environment that had fielded systems with substantial problems while acquiring new weapon systems to provide the warfighter an increased

Table 1. Comparison of testing and training cultures

Objective	Testing	Training
Community of interest	Acquisition	Operator (readiness)
Key concerns	Warfighter equipment	Warfighter operations
Key products	Materiel safety release	Warfighter readiness
	Material acceptance	Unit readiness
	Reliability certification	New equipment training
	Operational effectiveness	Increased warfighter effectiveness
	Suitability	
	Survivability	
Milestones	Production decisions	Combat requirements
	Fielding decisions	Unit readiness
Funding	Limited operations and maintenance (O&M)	O&M
	Limited procurement	Procurement
	Limited MILCON	MILCON
	Research and development (R&D)	Limited R&D

capability. Test and evaluation serves as what Secretary of Defense Perry called, "The conscience of acquisition" by providing a focused approach integrated into weapon system acquisition. The weapon system acquisition business model and its language sustain the testing community. Reimbursable range operations fund the T&E community workload with minimal institutional investment.

At the same time, the current training community construct arose from the readiness world, serving to prepare warfighters for combat. The warfighter focused approach, business model, and language keep the training community operating. Institutional funds provide the key support assets needed to keep warfighters in a state of readiness while expanding their capabilities to face combat challenges.

Thus, the two communities began as separate entities, and grew into distinct missions and cultures, united only by their support to the warfighter and few shared resources. Cost models, business enterprises, end objectives, and even the language of daily operations differ between them. All of these factors serve as obstacles to the Department's efforts to share resources to the benefit of the warfighter and the taxpayer.

Current situation

Currently, the test and training communities primarily attempt to resolve their individual challenges using community specific investments. This reinvestment approach generates community-wide savings but costs the DoD substantial resources when T&E systems typically are not applied to training applications, and vice-versa. Ultimately, the two communities established today's infrastructures which inherently inhibit the shared use of ranges, technologies, and mission space.

These policies and practices suffer from a lack of a strategic vision binding both communities and are divided by individual mission statements which are focused on acquisition or training thrusts rather than a unified thrust of victory in war. The emphasis needs to be focused on the warfighter as the ultimate customer rather than the missions of the two communities individually serving the warfighter.

Last year a policy letter was signed by the Under Secretary of Defense for Acquisition, Technology, and Logistics (AT&L), the Under Secretary of Defense Personnel and Readiness, and the Director of Operational Test and Evaluation (DOT&E) at the Office of the Secretary of Defense (OSD). This letter was sent to the three Service secretaries requesting their responses on how they would implement collaborative efforts between the two communities for activities requiring similar capabilities.

The BRAC Commission, viewing national assets to make more synergies for cost reduction, struggled to make simple definitions. Very little came out of its effort to drive the testing and training communities together.

There exists a demand for training as combat forces redeploy back to their home bases. However, limited site locations to conduct training (as well as test) exist. The demand on training is going to exceed that which is available.

Efforts to unify leadership (e.g., the Defense Test and Training Steering Group) have been thwarted by efforts to improve the two individual communities. The separate focus has led to an imbalance and diminished the stability of the shared testing and training environment. One of the problems is that warfighters view T&E as an encroachment into their critical domain, as T&E may force changes based on failures or safety risks within the inventory of military equipment. Sometimes this has been the case. By comparison, the training community sustains readiness.

Since T&E involvement in the training realm has defined ends and data requirements based on weapon systems, the warfighter perceives little benefit because the test functions are typically accomplished before the weapon system is widely distributed to the field, and is often perceived as delaying the receipt of the latest equipment into the field. Delays in receipt are tangible to the warfighter. Improvements in safety and effectiveness of an undelivered system are intangible. Thus, the warfighter perception is validated within his realm of awareness.

Today's growing financial and manpower constraints exacerbate the rice bowl syndrome. Program managers are being forced to concentrate their available resources on immediate requirements as opposed to contributing to long-term investments in the broad defense enterprise solution which will lead to more nearly global optimization. Consequently, today's existing incentives are often counterproductive to enterprise-wide solutions.

In the current environment, there have been numerous attempts to tie together test and training investments. Unfortunately, little has been achievable to translate early agreements and initiatives into meaningful long-term progress. One recent example was an initiative undertaken as a result of the OSD AT&L/P&R/DOT&E "Interdependency" memo. CRIIS, the "Common Range Integrated Instrumentation System," produced mixed results. After several months of negotiations, the closest the test and training communities could come to an interdependent agreement was to develop the hooks in CRIIS for "an open architecture system capable of supporting both missions." These hooks provide an "ability to grow the system over the next 5-10 years to meet training needs" as well as providing for the ability to "develop a radio capable of running Training's Range Instrumentation Waveform." Fiscal, mission requirements, and timing concerns of each community overshadowed the benefits recognized for the long term and thus sacrificed the future benefit for the current fiscal and business focus realities.

The objective situation for 2020

The ITEA Open Forum concluded with the participants bringing about a two-community perspective of the future path needed to establish an integrated testing and training operational basis. The following first two key points surfaced repeatedly:

Singular management. Testing and training cannot effectively merge common requirements and operations under the current management construct. A new paradigm needs to be created to establish a singular management approach across testing and training while securing the responsibilities of both the T&E and training communities. At the same time it is critical to take into consideration the Title X responsibilities of the Services. In OSD, T&E is divided into offices primarily covering operational test and evaluation (Director, Operational Test and Evaluation), resources (Test Resource Management Center), and developmental test and evaluation (Deputy Director, Developmental Test and Evaluation). By comparison, training falls under a single structure within the Under Secretary of Defense for Personnel and Readiness (USD P&R). These separate structures divide test and training objectives, plans, and funding, and further dissect the test community into operational and developmental focuses.

A singular management approach at OSD would swiftly enable the progress demanded for savings by the overall defense enterprise. This senior staff member should be at the DEPSECDEF or USD level to properly integrate the communities. An alternate is to have the director, Operational Test and Evaluation serve as the focal point. Whichever of these three options would be chosen, that individual needs the authority and resources — funding and manpower — required to properly execute the mission, an independent reporting system to the Secretary of Defense which promotes objectivity, and an enforcement system which promotes defense enterprise wide long-term solutions. As an interim step, the re-establishment of the Defense Test and Training Steering Group would help unite near-term coordination efforts between the communities.

This straightforward approach for change at OSD must take into consideration the multidimensional degrees of complexity involving the Services, joint commands, and program managers. Whatever recommendation is implemented at the OSD level must be mirrored swiftly within the Services and COCOMs to ensure that the streamlining and focus are made Defense-wide.

Incentives for shared investment. Singular management cannot succeed without the proper incentives to make it work. Testing and training, now managed under separate fiscal and managerial constructs, are hindered from establishing shared capabilities by distinctly different goals and timing. Each community must synchronize its priorities and funding with those of the other community to secure joint investments. Rarely do these priorities and funding opportunities completely intersect, leading to duplicative and stovepiped investments. Attempts to overcome these challenges lead to a situation where one community's high priority hinges on the other community's low priority, and thus unravel during the planning, programming, budget, and execution process. An incentive process rewarding efforts to link capabilities across testing and training is vital to future success. Within T&E, joint investments are encouraged by the use of the Central Test and Evaluation Investment Program (CTEIP). CTEIP provides funding for joint investments to encourage the Services to consider the needs of their sister Services and to share their future visions across the Department in building new test capabilities. This model, applied to joint test and training investments, would encourage similar sharing across these two communities. Further, this model would bridge today's two-community structure by encouraging the sharing of investments and requirements to develop singular solutions and capabilities.

Restructuring of investment processes into enterprise and service-based processes. Similar enterprise-level distributed capabilities like Real Time Casualty Assessment systems, backbone networks, fiber optic installations and maintenance, and standardized instrumentation systems, would enable additional savings on a Department level if provided by the DoD rather than left to be implemented according to individual Service requirements, schedules, and budgetary constraints. By contrast, individual instrumentation like sky screens, toxic fume detection, optical plume detection, and pressure sensors are best left to the current procurement construct. Individual requirement and budgetary processes are ill suited to the establishment of a corporate enterprise level capability within the DoD. Similarly, corporate enterprise investment cannot proactively address the requirements of near-term and Service-specific demands on testing and training. Corporate and individual investment need separate processes to achieve maximum benefit to the DoD, particularly if managed under a single oversight structure. Standardizing a backbone architecture will provide new and vital requirements at the Service level to invest in common and connective capabilities.

The Forum also identified the following major points:

- *Establish a shared, multilevel secure enterprise network for testing and training.* Currently, the training and testing communities struggle over when to use classified or unclassified versions of networks, instrumentation, and operations. These near-term savings are causing long-term detriment to the Department as individual solutions are being established, and integration opportunities are thus thwarted. A departmental decision to establish a singular standard for multi-level security across the communities would negate these problems while establishing a DoD standard for the virtual and live battlespaces. This would deliver long-term savings to the Services and to all joint operations involving the range infrastructure of the United States. This investment would also serve as the first corporate investment leading to a common electronic infrastructure across testing and training.
- *Shared, realistic joint battlespace.* Testers and trainers both seek to establish combat realism in their operations. Live, virtual, and constructive simulations have emerged as a vital tool to both communities, but continue to grow separately. Establishing a DoD initiative to provide a shared live, virtual, and constructive (LVC) architecture within the operating space of the multi-level security network above would immediately save resources and encourage shared investment across the communities. The Joint National Training Capability (JNTC) and Joint Mission Environment Test Capability (JMETC) both are taking key initial steps to make this effort a reality. But, they are chartered to perform other functions with the shared LVC environment as a byproduct of their efforts. A jointly managed capability, replete with networking standards and protocols, would ensure testers and trainers link to a common architecture. Ultimately, the communities would come together through their shared standards and investments in them.
- *Timing.* Today's national security environment mandates prompt attention to the issues above. Network Centric Warfare (NCW) is changing the test world to one in which the commander's decision making is a critical element of the test. Unlike past weapon systems, NCW systems enable the commander a multitude of choices in the solution to combat scenarios. No longer is the commander left to decide whether to fire a single gun, turn a single weapon system, or take other singular approaches. Instead, resources can be dedicated and rededicated in rapid succession. The commander's decision creates the ultimate pass-fail scenario for the weapon system. This enhanced capability has become inherent in the battlefield commander of the future as well as the warfighter commanding a singular weapon system such as the F-22 itself. Therefore, NCW is creating opportunities and critical needs pulling the training and the T&E communities together into a singular effort focused on supporting the warfighter.

Conclusion

T&E and training must converge to support the modern day warfighter to enable new opportunities for shared investment, development, and process improvement. Opportunities and models exist, such as the Defense Test and Training Steering Group and the CTEIP programs, which may be reapplied to this effort to secure immediate results. Longer term opportunities abound, but require hard choices for change in the managerial and fiscal models. This cannot be accomplished without the proper incentives. The DoD's investment in a shared LVC and multilevel secure backbone for testing and training can also be implemented in the near future to tie the communities together in ways never before realized.

The establishment of a singular backbone, and encouraging joint investment through a CTEIP-style model, will eliminate unnecessary duplication of effort. It will systemically bring about cost reduction and enhance effectiveness between the two communities in a proactive fashion. It will encourage creative solutions to problems for the warfighter. Such approaches will establish a new model for T&E and training that makes realism and instrumentation common across communities and Services. Ultimately, this test and training enterprise approach will enable greater realism for both communities at reduced overall cost. Lessons learned in the creation of the singular backbone could then be reapplied to a series of corporate investments that establish a universal digital battlespace for testing and training to secure further savings. Such a shared backbone would enhance testing, training, interoperability, and warfighting through the increased commonality and realism of Service systems prior to fielding. This increased commonality will also enhance "Joint Service" processes, by institutionalizing part of this shared framework during the testing and training phases of combat preparations. ❑

JIM SEBOLKA graduated from King's College with a bachelors of arts degree in mathematics and then volunteered to become a lay missionary. He has now completed the circle to work again with religious affairs as the vice president of The Paulus Institute with an emphasis on bringing together eastern and western churches through their liturgies. The intervening 40 years were spent in the USAF or working for DoD as a contractor. After obtaining a masters degree in industrial engineering from Texas A&M, his activities included serving in nine air campaigns during the Vietnam War. In Vietnam he participated in the planning and assessment of 400,000 combat sorties that included developing a quarterly risk factor for all of the combat missions flown and presenting a monthly briefing to the four star commanding general for the air war on combat aircraft lost. On the Air Staff, Sebolka served as the primary expert on developing and producing the world-wide air-to-ground conventional munitions requirements for USAF.

At the State Department, he provided assessments for the Comprehensive Nuclear Test Ban Treaty. As an adviser to the Thai government, his activities included writing and coordinating the Thai Operations Research Society charter for the Office of the Prime Minister and directing the definitive systems study for automation of logistics for the Thai Army. In Korea, he successfully led U.S. support to establish the Korean Institute of Defense Analysis as a world class think tank. In addition he was the contributing editor for the Korean Business Review of the Federation of Korean Industries, writing 80 percent of that journal. As military assistant to the NATO Adviser to the SECDEF, he initiated and led an international team to develop the methodology for munitions guidance for SHAPE.

As executive director for a joint office addressing the RDT&E of aircraft combat survivability, Sebolka's activities included selling the four Services to accept the OSD initiative for Joint Live Fire Testing and combining three different service program elements into one at OSD. He was one of the founders of the Survivability and Vulnerability Information and Analysis Center (SURVIAC). Subsequently, he provided support to the Live Fire Test and Evaluation (LFT&E) Office at OSD. He has also provided expert witness to Congress for changes to LFT&E legislation.

DAVID GROW has served as an instrumentation engineer, test director, dean of a test director college, product manager, and assistant project manager for the U.S. Army, and has served several details into the U.S. Navy, U.S. Air Force, and OSD to run financial programs for DoD. Grow is currently serving as the lead engineer and senior acquisition professional for Headquarters, Department of the Army Training Support Systems (DAMO-TRS) at the Pentagon. In that capacity, Grow is serving as the action officer coordinating the Training and Testing Interdependency Initiative (T2I2) effort for the Army's Director of Training, in coordination with the U.S. Army Test and Evaluation Command (ATEC).

COUNT BOYER "BO" TYE, JR. is a retired Air Force officer with over 26 years of experience in Air Force Special Operations, Test and Evaluation, and Acquisition. He served as the chief of special operations, Tactical Airlift and Trainer Division, Directorate of Global Reach Programs within the office of the Assistant Secretary of the Air Force for Acquisition. He also served on the Air Staff as the chief of policy and programs Division for Test and Evaluation. During his military career, Tye flew over 3,900 flying hours and flew 59 combat/combat support missions during operations URGENT FURY, DESERT STORM, and PROVIDE COMFORT. Tye currently supports the deputy director, Developmental Test and Evaluation (DT&E) for the Department of Defense. In this capacity, he develops and implements systems engineering and T&E policy to ensure Service programs are realistic, relevant, and in compliance with DoD and Congressional directives. His awards include the Legion of Merit, the Airman's Medal, and the Air Medal with oak leaf clusters.

ITEA Journal 2008; 29: 51–61

Evolving Enterprise Infrastructure for Model & Simulation-Based Testing of Net-Centric Systems

Steven Bridges
Joint Interoperability Test Command, Fort Huachuca, Arizona
Bernard P. Zeigler, Ph.D.
Arizona Center for Integrative Modeling and Simulation,
Electrical and Computer Engineering,
University of Arizona, Tucson, Arizona
James Nutaro, Ph.D.
Oak Ridge National Laboratory, Oak Ridge, Tennessee
Dane Hall, Tom Callaway, and Dale Fulton
Joint Interoperability Test Command, Fort Huachuca, Arizona

This article provides perspectives on how a test organization can organize and plan for enterprise-wide adoption of advances in emerging technologies and techniques, whether developed in-house or acquired from external sources. This article enumerates capabilities that greatly enhance a test organization's ability to support the impending testing demands from such GIG/SOA-based projects and presents an overarching strategic plan for integrating existing test technologies, identifying enterprise-wide technology gaps, and coordinating the development and acquisition of new test capabilities to greatly accelerate their readiness to meet impending net-centric testing challenges. The plan discussed in this article includes short-, medium-, and long-term horizon components to acquire or improve current test capabilities and offers a layered architecture that provides a framework for capability acquisition. Test organizations can incentivize their contractors to exploit the composability, reusability, and extensibility of technical attributes of SOA to support the development of the layered architecture. The authors conclude that the design of the test organization instrumentation and automation on top of the GIG/SOA infrastructure should be based on a model-driven software approach, systems-engineering modeling, and simulation principles and frameworks.

Key words: Global Information Grid (GIG), Service Oriented Architecture (SOA), net-centric testing, real-time interactivity, composability, reusability, extensibility, scalable.

Given Department of Defense (DoD) mandates for transition to net-centric operation, a test organization must acquire the ability to perform large-scale and fast-paced developmental and operational testing of Global Information Grid/Service Oriented Architecture (GIG/SOA)-based development projects. For example, the Joint Interoperability Test Command has the responsibility to test for GIG/SOA compliance for such projects as Net-Centric Enterprise Services and Net-Enabled Command Capability. A test organization's ability to support the impending testing demands from such GIG/SOA-based projects can be greatly enhanced by acquiring net-centric test capabilities. Although most test organizations already have the necessary capabilities to some extent, they can benefit from an overarching strategic plan for integrating existing test technologies, identifying enterprise-wide technology gaps, and coordinating the development and acquisition of new test capabilities to greatly accelerate their readiness to meet impending net-centric testing challenges.

Net-centric test capabilities

Several specific capabilities that a test organization must address to effectively conduct developmental and operational tests of net-centric systems are described below (Buchheister 2005, Carstairs 2005).

Composability

Composability is the capability to seamlessly compose the elements of the desired test environment by selecting and configuring live (e.g., human players, military systems) and/or virtual (digital representations of live components) versions of all test environment components. Test organizations can take advantage of the SOA and component styles that offer technical advantages for the composition of test instrumentation services and applications. Contractors should be incentivized to exploit the SOA constructs to build plug-and-play capabilities while meeting current and future needs.

Reusability and persistence

The test infrastructure persists over time and includes organized repositories to support the reuse of such elements as simulation models/digital representations, test development and implementation processes, and test experimentation components and tools (intelligent test agents, for example). This includes the capability to automatically store, catalog, and retrieve all information produced by any node on the network in a comprehensive, standard repository. A critical advantage of such repositories for the test organization is that they also help to avoid duplication of efforts by the test organization's multiple contractors.

Extensibility

The test infrastructure can be efficiently extended through the use of common architecture, interfaces, processes, and tools. Extensibility, composability, and reusability are mutually supportive attributes of model-driven software design methodology informed by engineering modeling and simulation fundamentals. The test organization must incentivize contractors to adopt such methodologies to achieve composability, reusability, and extensibility attributes in its developments.

Instrumented trustworthy measurement

Instrumented trustworthy measurement is the ability to instrument test environments in a manner that is principally nonintrusive and highly embedded, which provides real-time measures at the system and system-of-system (SoS) levels. Measurement is consistent and repeatable across experimental replications, providing reliable and trustworthy data. Specifically, instrumented trustworthy measurement includes the capability to

- Reproduce the test environment and play back segments of the test event in a manner that facilitates assessing the effects of modifying the experimental conditions with plug-and-play replaceable test components.
- Measure, compare, and evaluate experimentally specified architectural and parametric configurations of the system under test.
- Collect and segregate operational data (e.g., tactical and strategic data exchanged between systems under test) from test support data (e.g., instrumentation, simulation, analysis, and test control data).
- Seamlessly switch between real-time and after-test analysis of collected data.
- Perform the testing of net ready key performance parameters (NR-KPP) and compliance to the Net-Centric Reference Model for upcoming GIG/SOA and other net-centric developments.

Visibility and controllability

As net-centric systems under test become increasingly complex, the ability to visualize complex interactions and exert control over such interactions becomes increasingly vital for the test organization's ability to provide credible test results.

Real-time interactivity

Real-time interactivity includes visibility into events and processes through a display/representation of the test environment that is tailorable and provides accurate situational awareness of the test infrastructure and the tests that are underway. Currently, many test environments focus on relatively simple interactions and do not allow for highly complex many-on-many scenarios in which test environment components (networks, systems, and forces) react within a dynamic, closed-loop environment.

Features of advanced test organizations

The test organization should strive to be on the cutting edge of test organization capabilities, including

- Agility. Ability to automatically and adaptively monitor and manage selective functioning of the test infrastructures, test scenarios, networks, and systems and services under test.
- Automation. Ability to continually enhance the degree of automation of all the processes involved in defining, implementing, managing, reusing, and executing test events. This includes automated self-organizing recognition, initialization,

and control of plug-and-play test environment components.

- Scalability and Applicability to Full Life Cycle. Ability to scale the test infrastructure in terms of size, fidelity, and numbers of participants to accommodate the domains of systems engineering, development, development testing, operational testing, interoperability certification testing, and net-readiness and information assurance testing.
- GIG/SOA Integrated Robust Computer and Communication Infrastructure. Ability to provide high-performance computational support wherever needed in the configuration and execution of the test environment and the analysis of test data (in real time and after test). As the SoS and collaborations brought in by customers for testing become increasingly complex, the test organization will require increasingly powerful computing resources to manage all aspects of testing. The test organization will also require the ability to provide reliable, cost-effective, flexible, and GIG-enabled communication to all nodes.

(Note: Most of these requirements are not achievable with current manually based data collection and testing. Instrumentation and automation based on model-driven and systems-engineering modeling and simulation principles and frameworks are needed to meet these requirements.)

Proposed Acquisition Strategy

Acquiring all the assets needed for the above capabilities would significantly upgrade the test organization's capability for net-centric testing, but they will vary in degree of maturity. Some may be ready for implementation or purchase in the near term, and others may require significant investment in research and development. To help manage the acquisition of such assets, we propose an acquisition strategy having three levels corresponding to long-, medium-, and short-term planning horizons: (a) overall plan for test infrastructure evolution, (b) test infrastructure development to address test technology shortfalls, and (c) planning for individual test venues and events *(Figure 1)*. The underlying objective of the proposed strategy is to foster re-use of existing assets so as to maximize the cost-effectiveness of acquisition. The goal should be to set up a process for re-use, so that new capabilities are needed only when existing ones cannot be reasonably applied to the new situation.

Planning levels

Long-term planning

With respect to long-term planning, the objective is to look out past the horizon of imminent test events and current infrastructure improvement projects to identify emerging technologies and emerging system objectives and to lay out the broad approach to development of the test and evaluation infrastructure. As *Figure 2* illustrates, we suggest a planning approach to test individual customer projects and test events as part of the longer life cycle of the test infrastructure evolution. Key activities in the long-term strategic plan are as follows.

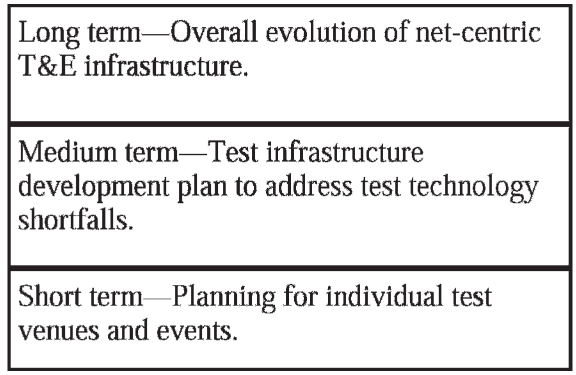


Figure 1. Net-centric testing planning levels

As new systems are defined and developed by a customer that will be subject to the test organization certification, the test organization must derive a coherent family of test objectives from the stated or to-be-developed system under test requirements and behavior specifications. Test events, venues, and infrastructure evolution must be synchronized with the customer system development schedule.

The high-level characteristics of the test development methodology and of the infrastructure to be used must be determined to meet the perceived complexity, volume, variety, and velocity of test challenges—with the objectives of furthering re-use of test resources and fostering cumulative knowledge management. This includes, among other things, establishing requirements for infrastructure development tools, such as formalizing and designing test models.

This long-term planning process passes technical shortfalls and their temporal attributes (e.g., "needed

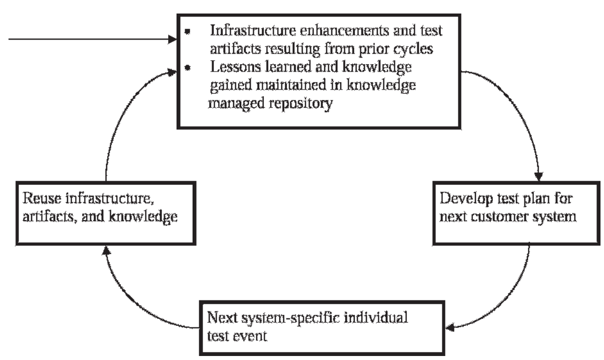


Figure 2. Long-term cycle of test activities

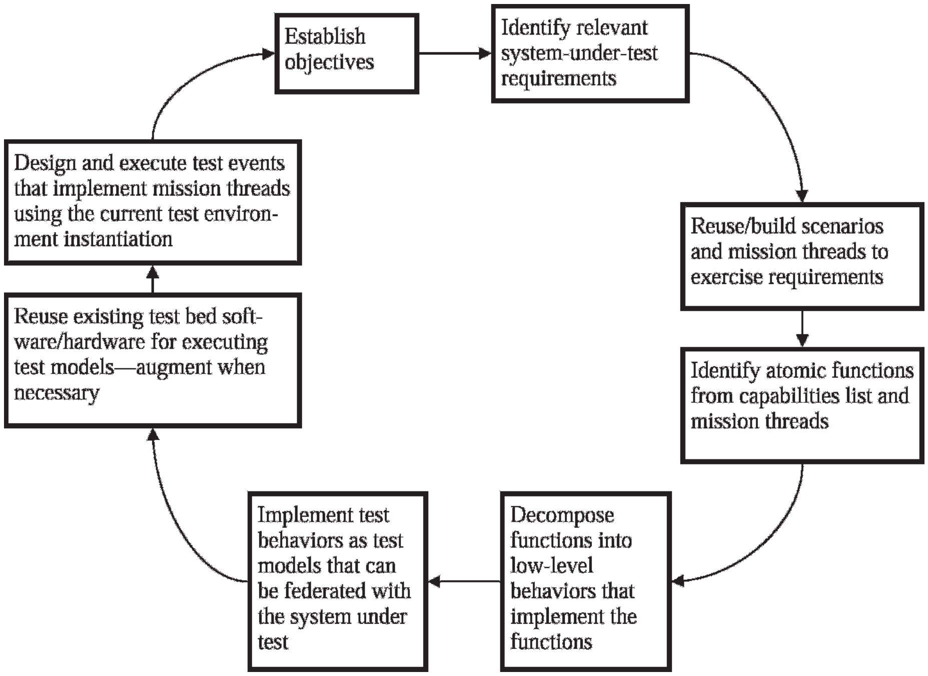


Figure 3. System-specific and individual event planning cycle

immediately," "needs can be foreseen for tests scheduled in the near future," or "is not critical now") on to medium-term planning.

Medium- and short-term planning

The planning for individual test venues and events consists of a cycle of activities that work within the structure established by the high-level planning. As *Figure 3* illustrates, this cycle consists of the following basic elements:

Establish objectives. The test objectives must provide an overview of the high-level system-specific test objectives and identify basic technical and operational evaluations that are needed to support future decision events. The objectives must

- Be tied to the system acquisition strategy.
- Establish the basis for a test and evaluation schedule in terms of test capabilities that will be available after each iteration of the test and evaluation process—this should include both anticipated costs and timelines. It is vital that the test organization and the customer agree to an integrated budget and timeline for each test objective.
- Be coordinated with the customer's strategy for system development and demonstration.
- Identify major strategic risks to achieving the identified test capabilities and lay out the activities necessary to mitigate the risks.
- Identify challenges, such as from complexity and need for testing that cannot be accomplished manually in sufficient volume, which must be overcome to effectively assess SoS and systems to contribute to their improvement. Update plans to meet these challenges.

Identify relevant test environment requirements. Once the test objectives are set, identify and evaluate specific test-support capabilities with respect to how they contribute to satisfying the test objectives. At this stage, a test environment description is constructed, which is tailored to the test objectives; relevant capabilities of the system under test are identified, and testable metrics are developed for those capabilities.

Reuse/build scenarios and mission threads to exercise given system under test requirements. The list of requirements for the system under test is linked to the underlying operational concepts and capabilities. With this list in hand, it is vital to develop specific mission threads that exercise these capabilities in a way that is relevant to the test objectives and anticipated operational environment.

Identify atomic functional units, decompose such functions into atomic behaviors, and implement test behaviors. The preceding three activities set the stage for technical development of the test environment. The technical development phase includes (a) identifying the atomic functional units of the system under test that comprise the identified capabilities, (b) decomposing these functional units into atomic testable behaviors, (c) combining these test behaviors as test models that can be compared with, and operated against, the system under test in the test environment. At this point, specific system under test components and/or subsystems are identified as being relevant to specific system capabilities in the context of identified mission threads, and the test machinery needed to stimulate and observe these components is ready to be put into place.

Build and/or reuse test bed software and hardware for executing test models; design and execute test events. Test events are planned to apply specific test bed items to the system under test. The test plan includes a test environment configuration for the test events, identifies the source of test data (e.g., live data, recorded system traces, simulations), and sets specific pass/fail criteria for the event. Acquire, build, and/or improve infrastructure development tools, such as tools for formalizing and designing test models.

This cycle of test activities defines an iterative process that allows for the evolution of each test phase as the system under test moves through its life cycle *(Figure 3)*. Throughout the cycle of test activities, there must be an emphasis on the reuse of proven, reliable, and efficient infrastructure elements and artifacts that were acquired as a result of earlier test projects. Efforts first capitalize on reusing existing software and hardware for executing test models. Of course, the requirements of each new project may exceed the capabilities of the current infrastructure and artifacts, in which case we seize opportunities to enhance the infrastructure. Thus, each specific system under test feeds back lessons learned and contributes to long-term capabilities and knowledge. This feedback loop is illustrated in *Figure 2*.

Proposed layered architecture

To support the acquisition of net-centric testing capability with the time horizons just discussed, we offer a layered architecture that provides a framework for such capability acquisition. We propose that the test organization develop an overall architecture for net-centric instrumentation as illustrated in *Figure 4.* The architecture is based on that presented in Sarjoughian, Xiegler, and Hall 2001 and refers to background in literature on modeling and simulation (Zeigler, Fulton, Hammonds, and Nutaro 2005; Zeigler, Kim, Praehofer 2000; Zeigler and Hammonds 2007; Traore and Muxy 2004); Systems of Systems (Sage 2007; Wymore 1992; Wymore 1967; Morganwalp and Sage 2004); model-driven software development (Dimario 2007; Dimario 2006; Object Modeling Group 2007; Jacobs 2004; Wagenhals, Haider, and Levis 2002; Wegmann 2002); and integrated simulation-based development and testing (Mak, Mittal, and Hwang [in press]; Mittal 2006; Mittal, Mak, and Nutaro 2006; Mittal 2007; Mittal, Sahin, and Jamshidi [in press]).

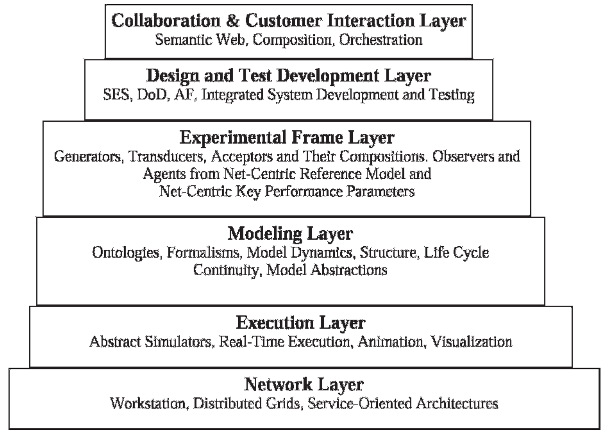


Figure 4. Architecture for net-centric test instrumentation

Network layer

The network layer contains the actual computers (including workstations and high performance systems) and the connecting networks (both local area network and wide area network, their hardware and software).

Execution layer

The execution layer is the software that executes the models in simulation time and/or real time to generate their behavior. Included in this layer are the protocols that provide the basis for distributed simulation (such as those that are standardized in the high level architecture). Also included are database management systems and software for controlling simulation executions and for displaying test results and animated visuals of the behaviors generated.

Modeling layer

The modeling layer supports the development of simulation models and other digital representations for net-centric testing in formalisms that are independent of execution layer implementations. At this layer, the test organization would compose services and applications. Also in this layer is support for the quality control of model acquisition, especially the key processes of verification and validation of models, simulators, and test tools.

Experimental frame layer

The experimental frame layer employs the artifacts and services of the modeling layer to develop test components, such as generators, acceptors, and transducers and their compositions, to provide test instrumentation services. Included are the observers and agents that run in the execution layer, and that interface with the systems and services under test to connect them to the experimental frame components. Also included are means to capture relevant measures of performance and effectiveness and instrument them as experimental frame compositions employing modeling layer and execution layer services. These measures are critical to the testing of NR-KPPs that the test organization must be able to accomplish.

Design and test development layer

The design and test development layer supports the ingestion and analysis of model-based system specification documents, such as in the DoD Architecture Framework, where the design is based on specifying desired behaviors through models and implementing these behaviors through interconnection of system components. In the modeling layer, results of this analysis of system behavior requirements will be used with automated generation of test models, which when deployed in the execution layer as automated test cases will interact with systems and services under test. The design and test development layer also includes maintenance and configuration support for large families of alternative test architectures, whether in the form of spaces set up by parameters or more powerful means of specifying alternative model structures such as provided by the System Entity Structure (SES) methodology. Artificial intelligence and simulated natural intelligence (evolutionary programming) may be brought in to help deal with combinatorial explosions occasioned by analysis for test development.

Collaboration and customer interaction layer

The collaboration and customer interaction layer enables people and/or intelligent agents to manage and control the infrastructure capabilities supplied by underlying layers. This includes interactions with the customer in which test results are conveyed and explained if needed.

Note that these layers describe functionalities that can be partially supplied by proven and reliable legacy tools in the test organization's inventory from earlier developments. However, the primary objective of such architecture is to facilitate carrying out the multi-horizon planning approach discussed earlier. As customer projects arrive, their testing requirements can be referenced to the elements within the layered architecture—the detailed test assets at the various levels are called out. Missing assets can be the cues to start an acquisition process to fill the gap. *Figure 6* illustrates the application of the layered architecture to sensor simulation infrastructure acquisition.

Artifacts, such as models and test and evaluation are results of processes (systems) that must not only have hardware and software support but must be done by competent people using competent methods in an environment that fosters each process. Indeed, to be

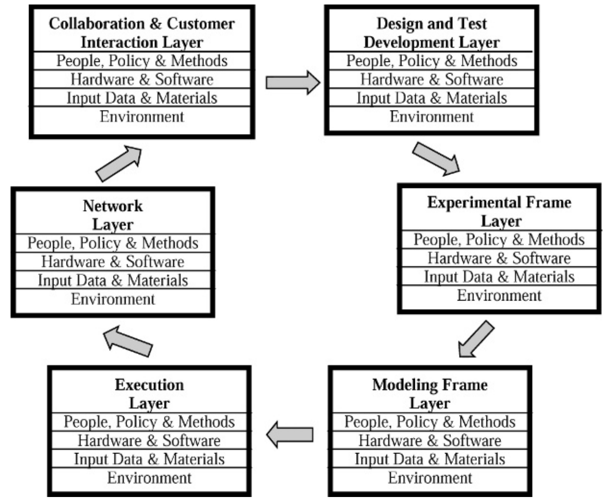


Figure 5. The layered architecture viewed from the DOTMLPF perspective

effective, there must be collaboration among layers and continuity of people, methods, software and hardware, good input and materials, and a supportive environment (e.g., from management and external networks). This collaboration is illustrated in *Figure 5*, employing the basic categories of People, Policy and Methods, Hardware and Software, Input Data and Materials, and Environment; expressing the areas DoD often refers to as DOTMLPF—doctrine, organization, training, materiel, leadership, personnel, facilities. To better communicate the main collaboration path, connections for exception handling and additional feedback have not been included in *Figure 5*. We recognize that a real-world portrayal of the collaboration would include numerous iterations, feedback, and exception handling.

Table 1 suggests how some of the identified layers can be further elaborated in terms of representative needs that must be met in the basic categories that are most pertinent to each layer.

We note that the table makes clear that besides the acquisition and application of test infrastructure elements, the Joint Interoperability Test Command (JITC) must plan for acquiring the right personnel and instituting the right organization. Specifically, JITC must develop a culture that will facilitate the interactions among personnel that are critical for the enterprise to be effective.

Mapping shortfalls to architectural layers

The proposed layered architecture will provide a framework for focusing the planning and acquisition of the test infrastructure capability. With the Xs in the cells of *Table 2* we offer a mapping to the shortfall areas that we think are best addressed in each layer.

Table 1. Illustrating the layered architecture in relation to doctrine, organization, training, materiel, leadership, personnel, and facilities (DOTMLPF)

Layer	People, Policy, and Methods	Hardware and Software	Input Data and Materials	Environment
Experimental Frame Layer	Experimental Frame Developers (1) are qualified, (2) have methodologies that are appropriate and effective, (3) have shared awareness of development plans, design decisions, and progress, and (4) have good access to model developers and to test development personnel who are prepared to clarify requirements and standards governing the systems under test.	(1) Access to relevant models and software to gather required measures (MOEs, MOPs), generate required stimuli and loads, and control. (2) Model development tools and software integrated design environments are adequate. (3) Access to JITC network and to test workstations.	(1) V&V experimental frame artifacts and test components from the Modeling Layer. (2) V&Ved data for DT, &V, V&T. (3) Good requirements and/or standards. (4) V&Ved means to capture relevant measures.	(1) Development, testing, and V&V are managed to plan. (2) Proper SW CM environment and practice.
Design and Test Development Layer	Design and Test Developers (1) are qualified, (2) have methodologies that are appropriate and effective, (3) have shared awareness with the JITC team, and (4) have good access to personnel who are prepared to clarify requirements and standards governing the systems under test.	Adequate tools to capture and characterize systems under test behaviors and interfaces.	(1) Adequate system specification documents and DoDAF documents, (2) Behavior requirements and/or standards are sufficiently well-specified. This applies particularly to GIG/SOA-based developments (e.g., NCES, NECC).	(1) Unplanned requirement additions are avoided. (2) Proper CM environment and practice.

The test organization should employ this architecture as the basis for its net-centric instrumentation plan.

Strategies for net-centric instrumentation planning

With the layered architecture as basis, the test organization can develop specific strategies that take into account long-, medium-, and short-term considerations for orderly acquisition of effective and reusable infrastructure. One alternative is to continue to rely on legacy tools while employing the architecture to plan for new tool acquisitions as the opportunities present themselves. Another alternative is to invest immediately in high priority tool developments that are compliant to

Table 2. Illustrating the mapping of shortfalls in architectural layers

	Layers					
	Network	Execution	Modeling	Experimental frame	Design and text development	Collaboration and customer interaction
Composability			X	X	X	
Reuseability and persistence	X		X	X	X	
Extensibility			X	X	X	
Instrumented trustworthy measurement				X		
Visibility and controllability	X	X				X
Real-time interactivity		X		X		
Agility					X	X
Automation	X	X	X	X	X	X
Scalability and applicability to full life cycle	X	X			X	X
GIG/SOA integrated robust computer and communication infrastructure	X	X	X	X	X	X

GIG/SOA, Global Information Grid/Service Oriented Architecture.

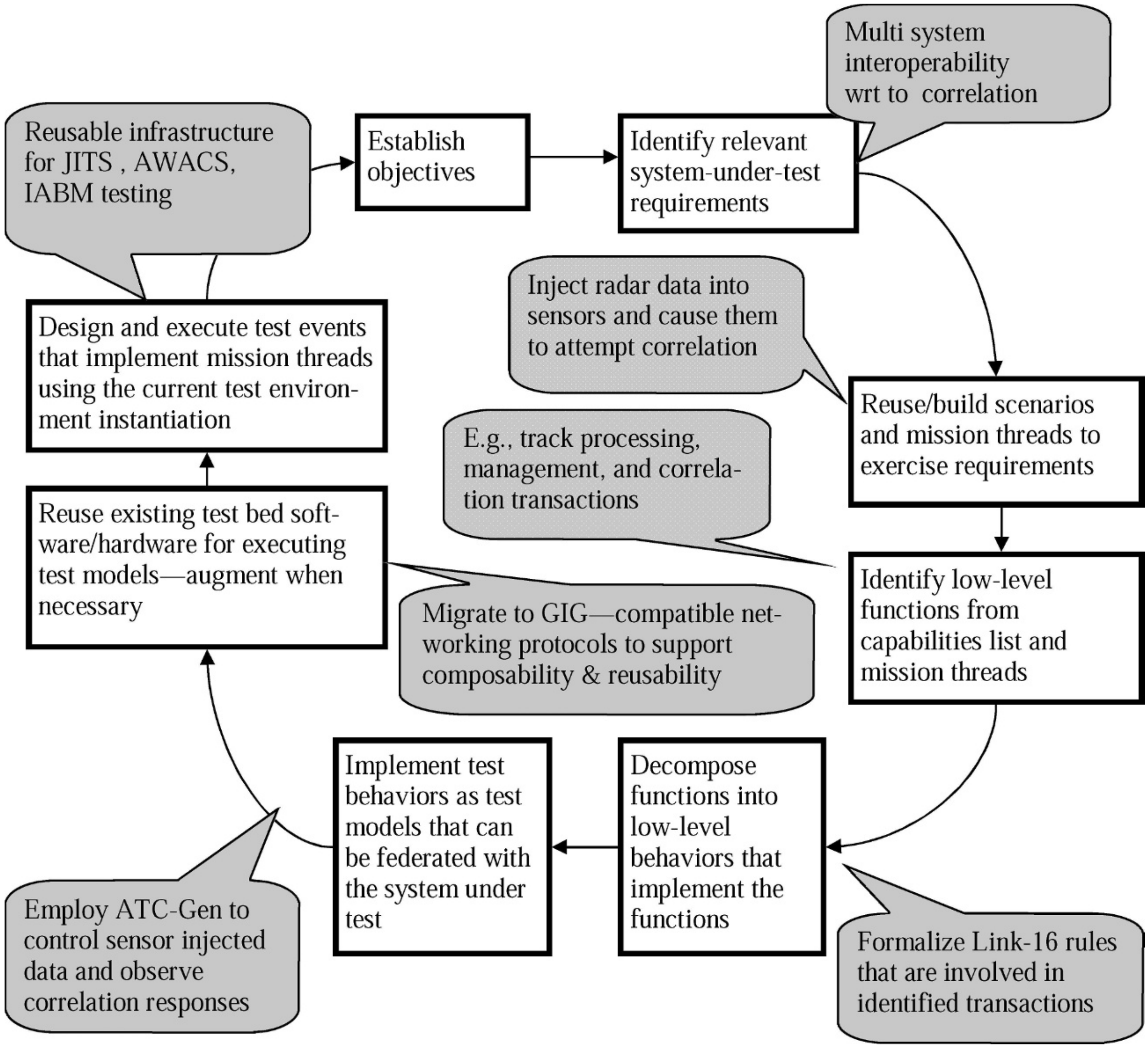


Figure 6. Illustrating event planning cycle for sensor simulation acquisition

such architecture and that implement nonexistent capabilities such as planning or automated testing and may not replace legacy tools in the near term.

Illustrative application to sensor simulation infrastructure acquisition

Figure 6 sketches how the planning cycle of *Figure 2* might apply to the acquisition of sensor simulation for net-centric testing. The perspectives offered by multi-horizon planning and layered test infrastructure architecture are intended to facilitate developing and evaluating acquisition strategies. By themselves, they do not decide the choices to make.

Summary and recommendations

A test organization needs an instrumentation development and maintenance system that can be considered an open subsystem of an open system—the test organization, test evaluation, and certification system, which produces results as shown on the left side of *Figure 7*. Shown on the left are the resources and funds leaving the system, and on the right are the funds and resources coming in. In addition, entering at the right is a seemingly high volume of a broad variety of not always clear or fixed system-under-design requirements, protocols, waveforms, standards, and mandated architectural styles (e.g., net-centric reference model and SOA). As shown at the bottom right, the test organization must encourage scientific research and technology development projects of the government, academia, and industry to develop methods and technologies needed to fill test capability gaps.

The specific inclusion of infrastructure development as an integral part of the top-down approach fosters

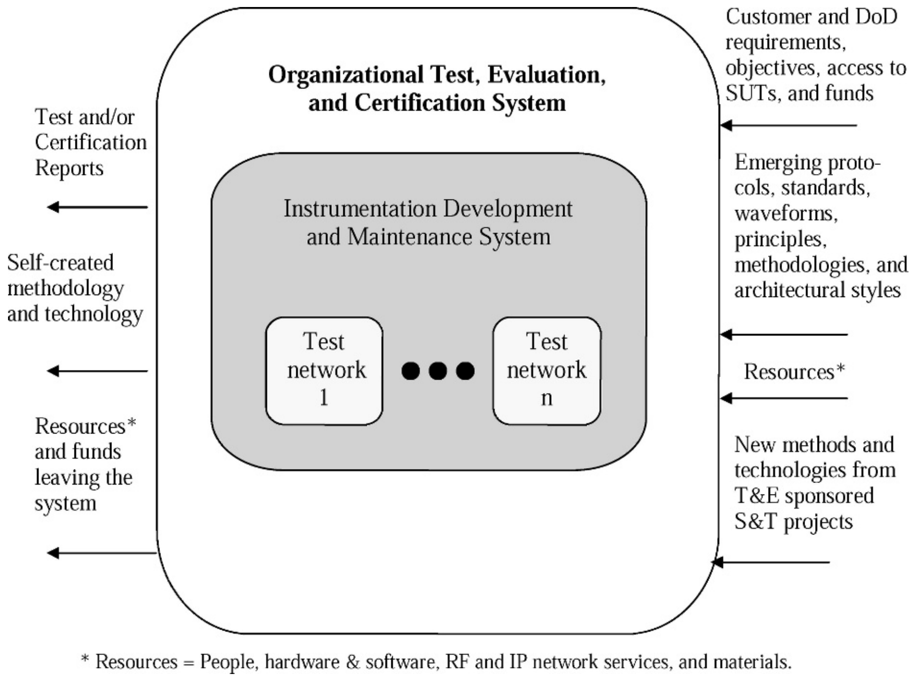


Figure 7. Instrumentation development and maintenance subsystem of the test organization test and evaluation and certification system

significant reuse of test resources and cumulative knowledge management of the products of testing. We recommend that in addition to basic test development, each iteration of the individual test event/venue planning cycle should *also target a small, well-defined, and incremental enhancement of the test environment functionality that we implement as components of the overall test infrastructure*. Iterations should refine and/or enhance test objectives and develop and/or modify the test bed technology as needed; and test events should realize these test objectives using the available test bed capabilities. In addition to supporting the planned test objectives, each iteration should to the extent possible include a test event that specifically demonstrates the new test environment functionality.

Testing in this paradigm is objective driven rather than event driven (i.e., test events must be traceable back to established test objectives). In most cases, major shortfalls of test technology should be identified early, either during the refinement/expansion of test objectives, or in the early phases of test event planning. Interim technology solutions to reduce shortfalls that are identified late in test event planning or even later during test event execution should be considered tentative pending review in the next iteration of the test bed development. These interim solutions should be the exception and not the rule.

We recognize that infrastructure development requires competent people using competent methods in an environment that fosters the development of each process and artifact. In this regard, we recommend including in the test organization team a test-infrastructure development component that supports testing for each customer project and its test events. The responsibilities of this infrastructure team would be to

- Identify existing, reusable testing tools and requirements that are common across test activities for use and for potential adaptation or conversion to a reusable component.
- Build and maintain reusable technical components of a common test infrastructure.
- Promote test asset reuse where appropriate.
- Advise test event planning and execution when the events rely on pieces of the common test infrastructure.
- Retain and disseminate lessons learned from a test event.

In addition to the net-centric test infrastructure components involved in specific customer projects, the test organization should stand up a global test infrastructure development team to operate within the larger framework of its enterprise level plans for coordinating instrumentation, automation, and architecture support across all the test organization portfolios. This team would

- Coordinate efforts for customer-specific developments with the test organization's enterprise level net-centric test infrastructure development and identify overlapping concerns and/or testing tools. Customer-specific testing requirements can be referenced to the elements within the layered architecture, calling out detailed test assets at the various levels. Missing assets can be the cues to start acquisitions.

- Provide *proactive* technical solutions to identified customer-specific test requirements. These solutions will be incorporated into test events that will be planned in detail later on in the test and evaluation process.
- Seek out and recommend best practices and cultural innovations that will facilitate effective coordination of the personnel working at the various architectural layers as customer projects arrive.
- Participate actively in teams responsible for test planning and developing test tools for specific events. Successful reuse requires positive involvement at all levels of the organization. Consequently, persons responsible for long-term infrastructure development must be constructively and actively engaged with the elements of the organization that they support. ❑

Steven Bridges is the chief engineer for the Joint Interoperability Test Command (JITC). In this capacity, he is responsible for the oversight of technical aspects of all JITC test programs, development of new test methods for the net-centric environment, and modernization of both test infrastructure and instrumentation. He serves as adviser to the JITC Commander on all technical issues and has worked on government technical acquistion projects for over 35 years. He received his bachelor of science in electrical engineering from Texas A&M, Kingsville in 1972 and was a recipient of the Defense Information Systems Agency (DISA) Director's Award for Achievement in a Scientific/Engineering Field. E-mail: steve.bridges@disa.mil

Bernard P. Zeigler is professor of electrical and computer engineering at the University of Arizona, Tucson, and director of the Arizona Center for Integrative Modeling and Simulation. He is internationally known for his 1976 foundational text Theory of Modeling and Simulation, recently revised for a second edition (Zeigler 2000). He has published numerous books and research publications on the discrete event system specification (DEVS) formalism. In 1995, he was named fellow of the IEEE in recognition of his contributions to the theory of discrete event simulation. In 2000 he received the McLeod Founder's Award from the Society for Computer Simulation, its highest recognition, for his contributions to discrete event simulation. He received the JITC Golden Eagle Award for research and development of the Automated Test Case Generator, 2005, and the Award for Best M&S Development in the Cross-functional Area, 2004/2005, by the National Training Simulation Association, May 2, 2006. E-mail: zeigler@ece.arizona.edu

James Nutaro is with the Modeling and Simulation group at the Oak Ridge National Laboratory in Oak Ridge, Tennessee. He obtained his Ph.D. at the University of Arizona with a dissertation entitled Parallel Discrete Event Simulation with Applications to Continuous Systems, and he has published papers describing his research in that area. He was a post-doctoral researcher at the Arizona Center for Integrative Modeling and Simulation. He has several years of employment experience developing and applying distributed simulation software in a defense systems context. E-mail: nutarojj@ornl.gov

Dane Hall is a certified systems engineering professional at JITC, Fort Huachuca, Arizona, who has 18 years experience as a system engineer working with acquisition programs of the DoD and other agencies. For 20 years he was an aeronautical engineering duty (maintenance) officer in the U.S. Navy. He has a master of science degree in systems management from the University of Southern California, Los Angeles, and a bachelor of science in aeronautical and astronautical engineering from Purdue University, Lafayette, Indiana. E-mail: dane.hall.ctr@disa.mil

Tom Callaway is president and principal engineer of Callaway Engineering Services Incorporated. He has worked as a program manager and communications test engineer for over 35 years for commercial and government technical acquisition projects. For over 20 years he was an Army signal corps officer. He has a bachelor of science degree in engineering from the U.S. Military Academy and a master of science degree in electrical engineering from the University of Arizona. E-mail: tom.callaway.ctr@disa.mil

Dale Fulton currently works at JITC, Fort Huachuca, Arizona, and has worked as an instrumentation designer and tester for 28 years in the DoD, Department of Energy, and commercial communication and data acquisition systems. His experience in test automation includes real-time stimulation and simulation for technical and operational test environments, and data acquisition and time-referencing in laboratory and field systems. He has a bachelor of science degree from the University of Arizona, Tucson. E-mail: dale.fulton.ctr@disa.mil

References

Buchheister, J. "Net-centric Test & Evaluation," Command and Control Research and Technology Symposium: The Power of Information Age Concepts and Technologies, 2004. Available at: http://www.dodccrp.org/events/2004/CCRTS_San_Diego/CD/papers/208.pdf. Accessed October 2005.

Carstairs, D. 2005. "Wanted: A New Test Approach for Military Net-Centric Operations." *The ITEA Journal of Test and Evaluation*. Volume 26, No. 3.

DiMario, M. "System of Systems Interoperability Types and Characteristics in Joint Command and Control." *Proceedings of the 2006 IEEE/SMC International Conference on System of Systems Engineering*. Los Angeles, California. April 2006.

DiMario, M. 2007. "SoSE Discussion Panel Introduction, From Systems Engineering to System of Systems Engineering." *2007 IEEE International Conference on System of Systems Engineering (SoSE)*. April 16–18, 2007, San Antonio, Texas.

Jacobs, R. 2004. "Model-Driven Development of Command and Control Capabilities for Joint and Coalition Warfare." *Command and Control Research and Technology Symposium*. June 2004. Available at http://www.dodccrp.org/events/9th/ICCRTS/CD/Papers/169.pdf. Accessed January 2008.

Mak, E., Mittal, S. and Hwang, M. 2008. "Automating Link-16 Testing Using DEVS and XML." *Journal of Defense Modeling and Simulation*. (Draft).

Mittal, S. 2006. "Extending DoDAF to Allow DEVS-Based Modeling and Simulation, Special issue on DoDAF." *Journal of Defense Modeling and Simulation*. 3(2).

Mittal, S. 2007. "DEVS Unified Process for Integrated Development and Testing of Service Oriented Architectures." University of Arizona Ph.D. dissertation.

Mittal, S., Mak, E. and Nutaro, J. 2006. "DEVS-Based Dynamic Modeling & Simulation Reconfiguration Using Enhanced DoDAF Design Process." Special issue on DoDAF. *Journal of Defense Modeling and Simulation*. III(4).

Mittal, S., Zeigler, B., Martin, J., Sahin, F. and Jamshidi, M. 2008. "Modeling and Simulation for Systems of Systems Engineering." *Systems of Systems Innovations for the 21st Century*. Edited by M. Jamshidi. John Wiley & Sons: New York (Draft).

Morganwalp, J. and Sage, A. 2004. "Enterprise Architecture Measures of Effectiveness." *International Journal of Technology, Policy and Management*. 4(1), pp. 81–94.

Object Modeling Group (OMG). 2007. http://www.omg.org.

Sage, A. 2007. "A System of Engineering an Integrated System Family." From Systems Engineering to System of Systems Engineering. *2007 IEEE International Conference on System of Systems Engineering (SoSE)*. April 16–18, 2007, San Antonio, Texas.

Sarjoughian, H., Zeigler, B. and Hall, S. 2001. "A Layered Modeling and Simulation Architecture for Agent-Based System Development." *Proceedings of the IEEE 89*. (2), pp. 201–213.

Traore, M. and Muxy, A. 2004. "Capturing the Dual Relationship Between Simulation Models and Their Context." *SIMPRA (Simulation Practice and Theory)* Elsevier: London.

Wagenhals, L., Haider, S. and Levis, A. 2002. "Synthesizing Executable Models of Object Oriented Architectures." *Workshop on Formal Methods Applied to Defense Systems*. Adelaide, Australia, June 2002.

Wegmann, A. 2002. "Strengthening MDA by Drawing from the Living Systems Theory. Workshop in Software Model Engineering."

Wymore, W. 1967. *A Mathematical Theory of Systems Engineering: The Elements.* John Wiley & Sons: New York.

Wymore, W., Chapman, W. and Bahill, A. T. 1992. *Engineering Modeling and Design* CRC Press: Boca Raton, Florida.

Zeigler, B., Fulton, D., Hammonds, P. and Nutaro, J. 2005. "Framework for M&S–Based System Development and Testing in a Net-Centric Environment." *ITEA Journal of Test and Evaluation*. 26(3), pp. 21–34, 20.

Zeigler, B. and Hammonds, P. 2007. *Modeling & Simulation-Based Data Engineering: Introducing Pragmatics into Ontologies for Net-Centric Information Exchange.* Academic Press: New York.

Zeigler, B., Kim, T. and Praehofer, H. 2000. *Theory of Modeling and Simulation.* Academic Press: New York.

ITEA Journal 2008; 29: 63–66

Towards Better Control of Information Assurance Assessments in Exercise Settings

David J. Aland
Wyle, Arlington, Virginia

By the adoption of certain limited techniques, the assessment of Information Assurance in both acquisition and fielded systems can achieve a higher level of rigor than available using current methods. These techniques do not replace traditional Blue/Red team activities but are used to augment them and provide a means by which replicable data may be recorded and analyzed without raising the level of risk to the exercise planner.

Key words: Exercise planning; network assessment; network penetration & exploitation; network protection; network vulnerability; risk; training.

The testing of Information Assurance (IA) in Department of Defense (DoD) information systems is addressed at numerous points throughout the life cycle of these systems, for the most part in the development and acquisition process. In 2002, a Congressional mandate added a requirement for post-fielding assessments of DoD networks. These assessments were to be accomplished during major exercises, a shared environment often familiar to the operational testing community. But this additional venue also created a challenge for both assessment and exercise planners—how to best integrate network evaluations into highly complex training events that depend upon the network that is also being evaluated. This would necessitate the integration of both the training events and assessment events, and a deeper level of synchronization between the two.

There are three key goals to such a process: (a) make the best possible use of the existing IA assessment capabilities; (b) provide meaningful and nondisruptive training in a warfare area (Information Operations) that had previously received little attention; and (c) structure events to gather meaningful observations and data regarding effectiveness of IA systems, practices, and policies. In order to accomplish this, it is necessary to design exercise events that emphasize the various aspects of IA in a manner that adds value to the training exercises and is consistent with the skills and expertise of the teams from the agencies that normally conduct DoD network assessments. This also requires IA teams to adhere in some degree to scripted events and timelines. In addition, it requires exercise planners to place greater emphasis on IA events, an area which is only now growing in prominence in most exercise scenarios. For the IA teams, this means greater constraints, and for the exercise planners, greater risk. For the operational evaluator, this could only mean many more variables in the shared testing environment.

Assessment process

Inherent to the DoD IA assessment process is the use of traditional DoD IA teams: Blue teams (technical and nontechnical vulnerability audits) and Red teams (technical adversarial penetration and exploitation tests). The missions of these teams differ, despite the common focus. The Blue team assessment most frequently consists of a collaborative review of technical and administrative support to a system or network, often including the use of scanning tools, password crackers, and low-intensity penetration tests. The goal of a Blue team assessment is to identify and document vulnerabilities caused by configuration, process, or management shortfalls. Conversely, a Red team assessment is usually a limited-duration "attack"—a network-based adversary, operating within some preset limitations, which attempts to find and exploit at least one area of vulnerability to gain internal access to a network or system. In many cases, such an attack will be accompanied by modest exploitation of that access, usually in the form of data exfiltration or modification, in order to demonstrate the operational impact of the vulnerability exploited. For these reasons, as well as others (including technical limitations, operational considerations, and resources), the Blue team activities could be described as being "a mile wide, but an inch deep," whereas the Red team activities would be "a mile deep, but an inch wide." The differing focus of each team provides very different products.

Information Assurance is normally described as consisting of four fundamental tasks or principles: protect, detect, react, and restore. Due to the fundamental character of the established DoD IA controls (DoD Instruction 8500.2), the focus of most DoD IA assessments (and pre-acquisition testing) is on network protection, with limited insight or investigation into network detection, reaction, and restoration capabilities. Most Blue team events similarly focus on protection, with some view to detection. Red team assessments also focus on protection (through penetration and exploitation events) but can allow greater assessment of the other three tasks, if structured to do so. However, because of the limitations most often imposed on Red team events (whether technical, operational, or resource), the detect, react, and restore functions are not often examined in any depth, nor in a reproducible fashion. The "traditional" modus operandi of most Red teams is to find and exploit a single vulnerability, making comparison of one event to another relatively difficult, with only a few common characteristics. Employment of wider testing can significantly expand the cost, in both time and resources, of any given Red team event, making such an expansion typically impractical. Furthermore, such an expansion may be contrary to the interests of the exercise planner, as they may increase risk to other training objectives.

Overcoming obstacles

The agencies that sponsor Blue and Red teams are experiencing a growing demand for their services, as the number of critical mission functions migrating into automated information systems grows. Working within limited budgets, and facing a long lead-time for the development, training, and employment of skilled operators, the Blue and Red teams cannot practically expand the scope of their assessments without having to reduce the quantity of assessments they can perform. Given the limits in funding and manpower, one possible solution would be to establish means by which these assessments can provide greater depth of assessment without requiring additional time, personnel, or other resources.

For the "customer"—that is, the unit being assessed or sponsoring the assessment—a very robust IA assessment can potentially derail other testing or training objectives, and for that reason, most Blue and Red teams must operate within a series of constraints or written ground rules established in advance of the event. These ground rules serve to protect critical training events from disruption and yet create de facto limits on the scope and quality of the IA assessment. Most unit commanders would be reluctant to expand the scope of IA events without some form of assurance that critical functions or events would not be impaired.

For both reasons, the assessment planner is faced with limitations that all too often render the assessment findings for any one event essentially unique—a product of the variable selection of limitations imposed by both the assessment agency and the one being assessed. In order to widen the available data for analysis, trending, and long-term issue identification, the evaluator working in this shared environment requires a better form of controlled metrics and conditions but often has the least influence over the environment itself. From an operational test and evaluation standpoint, this is a considerable obstacle: conducting an assessment in an environment that is not controlled by the assessor, using resources that are, to a greater or lesser extent, also not controlled by the assessor.

A better way

The needs of all three stakeholders—the Blue/Red teams, the assessed unit, and the operational assessor—can be met by the application of a common solution: establishment of a set of core events that are more closely controlled but do not raise the cost of conducting an assessment, and that do not increase risk but do improve the consistency of the data gathered.

In order to do so, these events must: (a) leverage tasks already being performed (or that can be performed) by the Blue and Red teams; (b) maintain or decrease the level of risk currently available through existing limitations; and (c) be sufficiently consistent that they can be performed repeatedly, and in the same manner, during a variety of assessments of systems, networks, and locations. This may require all three parties to make adjustments to their current processes, but these adjustments are relatively small, particularly in view of the gains to be realized.

The implementation of more controlled test events must make use of the highly developed skills of Blue and Red teams in achieving system or network penetration, and exploiting those penetrations; demonstrate the operational/training risk such penetrations and exploitations produce without actually incurring any significant risk; and provide a consistent set of tests that can be repeated and compared in subsequent evaluations and assessments. The main attribute in achieving all three goals is control.

Such control can be achieved in a number of ways: (a) by establishing alternative, but equally fixed, boundaries for test events; (b) by conducting tests against non-operational entities; (c) by applying precise amounts of force/stimulation during tests; (d) by segregating tests into discrete events or phases; or (e) by limited automation.

Examples of the kinds of controlled test events that might meet these conditions include

- Mission-focused assessments (alternative limits). Assessment plans are designed around one or more specific mission areas and are limited to impacting those missions and the network components supporting the missions designated. For the purposes of an IA assessment, risk would be limited largely to the system or systems targeted, and the assessment focuses on determining the impact to the designated mission supported by the targeted systems. This method would also allow extrapolation from prior acquisition testing into the broader testing of systems in their intended operational environments while limiting "spillover" effects into other systems or portions of the network.
- Repetitive vector assessments (alternative limits/precise force/segregation). Assessment team activities are organized as a series of repeated events, with each event specifically focused on testing a discrete segment of a system/network, or functional attribute. Such events can be conducted as multiple attacks along a limited set of identified attack vectors (authentication, known vulnerabilities, etc.) to statistically determine the rate of success and/or failure, as well as root causes. They can also be conducted as a series of events constructed to be increasingly detectable over time to statistically determine thresholds of sensitivity.
- Automated test events (alternative limits/automation). These events would be a controlled series of indicators (which may not necessarily require the services of either a Blue or Red team) that replicate the symptoms of abnormal network activity, internal traffic loading, or data-exfiltration. These would be used to evaluate network team responses and detection capability. Such automated events would be useful in accomplishing repetitive vector assessments as well as proxy target events.
- Proxy target events (alternative limits/non-operational). Assessment teams focus on locating and exfiltrating target files specifically placed at critical network locations as a means of determining depth of penetration, potential mission impact (without actually disrupting operations), attack pathways, and effectiveness of specific defense and detection devices ("Capture the Flag"). Alternatively, essentially harmless target files (or limited purpose macros constructed to replicate unauthorized activities) can be planted at critical network locations as a means of determining the ability of the network management and defense systems/personnel in detecting and reacting to these activities ("Scavenger Hunt").
- Adversary Level-of-Effort Metrics (alternative limits/precise force). If the level of effort expended by a Red team is one de facto measurement of the level of network protection, detection, and reaction (just as the level of force applied in kinetic testing is a de facto measurement of material strength), then the need to more precisely measure and express the level of effort brought to bear against the network or system is essential to scoping an assessment and analyzing the results. These metrics would include observation of success/failure along selected Red team attack vectors, time expended, manpower/tool levels, and possibly time-sensitivity factors (i.e., Was a successful attack achieved within a critical time-span?).
- Test Range events (non-operational/segregation). While the best method for observing risk to operational networks is to conduct tests on the operating network, one method for reducing actual risk to those networks is to conduct discrete or high-intensity tests on a simulacrum—a similarly configured test network that does not convey risk to actual network components or systems. While this type of test is more akin to laboratory testing than to live system testing, the use of a test network (and, potentially, simulations or models) allows the assessment of specific issues that would otherwise induce unacceptable degrees of risk to operating and operational systems and networks.
- Casualty testing (non-operational/precise force/segregation). One of the most critical IA precepts is the ability to reconfigure or restore a system following a casualty, system attack, or other debilitating event. The very nature of such events causes most network owners to shun such testing. The risk incurred in "bringing down" any portion of the network, however, can be ameliorated by inducing the casualties in a very limited scope (specific systems, specific durations, specific network segments) and observing the subsequent actions.

Conclusion

Implementation of some, or all, of these types of assessment/test events can meet the goals of all three stakeholders in the IA assessment process: (a) they are intended to provide a baseline for Blue and Red team activities, but only a baseline—they do not replace the existing skills and techniques employed by these teams, nor do they represent any significant expansion to their tasks; (b) they serve to increase the degree of control and decrease the risk present in conducting such assessments in operational environments, while preserving the most critical attributes of those environments in the scope of the assessment; and (c) they provide a standardized basis by which multiple assessments can be compared, either of the same system, or of same/similar networks and environments.

Each of the three major stakeholders must accept some change to the way they currently conduct these assessments. For the Blue and Red teams, it means incorporating a more scripted structure to the often more freely executed penetration and exploitation efforts, but it does not replace the element of "free-play" in the assessment. All of the tasks described above are within the current scope of skills and expertise for these teams and should not require additional personnel, time, or significant resources. For the exercise planner, it means incorporating more aggressive events into the exercise structure, but it also means a significant reduction in the risk represented by those events. For the operational evaluator, it means developing more specific assessment plans, but it also means a greater return in terms of observations and replicable data.

For each of the stakeholders, the greatest obstacle to implementing such an approach may be essentially cultural. It will require IA teams to think like exercise planners, assessment planners to think like IA teams, and exercise planners to think like operational testers. In the end, however, all three are likely to find that the final product of the assessment/exercise event is a better view to how well DoD networks are performing. ❑

David J. Aland is an employee of Wyle, supporting the Office of the Secretary of Defense, Operational Test and Evaluation Directorate (OSD DOT&E) in the assessment of Information Assurance and Interoperability during major DoD exercises. He is a graduate of the U.S. Naval Academy and U.S. Naval War College, and a retired Naval officer with prior experience as Sixth Fleet Communications and Information Systems Officer (N6) and as deputy to the Navy Chief Information Officer. E-mail: david.aland@wylelabs.com

ITEA Journal 2008; 29: 67–74

Best Practices for Developmental Testing of Modern, Complex Munitions

Capt Joshua Stults
780th Test Squadron, Eglin AFB, Florida

The growing cost and schedule constraints on government weapons development programs as well as their rising complexity increase the need for a decision theoretic-framework for product development. This framework must rely on insight gained from a variety of sources for test planning, test evaluation, and decision support. The best practices presented in this article for system-level developmental test planning and execution are collected from reported experience and criticism of industry and government product development programs. These practices and methodologies are applied in a coherent framework that allows a formal combination of the disparate sources of product knowledge available to decision makers in the early stages of development.

Key words: Bayes Theorem; best practices; complexity; external validation; knowledge-based acquisition; weapons systems.

This article illustrates a formal decision support framework for program managers and testers that embodies the ideas of knowledge-based acquisition and incorporates best practices identified from historical product development programs in the government and commercial sectors. Emphasis is on system-level developmental test and evaluation (DT&E) in support of risk reduction for production decisions. The framework consists of four basic steps: identify relevant system performance factors, use prior knowledge to evaluate system level outcomes, incorporate validated knowledge into product improvements and evaluate sufficiency of testing through external validation. The motivation for such a formal decision support framework is the growing complexity of modern weapon systems. While complexity is not easy to define or measure consistently, indicators of complexity are type and number of weapon sensors, multiple operational modes, multiple communications links, software for autonomous loitering or targeting, etc. These indicators have been shown to increase the cost of test and evaluation (T&E) despite the significant constraints currently being placed on weapons development funding (Fox et al. 2004).

The motivation for knowledge-based acquisition is to improve product development outcomes using "quantifiable and demonstrable knowledge to make go/no-go decisions" (GAO 2005). It is based on ensuring that the proper product knowledge is validated at critical decision points (DoD 2003). Central to this acquisition approach is the progression of the product through well-defined maturity levels, driven by validated product knowledge.

Three main product maturity levels have been identified through analysis of successful product development practices in industry. The product progresses through these levels based on specific events that demonstrate validated product knowledge rather than schedule driven milestones (GAO 2000). Heuristics learned from commercial and government product development programs can guide the planning of a knowledge validation (testing) program to successfully progress through the product maturity levels. Ideas such as "break it big early" are examples of these sorts of experience-based rules of thumb (GAO 2000).

In addition to informal rules of thumb, there are rigorous inference methods that can support knowledge validation and decision making even in the system development phase when sample sizes are too small for standard large sample size statistical methods to apply. For example, approaches based on Bayes theorem which incorporate prior knowledge in evaluating new knowledge as it arrives can ensure that product developers are making informed decisions even in the face of few samples. Sequential Design of Experiments is another method that allows for smaller expected numbers of test events to achieve a given statistical power by using some sort of stopping rule (Cohen and Rolph 1998).

The product maturity paradigm, experience-based heuristics, formal inference and design of experiments methods can be tied together into a coherent decision support framework by a high-fidelity system performance model as suggested in (Cohen and Rolph 1998). System performance models provide a repository for the product knowledge gained as the system matures, so that successive testing can be planned based on validated knowledge. They can support a constructive approach to testing that leverages knowledge discovery from the early phases of product maturity for more efficient system level DT&E. Likewise, as has been previously suggested, the knowledge gained from DT&E to develop and validate the system performance model should be used for efficient operational test and evaluation (OT&E) planning (Cohen and Rolph 1998).

A recurring criticism of Department of Defense product development is that programs proceed without the right kind of knowledge gained from test efforts. When this happens cost, schedule, and performance problems often result (GAO 2003). As has been observed, "It is possible to conduct a test or simulation that does not contribute worthwhile information" (GAO 2003). By focusing on knowledge validation and knowledge driven product maturity rather than specific test schedules or events, we hope to avoid this waste of effort and ensure that all planned test events validate the right knowledge at the right level of product maturity.

Product maturity levels

Three levels of product maturity identified in (GAO 2000) are:

1. Technologies and subsystems work individually;
2. Components and subsystems work together as a system in a controlled setting;
3. Components and subsystems work together as a system in a realistic setting.

This article will focus on the second and third levels of product maturity which correspond to system-level DT&E. Oftentimes because the number of system-level tests during the DT&E phase of weapon development is not large enough for statistical significance in the classical frequentist sense, these tests are relegated to "demonstration" status. When incorporated into a Bayesian inference framework, these tests can support a meaningful estimate of parameters important to programmatic decisions from the first test event. In addition, the marginal value (reduction in risk) of additional testing can begin to be compared to the marginal cost of that testing. This comparison is critical to allowing for a decision theoretic approach to answering the question of how much testing is enough (Cohen and Rolph 1998).

Knowledge validated by testing drives the progress of a product through the stages of development. Incorporating the knowledge gained from each phase of testing and development can guide the test plan to be more efficient than starting from assumed ignorance at each stage. Assuming ignorance is conservative as far as technical risk goes, it drives larger and less efficient test plans than if prior knowledge is incorporated into the planning effort.

Historically based heuristics for test planning and product development

A very disciplined approach to maturing a product is required to avoid costly rework late in product development. The three critical factors that underlie this disciplined approach ensure that:

1. Validation is event based rather than schedule based;
2. The quality of the knowledge validated in each event is not sacrificed;
3. The knowledge validated in each event is used to improve the product (GAO 2000).

One of the most important heuristics identified from successful commercial product development efforts is known as "break it big early", or "move discovery to the left" (GAO 2000). This means that challenging validation events are planned early to expose areas of weaknesses in the new design.

Rigorous subsystem verification has been identified as one of the means to reduce the burden of discovery on the later system level test events. This is a way to ensure that the quality of knowledge gained from test events does not suffer due to immature test articles. Aggressive development schedules can often result in an undue burden of discovery on system-level flight testing. Experience in the Theater High Altitude Air Defense (THAAD) program illustrated that shortcomings in component and subsystem validation lead to very expensive failures in the flight test program (GAO 2000). Sacrifices were made in the first two stages of product maturity to keep system level flight testing on schedule. The problems experienced by THAAD were not that tests failed or discoveries occurred, which is the very purpose of testing. In fact, it has been pointed out that "...bad things happen in test and that those bad things are valid results just as successes are" (DOT&E 2007). The object is to find those bad things early in component level and subsystem integration testing, so that the discoveries during more expensive full-up system level testing are small and affordably corrected.

Also in line with the "break it big early" philosophy is to test at factor levels that give the most variation in system performance. System response in most real systems is nonlinear, so the factor level matters. The most knowledge can be gained from a limited number of test events by testing at the most stressing factor levels.

In keeping with the third element of disciplined product development, information gained from initial test events must be incorporated into improving the product. Using knowledge to mature the product and getting the right knowledge to decision makers is the focus rather than sacrificing the quality of test events to maintain schedule goals. The DarkStar Unmanned Aerial Vehicle program experienced significant flight test failures and was eventually terminated due to problems that surfaced during initial flight testing which were not addressed and fixed before subsequent testing continued (GAO 2000). The point here is not that flight test failures cause program termination, but that sacrificing knowledge validation and product improvement based on validated system knowledge to maintain schedule is counterproductive.

If these heuristics are applied to the first two levels of product maturity, then the burden of discovery on system-level DT&E will be reduced (GAO 2005). This allows more operational realism to be incorporated into DT&E, thus improving the quality of knowledge gained from these test events.

The Stand-off Land Attack Missile – Expanded Response (SLAM-ER) system experienced failures during OT&E that were masked in earlier testing because of unrealistic DT&E test conditions and immature test articles (GAO 2000). This shows how the heuristics identified can complement each other, mature test articles support more operational realism in DT&E which in-turn supports "moving discovery to the left."

To summarize the above discussion, here is a collection of some of the experience-based rules of thumb:

- *Break it big early, move discovery to the left*
  - Rigorous subsystem verification and integration minimizes discovery burden on the final, most expensive system-level development effort;
  - Test difficult technology or design features early;
  - Test at factor levels that give the most variation in system performance: System response in most real systems is nonlinear, the level matters.
- *Focus on getting necessary knowledge to decision makers rather than specific events, techniques, or schedules*
  - Incorporate information from early test events to improve the product before proceeding to future test events;
  - Do not curtail early testing to stay on schedule;
  - Do not sacrifice test-item fidelity to stay on schedule: Unrealistic system level test events lower the amount of useful information gained from those events.

Importance of system performance models

Incorporating knowledge gained from disciplined component and subsystem validation into a high-fidelity system performance model informs decision makers about development and production risk. This can also lead to more efficient test planning and analysis. The system performance model tracks the system through the product maturity levels. As product knowledge is validated in each level, that knowledge is incorporated into the model. The model provides a means for the heuristics identified in Section 3 to be rigorously applied. It allows the test planner to answer the questions like:

- Where can I expect the most variation?
- What level of product maturity is the modeled performance based on?
- What discoveries have been made, and has that knowledge been incorporated into the product (and its model)?

The test planner can make basic decisions about influential factors and their likely critical levels before design details of the actual test article are finalized. In other words, "one can design an effective test for a system without understanding precisely how a system behaves" (Cohen and Rolph 1998). This allows testing for the later levels of product maturity to be based on knowledge gained during the initial levels. *Figure 1* illustrates the progression of model maturity. Initially, the insight for test planning comes from physics-based simulation and other analysis tools. As the product matures and component and integration testing data become available these can be used for test planning and decision making. The fast running engineering models are based on the more fundamental information in the detailed physical models. Component performance and integration testing data are incorporated as they become available.

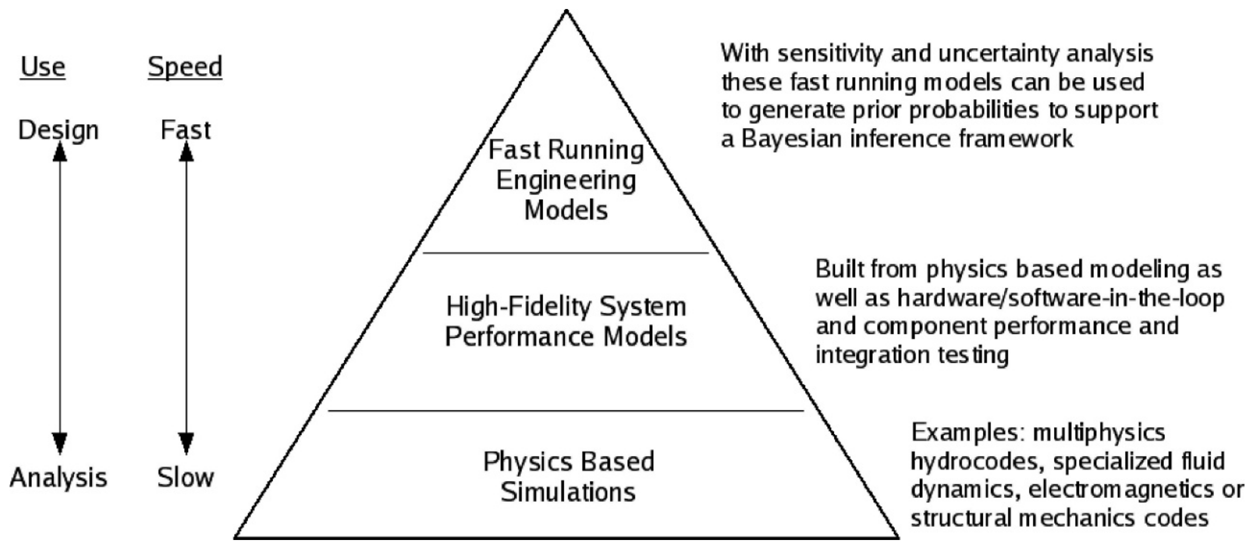


Figure 1. Modeling hierarchy

Incorporating prior knowledge

Knowledge captured in the system performance model (based on component level testing and system design analysis) can be used to generate prior probabilities in performance metrics of interest. These prior probabilities, or degrees of belief, are useful for a Bayesian inference method.

The Bayesian approach has advantages over approaches which do not adjust their prior probabilities based on experience (Robbins 1964). It is desirable because it gives an optimal prediction: given the hypothesis prior probabilities, any other prediction will be correct less often (Russell and Norvig 1995). Bayes Theorem is shown in Equation 1.

$$P(H_j|E_i,I) = \frac{P(E_i|H_j,I)P(H_j|I)}{P(E_i|I)} \qquad (1)$$

Where the posterior, or final, probability of the hypothesis, H_j, being true given the new data, E_i, and the background information, I is updated by the likelihood, $P(E_i|H_j, I)$, and the prior or initial probability, $P(H_j|I)$. Beliefs about the system under test are updated by new information gained from each test event.

A common criticism of the Bayesian approach is that there is subjectivity in choosing the prior probabilities. This is true, but the benefit is that an explicit exposition of the assumptions underlying the test planning and analysis has been made, which is often not the case for other test planning approaches. In addition, the dependence of the result on the prior probability decreases as the sample size increases. In the large sample size limit, for certain model assumptions the Bayesian approach matches the more standard frequentist result (D'Agostini 2003).

High level test planning for weapon development programs tends to focus on the number of end-to-end flight tests because this is a significant contribution to overall test program cost and schedule. Performing enough end-to-end testing to build confidence intervals based on large sample-size theories is cost and schedule prohibitive, so the end-to-end testing is many times relegated to a demonstration only status. If the system level test events are merely demonstration, there is little rigorous or quantifiable connection between those small samples and knowledge gained to support decision criteria.

Since there is no quantifiable connection the argument is often put forth that a sample of 1 is as good as 1 + m, where m is some number small enough that large sample theories still do not apply with sufficient power. This argument is fallacious because large sample theory is not meant to measure the difference in marginal information gained between two small samples. It does not follow that there is no difference in value to the decision maker because large sample theories cannot measure that difference.

A Bayesian approach incorporates assumptions and prior knowledge about the system under test in a formal way so that information gained beginning with the first test event improves the certainty of the knowledge about the system in a quantifiable manner. Some estimation of the marginal value of n and $n + 1$ samples can be evaluated even though n is far too small for frequentist statistical approaches to apply. There is no free lunch here. With very small n the inferences supported by a Bayesian approach will be quite sensitive to the priors; however, that sensitivity information can be provided to decision makers so

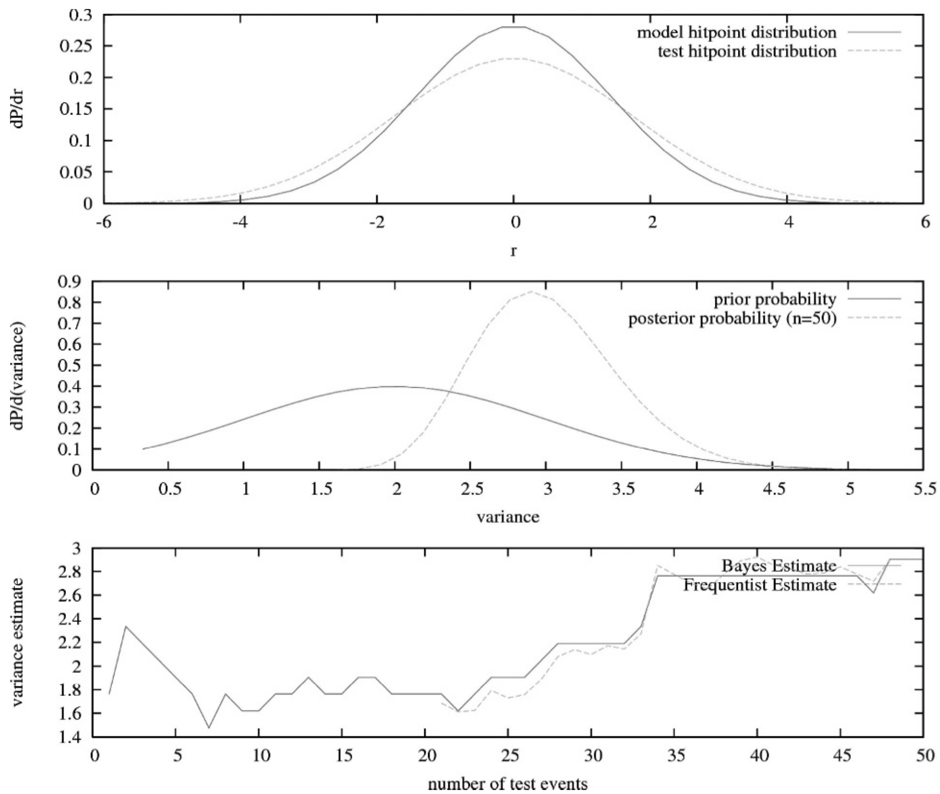


Figure 2. Estimating variance in hit-point distribution

that they understand what increasing n will mean in terms of reduced risk.

Hit-point distribution

This section presents an example of the Bayesian approach evaluating hit-point distributions for a munition with some type of smart terminal guidance based on a multimode seeker and target recognition algorithms. The seeker component level testing and closed-loop guidance and control simulation can provide a probability density for the hit-point in the plane normal to the weapon's attack vector. This information provides a prior probability for evaluating the hit-point from the very first end-to-end flight test. For smaller, smarter munitions this hit-point becomes increasingly important. Great variations in system effectiveness (i.e., killing the target) might be expected for small variations in hit-point.

Figure 2 illustrates using the Bayesian approach to estimate the variance in hit-point distribution. The model predicts a radial distribution of hit-points with a variance of two, while the actual performance is drawn from a distribution with variance of three. The variance in this example is our hypothesis, and the prior probabilities (see Equation 1) for the hypothesis could be generated from sensitivity and uncertainty analysis of the model. The actual form for the prior is not critical as long as there is some finite probability assigned to the true answer (Russell and Norvig 1995).

The lowest graph in *Figure 2* shows the maximum probability estimate of the Bayes method and compares it to the standard frequentist result (for $n > 20$). Rather than integrate over the continuous hypothesis space (variance in this case), a discrete set of hypotheses is evaluated. This is why the Bayesian estimate in *Figure 2* jumps discontinuously between levels. The method allows significant insight into the problem while the sample size is still small compared with more standard estimation methods.

Model output for prior probabilities

Suppose the output of an uncertainty analysis for a simple fast-running model can be given by Equation 2,

$$y = \beta_0 + e_0 + (\beta_1 + e_1)x \tag{2}$$

where $\beta_0 = 1$, $\beta_1 = 3$, and e_0 , e_1 are normally distributed errors with zero mean and 0.25 standard deviation. The variation simulated here by e_0 , e_1 can be generated by sensitivity and uncertainty analysis in a fast running engineering model. The prior distributions for the model parameters can be estimated by holding the other parameters constant at their expected value and treating each data point as a measurement of the parameter of interest.

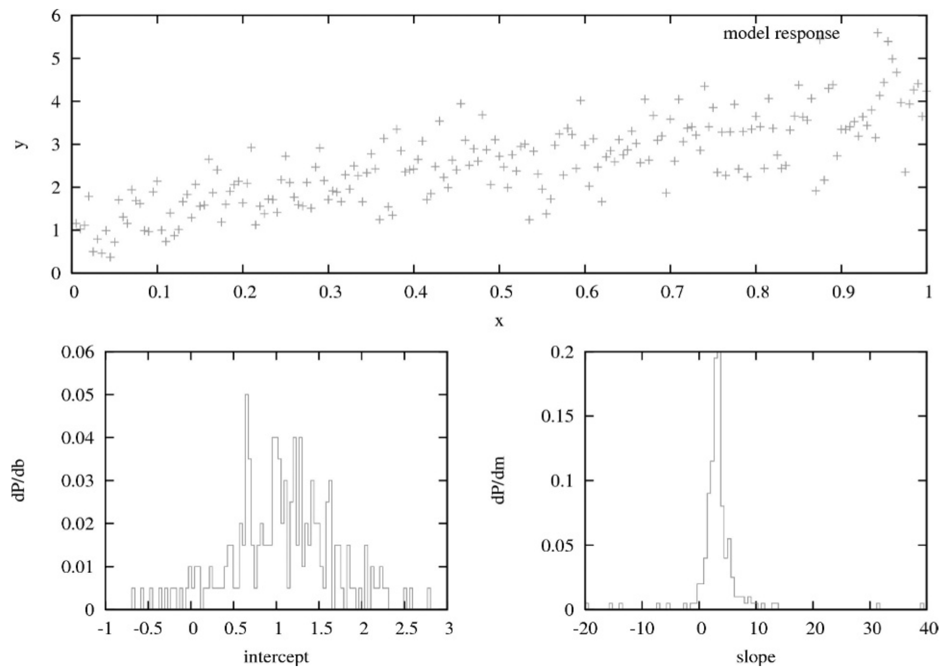


Figure 3. Estimation of prior probability from model output

Figure 3 shows the probability distributions for the slope and the intercept of the model's output following this method. These prior probabilities can be used to guide test planning by identifying where variation or uncertainty is greatest, which leads naturally to where testing will be most profitably executed. The best practice heuristics previously discussed become more than just good rules of thumb when informed by a Bayesian planning and analysis framework. This framework provides insight into where the variation in system performance can be expected, because it explicitly incorporates the prior knowledge from component-level testing residing in the system performance model.

Sequential design of experiments

The basic idea of sequential design of experiments is to test progressively from the outside of the parameter space, capturing linear effects, towards the inside of the parameter space, capturing higher-order interaction effects if needed (Curry and Lee 2007). A comprehensive review of the field is given in (Lai 2001). At each level, the predictive power of the effects measured so far is evaluated and a decision is made about whether additional testing is required.

For example, perhaps the product development team has identified some significant factors for a notional munition with terminal phase guidance and in-flight communication as follows: target aspect (TA), target speed (TS), target movement duty cycle (TMDC), impact angle (IA), engagement mode (EM), and target type (TT). Factors such as noise environment or weather are generally uncontrollable by the testers, but it is worthwhile to note their significance and then record their levels during test events so their influence on performance can be quantified (Cohen and Rolph 1998).

An initial experimental design will attempt to measure the linear or "main" effects. For the six controllable factors identified above, a seven-parameter model results, requiring seven tests at the minimum to make point estimates of the parameters (shown in Equation 3). Two additional tests are added to the design so that some estimate of the process variability can be made, and a final confirmation test is added to evaluate the sufficiency of the linear model.

$$Y = \beta_0 + \sum_{i=1}^{n} \beta_i x_i \qquad (3)$$

Given ten test events and minimum and maximum levels for each of the factors, a constrained optimization method can be applied to find the combination of factor levels across the tests that gives the lowest factor correlation. This is known as a d-optimal test design since it maximizes the determinant of the factor correlation matrix (Curry and Lee 2007).

One method of reaching an approximate optimum is simulated annealing (exactly orthogonal test series exist only at multiples of four tests). It is a heuristic optimization method that combines both divide-and-conquer and iterative improvement strategies (Kirkpatrick and Gelatt 1983). The method starts with a feasible set of factor levels for the test series and then swaps factor levels and evaluates if this improves or

Table 1. Approximately d-optimal test design

Test	TA	TS	TMDC	IA	TT	EM
1	360	20	0.1	15	1	1
2	360	20	0.9	75	−1	−1
3	180	4	0.9	15	−1	1
4	360	20	0.9	15	−1	−1
5	180	4	0.9	15	1	1
6	180	20	0.9	75	1	1
7	180	4	0.9	15	1	−1
8	180	20	0.9	75	−1	1
9	360	4	0.1	75	1	1
10	180	4	0.1	75	−1	−1

TA, target aspect; TS, target speed; TMDC, target movement duty cycle; IA, impact angle; TT, target type; EM, engagement mode.

degrades the orthogonality of the tests. If the change improves the orthogonality, it is accepted with probability, $P = 1$. If the change degrades the orthogonality, it is accepted with probability relation shown in Equation 4.

$$P = e^{d_1 - d_0 / T} \tag{4}$$

Where d_1 is the determinant of the correlation matrix (a measure of orthogonality or "goodness") and T is the temperature, a parameter that is gradually reduced during the optimization. This allows the process to avoid being trapped by local minima because it accepts moves which are "bad" according to the difference $d_1 - d_0$ and the cooling schedule in T. As cooling progresses the algorithm accepts "bad" moves with less and less probability.

A test series developed by the simulated annealing method is shown in *Table 1*. The correlation of factors across the test events for this design is shown in *Table 2*.

An exactly orthogonal series would have no nonzero off-diagonal terms in the correlation matrix. The goal of the optimization is to make these terms approximately zero. The advantage of using an optimization technique like simulated annealing is that constraints on the test design can easily be added and optimization can proceed exactly as before, only within the reduced set of feasible designs. For example, the factors describing an important operationally representative scenario can be constrained to occur a given number of times.

Importance of external validation

In a test program that relies heavily on modeling and simulation, it is critical to guard against over-fitting the model. The basic algorithm to avoid such over-fitting is known as "model-test-model-test" (Cohen and Rolph 1998). The final validation tests are outside the scenarios which were used for parameter tuning. Sequential design of experiments naturally provides the framework for such an approach. The stopping rule in a standard sequential design depends on evaluating the predictive power of the simple empirical model using the final additional test.

When a high-fidelity system performance model is available the stopping rule should be modified to depend on an external validation of the system performance model as well as the more standard stopping rule. The initial tests used to develop the simple linear empirical model can also be used for parameter tuning of the high-fidelity model and the final test serves as an external validation of the high-fidelity model as well.

Conclusions

High-fidelity system performance models along with full-up system level test events incorporated into a formal inference framework provide rigorous support to decision makers in developing and acquiring modern weapon systems of ever-increasing complexity. The proposed framework for knowledge-based test planning and execution consists of four basic steps:

1. Identify significant factors and levels based on a high-fidelity system performance model;
2. Use the model for prior distributions (context, background knowledge) with which to analyze full-up system level test outcomes;
3. Incorporate discoveries into product improvements and improved performance model;
4. Evaluate sufficiency of testing based on predictive power of high-fidelity system performance model, i.e., model-test-model-test.

Table 2. Factor cross-correlation matrix

	TA	TS	TMDC	IA	TT	EM
TA	1	0	0	0	0.2	0
TS	0	1	−0.16667	0.16667	0	0.102062
TMDC	0	−0.16667	1	−0.16667	0	−0.102062
IA	0	0.16667	−0.16667	1	0	0.102062
TT	0.2	0	0	0	1	0
EM	0	0.102062	−0.102062	0.102062	0	1

TA, target aspect; TS, target speed; TMDC, target movement duty cycle; IA, impact angle; TT, target type; EM, engagement mode.

The exact mechanics of the approach presented in this article are not critical. Any integrated method that gives some measure of the marginal value of system-level test events when sample sizes are small can provide useful support to decision makers. This support will begin to allow hard risk management decisions about how much testing is sufficient to be made in a more decision-theoretic framework.

The critical aspect of the approach is the knowledge warehouse known as the system performance model. The knowledge it contains at the same time informs decision makers and test planners, and provides a repository of validated knowledge from test conductors. The execution of a knowledge-based test program supports decision makers with solid information about test sufficiency and risk. Through improvements incorporated into the product and its model, it ensures that decisions made about the system are based on the highest quality of information available. ❑

CAPT JOSHUA A. STULTS is the deputy live fire agent for conventional munitions, Air Force Live Fire Office, 780th Test Squadron, Eglin AFB, FL. Prior to his current duties supporting U.S. Air Force weapons programs in planning and executing live fire test and evaluation, he was a test engineer for the 780th supporting weapons development flight test. He holds a master of science degree in aeronautical engineering from the Air Force Institute of Technology, Dayton, Ohio; and a bachelor of science degree in aeronautical engineering from the U.S. Air Force Academy, Colorado Springs, Colorado. E-mail: joshua.stults@us.af.mil

References

Cohen, M. L. and Rolph, J. E. (eds.) 1998. "Statistics, Testing and Defense Acquisition: New Approaches and Methodological Improvements." *Washington, D.C.: National Academy Press*.

Curry, T. F. and Lee, S. J. 2007. "Using Sequential-Designed Experimentation to Minimize the Number of Research and Development Tests." *The ITEA Journal of Test and Evaluation,* Volume 28-2, pp. 41–47.

D'Agostini, G. 2003. "Bayesian Inference in Processing Experimental Data: Principles and Basic Applications." *Rept. Prog. Phy*. 66, pp. 1,383–1,420.

DoD, (Undersecretary of Defense for Acquisition, T. and Logistics). 2003. "Operation of the Defense Acquisition System, No. 5000.2." In: *Department of Defense Instructions.* May 2003, pp. 1–50.

DOT&E. 2007. *"Lessons Learned from Live Fire Testing: Insights Into Designing, Testing, and Operating U.S. Air, Land, and Sea Combat Systems for Improved Survivability and Lethality."* O'Bryon, J. F., (ed). Washington, D.C.: Secretary of Defense, Operational Test and Evaluation Directorate (DOT&E), Live Fire Test and Evaluation, pp. 3–15.

Fox, B., Boito, M., Graser, J. C. and Younossai, O. 2004. "Test and Evaluation Trends and Costs for Aircraft and Guided Weapons, Tech. Rep. MG-109." Arlington, Virginia: RAND Corporation.

GAO (General Accounting Office). 2000. "Best Practices: A More Constructive Test Approach is Key to Better Weapon System Outcomes, Tech. Rep. GAO/NSIAD-00-199." Washington, D.C.: U.S. General Accounting Office. Available online at http://www.gao.gov. Accessed January 18, 2008.

GAO (General Accounting Office). 2003. "Defense Acquisitions: Assessment of Major Weapon Programs, Tech. Rep. GAO-03-476." Washington, D.C.: U.S. General Accounting Office. Available online at http://www.gao.gov. Accessed January 18, 2008.

GAO (General Accounting Office). 2005. "Best Practices: Better Support of Weapon System Program Managers Needed to Improve Outcomes, Tech. Rep. GAO-06-110." Washington, D.C.: U.S. General Accounting Office. Available online at http://www.gao.gov. Accessed January 18, 2008.

Kirkpatrick, S. and Gelatt, M. V. 1983. "Optimization by Simulated Annealing." *Science,* Volume 220, No. 4598, pp. 671–781.

Lai, T. L. 2001. "Sequential Analysis: Some Classical Problems and New Challenges." *Statistica Sinica,* Volume 11, pp. 303–408.

Robbins, H. 1964. "The Empirical Bayes Approach to Statistical Decision Problems." *The Annals of Mathematical Statistics,* Volume 35, No. 1, 1964, p. 1.

Russell, S. and Norvig, P. 1995. "Artificial Intelligence: A Modern Approach." 2nd ed. Upper Saddle River, NJ: Prentice Hall. p. 1132.

Acknowledgments

The author thanks Maj David Winebrener for the lively discussions on test planning and lessons learned that led to this research effort.