The ITEA Journal

June 2011
Volume 32, Number 2

Organizing to Accomplish the Mission

Published quarterly by the International Test and Evaluation Association

CONNECT with ITEA to LEARN, SHARE, and ADVANCE!

NACMAST Cyber Test Range

EWA GSI is the primary component of the Army Research Lab (ARL) Network Attack Characterization Modeling and Simulation Testbed—NACMAST Enterprise. The NACMAST testbed environment facilitates modeling and simulation tools to detect and characterize network attacks in near-real time, provide software solutions to mitigate attacks, and establish a testing environment to certify network protection tools. With a consortium of universities and the ARL, we've established a new laboratory dedicated to computer network intrusion detection, and an ancillary test range for immediate testing and implementation of new computer network defense (CND) technologies. EWA GSI maintains a customized training program on intrusion detection techniques for all test range participants.

EWA Government Systems, Inc.
"Enabling a more secure future"

CONTENTS

OPERATIONAL TEST AGENCIES ORGANIZING TO ACCOMPLISH THE MISSION

ITEA BEST PAPERS

TECHNICAL ARTICLES

DEPARTMENTS

ON THE COVER: The precursor to current day simulation began in the 1940s with the development of the finite element method for structural analysis, and computational fluid dynamics for weapon design, aircraft design, and weather forecasting. In the 1960s it came into its own with improved understanding of numerical methods and development of scientific computers, especially those of Univac and Control Data Corporation. Vector computers, parallel computers, and commodity clusters have continued to change the face of simulation and made ever more complex configurations, scenarios and multi-physics phenomena accessible. Instead of analyzing an individual platform we can represent a battlespace with its plethora of vehicles, weapons, sensors, players, dynamics, and interactions. It may now be the case where computational capacity outstrips our understanding of phenomena to be modeled, and legacy algorithms, programming languages, and operating systems cannot fully utilize the emerging hardware. This issue addresses the state of simulation, sheds light on some limitations in realizing past expectations, and provides suggestions for moving forward. (Cover graphic courtesy of Mr. Matt Guy of the Air Force Operational Test and Evaluation Center.)

CONNECT with ITEA to LEARN, SHARE, and ADVANCE!

■ 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.*

■ For over thirty years the International Test and Evaluation Association (ITEA), a not-for-profit education organization, has been advancing the exchange of technical, programmatic, and acquisition information among the test and evaluation community. ITEA members come together to learn and share with others from industry, government, and academia, who are involved with the development and application of the policies and techniques used to assess effectiveness, reliability, interoperability, and safety of existing, legacy, and future technology-based weapon and non-weapon systems and products throughout their lifecycle. ITEA members embody a broad and diverse set of knowledge, skills, and abilities that span the full spectrum of the test and evaluation profession. All of which is shared with others through the ITEA Journal–the industry's premier technical publication for the professional tester–and at ITEA's Annual Symposium, regional workshops, education courses, and local Chapter events. Join the thousands of ITEA members–your peers in the industry–in contributing to the ITEA Journal and participating at ITEA events so that you also can benefit from the opportunities to learn from others, share your knowledge, and help advance the T&E industry.

■ *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 offers annual membership for individuals, full-time students, and any organization interested in advancing the test and evaluation profession. Annual membership dues are $50 for individuals ($70 if located outside of the U.S.) and $25 for full-time students. Annual Corporate Membership dues are $500 for small organizations (less than 50 employees), and $1,000 for large organizations (50 or more employees). Corporate Memberships include three (3) Individual Memberships for small organizations and five (5) Individual Memberships for large organizations. 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 2011; 32: 117–118

What is the current Department of Defense test community (i.e., organizations and capabilities)? Have severe budget constraints led the defense test and evaluation (T&E) organizations to a wrong focus on investments and expenditures divorced from the mission of preparing and equipping the warfighter? The aging workforce continually challenges us to find critical skills while more complex systems demand not only a flexible workforce, but more "System of System" engineers. Insourcing is the administration's attempt to ameliorate some of these concerns but it has short- and long-term consequences. These questions and current real world challenges were the ITEA Publications Committee's focus when planning this issue. Their objective was to take a candid look at defense T&E organizations and capabilities; academic preparation, certifications, project and program management skills; and what we need to do in the future in training, hiring, and improving our testing—especially as budgets get tighter and the scrutiny on programs increases.

Stephanie H. Clewer

Ultimately, T&E enables the acquisition community to learn about limitations (technical or operational) of a system under development, so that they can be resolved prior to production and deployment. Of course, all acquisitions are plagued with cost and schedule constraints as well. I continue to hear increasing references to the T&E roles in initiatives such as rapid, Agile, and evolutionary acquisition processes. In all of these initiatives, the overarching objective is to balance needs and available capability with resources to put capability into the hands of the user quickly. The success of the incremental Agile and evolutionary approaches depends on the phased definition of capability needs and requirements and an awareness of the technology maturation that will enable increasing capability over time. The T&E strategy of these systems must allow for adequate T&E of each increment intended for fielding with emphasis on regression testing of operational and/or suitability requirements of subsequent increments. The bottom line is that the acquisition process is changing…and the test community must change in concert. First and foremost, incremental acquisitions require earlier collaborations among the testers, users and evaluators and a long-term strategy that ensures a tightly coupled linkage between evolving technologies and requirements and a responsive acquisition, test and sustainment strategy.

The Office of the Deputy Assistant Secretary of Defense, Developmental Test and Evaluation (DT&E), created in May 2009, is focusing on early involvement in the acquisition life cycle by testing and evaluating earlier and more effectively. By advocating T&E infrastructure support to rapid prototyping and medium- and small-scale modifications to existing weapons systems, the DT&E goals are commensurate with the rest of the acquisition community—to decrease costs, enhance performance, and retain program schedules.

I especially like the theme of this issue…*Organizing to Accomplish the Mission*. As I have reiterated in the last two issues of the ITEA Journal, that has been our theme at ITEA this year as well. The ITEA mission is to advance the field of test and evaluation worldwide in government, industry and academia. Therefore, to remain relevant, ITEA must also evolve in step with the changes in the Department of Defense acquisition processes.

Another mission of DT&E is to develop strategies to enhance the technical management workforce via practical experience and training, redefine the T&E acquisition certification standards, and ensure education requirements are balanced with job demands for military and civilian personnel. ITEA also recognizes the importance of these workforce-related strategies and considers it our fundamental responsibility to play a role in supporting them. I am pleased to announce the addition of John "Jay" Weaver to the ITEA Executive Office staff as the Association's Director of Education. Jay brings a wealth of experience including military service with the U.S. Air Force and positions with other not-for-profit associations where he led professional development, adult education, and individual certification programs. Jay's primary responsibilities are to augment the current ITEA educational content, develop new curriculum, promote continuous learning opportunities, collaborate with other academic institutions, enhance and facilitate our academic scholarships, and develop and manage potential certification programs and grant funding.

The Director of Education is a new staff position established by the Board of Directors with the overarching goal to continue ensuring that ITEA remains recognized as the premier provider of a "continuum of learning" meeting the evolving needs of the T&E community through a variety of timely, relevant, and easily accessible education opportunities. Jay is working closely with the ITEA Education Committee to identify gaps in our current educational offerings. If you have not already done so, please take a few minutes to complete their "Recommended Topics for ITEA Education Events Survey" on the ITEA website. Input from our members is very important to us! Also part of our "organizing to meet the ITEA education mission," we recently partnered with Air Academy Associates to augment our educational content by offering additional courses related to Combinatorial Testing with Design of Experiments.

Although improving the quality and relevance of our education offerings is one of our primary goals, the Board of Directors is also taking a hard look at ITEA's Strategic Plan and the associated Five-Year Action Plan for meeting our other high priority goals:

1) Broadening participation;
2) Strengthening our chapters;
3) Increasing opportunities for information exchange and networking through association events, publications, on-line forums and other avenues;
4) Increasing opportunities to recognize accomplishments of T&E professionals; and
5) Ensuring ITEA's governance is effective and efficient

The Vice President of the Board, Mark Brown, is leading a strategic planning tiger team to ensure that the Association's high priority goals remain aligned with the needs of our members and the T&E community as a whole. In addition, we have already begun an exciting Membership Recruitment Promotion, the "Member-Get-A-Member Campaign," aimed at both broadening participation and strengthening the ITEA Chapters. Check the website for further details and be sure to take the opportunity to possibly win a free Annual Membership for yourself as well as additional scholarship funding for your chapter!

I hope to see many of you at our 2011 Technology Review in Annapolis in July and at our Annual Symposium in Orlando in September. The theme for the Annual Symposium, *Fostering Partnerships in T&E and Acquisition*, will be addressing a number of the challenges I mentioned above, to include both a panel and poster session on Improving the Current and Future T&E Workforce.

As always, I welcome any comments or suggestions sent to me personally via president@itea.org.

Issue at a Glance

ITEA Journal 2011; 32: 119

O*rganizing to Accomplish the Mission*. Just as we face a rapidly and continuously evolving threat, the Department of Defense test community must adapt to ever changing test requirements and circumstances, for example, more complex systems, shorter schedules, reduced budgets, a leaner workforce, and rapid acquisition to name a few. The aging workforce continually challenges us to find critical skills while more complex systems demand not only a flexible workforce but more "System of System" engineers. Doing more with less is not new but remains a mantra. This issue addresses ways the test community is responding and delivering by organizing to accomplish the test mission.

Dr. David Alberts, the Director of Research in the Office of the Secretary of Defense Chief Information Officer (DoD CIO), provides in the *Guest Editorial* a call for partnership between the test and evaluation (T&E) and research communities, to better understand agility, incorporate the consideration of agility in T&E, thus improving the agility of the capabilities we acquire. Daniel Laird of the Air Force Flight Test Center analyzes communication link resynchronization, critical information for the performance of new telemetry systems. Randy Herrin of the Joint Interoperability Test Command (JITC) and Chris Watson of the Defense Information Systems Agency present the history and evolution of the JITC for *Historical Perspectives*.

Commanders of three of the Department of Defense Operational Test Agencies offer perspectives on *Organizing to Accomplish the Mission*. Major General David Eichhorn, Commander of the Air Force Operational Test and Evaluation Center, discusses the impact of the national debt and budget deficits on acquisition, and how T&E as the conscience of acquisition is more critical today than it has ever been. Rear Admiral David Dunaway, Commander, Operational Test and Evaluation Force, describes improved operational test efficiencies and better early involvement through internal process improvements including use of integrated testing, mission based test design, and design of experiments. Colonel Joe Puett, Commander, Joint Interoperability Test Command, presents the challenges of balancing warfighters' demands for increased IT capabilities with strengthening our national security capability while DoD seeks to achieve budget efficiencies.

Two best paper awards from the 2011 LVC Conference appear in this issue. In the *Best Undergraduate Student Paper*, Marjorie Ingle *et al.* from the Center for Space Exploration Technology Research at the University of Texas at El Paso, design a thrust balance and test micropropulsion thrusters intended for space applications. In the *Best Graduate Student Paper*, Jose Barajas and colleagues from New Mexico State University use a rigorous systems engineering approach to identify environmentally benign candidate technologies for energy harvesting and scavenging.

Dr. James Valdes *et al.* of the National Defense University survey requirements and the state of the art in ground truth determination for bioaerosol detector testing. Brian Weiss from the National Institute of Standards and Technology and Dr. Linda Schmidt from the University of Maryland continue work on tools for complex evaluation designs of advanced and intelligent systems. Mark Jones reviews five proposed changes for test and evaluation to improve our ability to identify critical failure modes and design flaws early in a system's lifecycle.

The issue closes with two articles from the 2010 Technology Review. David Bate of Scientific Research Corporation examines the effects of sleep deprivation on working memory and cognitive performance, concluding that no single metric can characterize the effects. Dr. Joshua Gomer of Human Factor Engineering and Dr. Christopher Pagano of Clemson University demonstrate the utility of the NASA Task Load Index for assessing subjective workload data in complex human system technologies, valuable in operator selection and training as well as in design of human system technologies.

Guest Editorial

ITEA Journal 2011; 32: 123–125

Rethinking Test and Evaluation for a *New Age*

David S. Alberts, Ph.D.
Director, Research, DoD CIO

Test and Evaluation (T&E) is a critical component of the processes that develop and/or acquire capability. Done properly, T&E greatly increases the likelihood that the capabilities in question will be effective and efficient; however, organizations and the systems they conceive and develop actually adapt to how they are evaluated. If T&E criteria and/or processes are not well-matched to the demands of the operational environment and the missions to be undertaken, the probability of getting the capabilities one really needs at an affordable price can be significantly reduced.


Dr. David S. Alberts

Although there is always room for improvement, if one knows the capabilities needed and can adequately specify a set of relevant characteristics and performance measures, then existing T&E processes and practices are up to the task at hand. Unfortunately, this foundational assumption upon which our entire approach to T&E is based, that is, that we know and can adequately specify the capabilities we need, is no longer true. Furthermore, even if one could adequately specify what is currently needed given the nature of the Age in which we live, our needs are bound to change, if not during the process of development, then certainly within the economic life of the capability being acquired.

We can no longer ignore the fact that we find ourselves in a new, not-yet named Age, an Age in which we are increasingly interconnected, interdependent, and pressed for time. We no longer can adequately understand our requirements and predict the nature of the circumstances in which we operate. This is not because we are deficient in some way; it is an unavoidable consequence of our Age.

Survival in this *New Age* requires, above all else, *Agility*.[1] This editorial briefly explains why this is so, and explores the implications for T&E.

A *New Age*

Information Age technologies have not only improved our ability to collect, store, process, and disseminate information, but they have also have had a profound effect on a variety of interactions. The products and services that have been enabled by these technologies create a need for interactions that previously were unimagined and have facilitated long-desired interactions that were too difficult or expensive. As a result, these products and services have improved both the speed and the "richness" of our interactions.[2]

> *As a result, we are seeing an explosion of interactions that, in and of themselves, has altered the world in which we live.*[3]

The explosion in the number of interactions has created new feedback loops, increased the number of entities and variables involved in existing feedback loops, and increased the links between and among previous independent domains. As a result, the world is growing more and more interconnected. This increase in interconnectedness has created both possibilities and interdependencies between and among entities.

Our most challenging problems in this *New Age* have two characteristics in common:

- The problem to be solved involves a multidimensional effects space.
- No single entity, however large or powerful, can make progress by itself; any effective solution requires collective action.

These aspects of the endeavors we must undertake are, in part, a result of the changes to our world brought about by Information Age technologies and the capabilities they have enabled. If we look at these endeavors from terrorism and cyberterrorism to natural and man-made disasters, we find that they have not only, e.g., a military or security dimension, but also political, social, and economic dimensions. Tackling these multidimensional challenges, as we have repeatedly found, cannot be accomplished by a single entity. In some cases, they require a "whole of government" solution, while in other cases, an international coalition and/or a public–private effort is necessary.

As a consequence of these two aspects of complex endeavors, the information needed will come from a large number of sources, both known and previously unknown. This information will need to be widely shared among individuals and organizations with whom we have established relationships and those with whom we have not. Furthermore, to develop the critical mass of effort, individual entities will need to coordinate, collaborate, and self-synchronize their efforts in a manner appropriate for the situation. However well intended, narrowly focused actions and/or unilateral action can often have an adverse impact.

The complex nature of *New Age* mission challenges makes it unwise for us to rely on statements of requirements and prognostications of environmental conditions and stresses as the sole basis for T&E. How can we be sure which specific entities will be our partners? How can we know exactly what information we will need to have? What information will we need to share? Is the 80 percent solution going to be enough?

Therefore, testing our systems is necessary, but not sufficient, to ascertain whether or not a real capability is being acquired. Our systems need to support not only our soldiers, sailors, airmen, marines, and civilian employees, but also the individuals we partner with in many other entities—military and civilian, governmental and nongovernmental, public and private. Do we think we know these entities' requirements? If we only test our systems, how can we know that the overall characteristics and performance of this ever-changing collection of systems is adequate?

To make matters worse, we can no longer act as if cyberspace is a benign environment. Is it reasonable to assume that we will be able to function at an acceptable level in the face of a never-ending and evolving stream of attacks? What we know about the likely consequences of such attacks comes mainly from looking at specific systems, not the dynamic collections of systems needed to support complex endeavors.

> Can we assume we will have an adequate solution to the insider threat?

Cyber attacks do more than impact system performance—they can destroy trust. Trust is the key to information sharing. A loss of trust, in systems, information sources, and in the partners upon whom we rely to accomplish our missions, has at least the same adverse impact as a loss of connectivity or performance would have. Furthermore, the damage sustained could be far more difficult to repair and could require much more time to rebuild. Thus, we need to understand the full consequences of our vulnerabilities to cyber attacks, not just the immediate impacts on the systems we are testing.

A way forward

The logical response to the residual uncertainty that remains after we have tested a capability against the stated requirement in the predicted environment(s) is to attempt to measure potential agility.

> *Agility is the capability to successfully cope with and/or exploit changes in circumstances.*[4]

Measuring agility involves a Catch 22. While one can observe manifest agility by observing how an entity or system actually responds to observed or simulated changes in circumstances, since we cannot adequately predict what future circumstances will be, we cannot measure an entity's or system's future or potential agility in this manner. To do so requires both a conceptual model of agility and an instantiation that is tailored to the capability being acquired.

Fortunately, experiments with agility as it applies to command and control, particularly in the context of complex endeavors, have provided a better understanding of agility. We can now use this knowledge to improve current T&E approaches that rely on statements of requirements alone. However, we can greatly improve both our understanding of agility, and our ability to measure the potential agility of entities and systems if we enlist others in this quest.

I am therefore proposing a way forward that involves a partnership between the T&E and research communities. Together we can make progress in better understanding agility and by incorporating the consideration of agility in T&E, we can improve the agility of the capabilities we acquire. Such a partnership will create a positive feedback loop by creating a better appreciation of the costs of a lack of agility, accepted metrics for agility, and a definitive, empirically supported, quantitative link between levels of agility and performance or effectiveness.

This editorial is a call to action. If we get T&E right we will have created a forcing function that will shape the capabilities we acquire, making them more agile. As a result, the likelihood of mission success will be increased while, at the same time, total expenditures can be reduced. Past failures have taught us that "doing it right the first time," even if it costs more and takes more time in the short run, turns out to provide more capability at a lower cost in the long run. In this case, getting it right means greatly increased attention to agility. ❑

DR. DAVID S. ALBERTS is currently the director of research in the Office of the Secretary of Defense Chief Information Officer (DoD CIO) where his research initiatives are focused on issues related to agility,

information sharing, collaboration, trust, collective command and control, and the future of cybersecurity.

Prior to this he was the director, Advanced Concepts, Technologies, and Information Strategies (ACTIS), deputy director of the Institute for National Strategic Studies, and the executive agent for DoD's Command and Control Research Program. This included responsibility for the Center for Advanced Concepts and Technology (ACT) and the School of Information Warfare and Strategy (SIWS) at the National Defense University. He has more than 30 years of experience developing and introducing leading-edge technology into private and public sector organizations. This extensive applied experience is augmented by a distinguished academic career in computer science and operations research and government service in senior policy and management positions.

Dr. Alberts' experience includes serving as a Chief Executive Officer (CEO) for a high-technology firm specializing in the design and development of large, state-of-the-art computer systems (including expert, investigative, intelligence, information, and command and control systems) in both government and industry. He has also led organizations engaged in research and analysis of command and control system performance and related contributions to operational missions. Dr. Alberts has had policy responsibilities for corporate computer and telecommunications capabilities, facilities, and experimental laboratories. His responsibilities have also included management of research aimed at enhancing the usefulness of systems, extending their productive life, and development of improved methods for evaluating the contributions those systems make to organizational functions. Dr. Alberts frequently contributes to government task forces and workshops on systems acquisition, command and control, and systems evaluation.

His recent honors have included the Secretary of Defense's Outstanding Public Service Award, Aviation Week and Space Technology's Government/Military Laurel, and the inaugural Network Centric Warfare (NCW) Award for Best Contribution to the Theory of NCW presented by the Institute for Defense and Government Advancement (IDGA).

Dr. Alberts currently chairs a North American Treaty Organization (NATO) research group, SAS-085, which is tasked to define and explore C2 Agility in the context of complex endeavors. He received a doctorate in operations research and a master's degree from the University of Pennsylvania. His undergraduate work was at City College of New York where he received a bachelor of arts degree in statistics. *E-mail: David.Alberts@osd.mil and davidsalberts@gmail.com*

Endnotes

[1]This is the conclusion of a forthcoming book. Alberts, D. S. The Agility Imperative. A précis can be found at http://www.dodccrp.org/html4/books_downloads.html#agility_precis

[2]The adoption of Information Age technologies changed the economics of information. Originally, this impact was discussed in terms of the relationship between information richness and reach. The concept of "richness of interactions" was introduced by the author to reflect the fact that these changes were about more than information, they were about the nature of the interactions between and among entities.

[3]Alberts, D. S. 2011. unpublished draft. *The Agility Imperatives: Survival Guide for a New Age, Part II*, page 63. Draft is in peer review.

[4]This is the working definition adopted by the North American Treaty Organization (NATO) Research Group on C2 Agility (SAS-085), which is chaired by the author. Case studies and experiments are currently underway to ascertain if this definition can be successfully applied to real world entities and situations.

ITEA Journal 2011; 32: 127–131

To Determine a Networked Telemetry Transceiver Resynchronization Time

Daniel T. Laird

Air Force Flight Test Center, Edwards Air Force Base, California

In the development of the new Telemetry (TM) Transceivers (TxRx), one of the unique characteristics of the TM channel is its signal dynamics due to obstruction (i.e., terrain, weather, aircraft attitude, manmade objects, etc.). One of the key measures of TM-TxRx reliability is time to resynchronize after link loss. To estimate this resynchronization time—which we expect to be relatively short with respect to the transmission epoch—we created a model of epoch time variation based on known link availability. To simulate channel degradation, we generated traffic with a network emulator and switched the channel into binary states (ON | OFF) to force a resynchronization. We processed time measures based on stale packet data, as described in this article, to estimate the resynchronization time.

Key words: Confidence; one-way ANOVA; radio frequency signal loss; wireless communication.

The objective was to measure the time for Techtronics Technology Corporation (TTC) Telemetry (TM) Transceivers (TxRx) to resynchronize after radio frequency (RF) signal loss in a two-node degraded channel for four modulation schemes[1] of an Orthogonal Frequency Division Multiplexed (OFDM) signal under three Time Division Multiple Access (TDMA) timing *epochs*. We degraded the *channel* with a high dB attenuator to impede signal transception between two wireless (radio frequency [RF]) TM-TxRx carrying unidirectional Universal Datagram Protocol (UDP[2]) payload from emulated wired Ethernet terminals. For discussion, the traffic flow is unidirectional from *source* to *sink*. We "tune" traffic payload in bytes and frames per second (FPS) to maximum transmission units (MTU)[3] for each modulation scheme, such that there are no significant packet drops in the pristine channel. We then capture transmission packets with an emulator application that samples time-tagged byte counts of the tuned traffic at the data-link layer and stores count-time samples in a spreadsheet for analysis.

Using a Spirent Ethernet traffic emulator, we configured two wired nodes in a two-node wireless TM-TxRx RF channel linked via a switched (0|80 dB) attenuator over a 2-second, 50-percent duty-cycle clock (one-half period = 1 second, our *epoch-time*), which is triggered by the source Global Positioning System (GPS)–1 pulse per second (pps) clock. We configured the emulator to capture time-stamped Ethernet packets at 1-millisecond accuracy with precision 0.1 microsecond over approximately 60-second intervals, for each *modulation scheme* and *epoch timing*. The emulator is also GPS-1 pps triggered for clock alignment.

Using the TTC Command Line Interface, we configured each TM-TxRx for three timing epochs of 1 : 1, 10 : 10, and 4 : 16. These designators signify *sink : source slot-time* message ratios of the epoch.[4] For each epoch, we test four OFDM modulation modes: BPSK4, QPSK4, QAM164, and AUTO. The "4" represents three-fourths forward error correction (¾FEC) coding. AUTO selects a modulation mode based on channel conditions and is not restricted to ¾FEC; it can and does run ½FEC. The power leveling is automated by the TM-TxRx. An epoch is a composite of transceiver *slot-times*, in which a node either receives (*sinks*) or transmits (*sources*) message traffic. Slot-time is a composite of 1-millisecond *frames,* and minimum slot-time duration is 5 milliseconds, or 5 frames.

Data sampling and processing

The Spirent emulator captures messages at the data-link layer and time-tags the data accurate to 1-millisecond intervals resolved to 0.1-microsecond precision. There are $\sim 65 \times 10^3$ samples over a 65-second sample interval, of which $\sim$32 are in the "OFF" state and ~32 are in the "ON" state of 1-second each, from which we sample data and cull timing information. The filter algorithm is as follows:

1. Get time and count *arrays*: yields time-series matrix $[\mathbf{c},\mathbf{t}]$.
2. First delta-counts and time, $\Delta[\mathbf{c},\mathbf{t}]$ with zeros: yields matrix $[\Delta\mathbf{c},\Delta\mathbf{t}]$, adjacent differences.
3. Second delta zeros removed: yields matrix $[\Delta^2\mathbf{c},\Delta^2\mathbf{t}]$, interval difference: *"delta-time"* and *"delta-counts."*
4. Filter $\Delta^2\mathbf{t}$ with "time-offsets" via "*time-window*": $W[\Delta^2\mathbf{t}|\ T_L,T_U]$: yields *"resync-time"* ($\delta\mathbf{t}$) we seek.

From the 30-plus OFF and ON intervals, each duration >1 second, we filter $\sim 65\times10^3$ samples down to 30-plus data points of counts and time. We create a *means-model* based on the equivalence of first *variation* of a 1-second epoch and second *difference* of our processing algorithm:

$$\Delta^2\mathbf{t}(1) \leftrightarrow 1+\delta\mathbf{t} \equiv 1+\mu(\delta\mathbf{t})+\boldsymbol{\varepsilon}. \qquad (1)^5$$

Where $\mu(\delta\mathbf{t})$ is the mean of all measures, $\Delta\mathbf{t}$; $\boldsymbol{\varepsilon}$ is a vector of *sample residuals*, of an assumed zero mean ($\mu=0$) normal-, $\mathcal{N}(0,\sigma^2)$, time series. In our model, 1 is an array of units modulo our timing, and $\Delta\mathbf{t}$ are measures modulo this unit: the epoch period plus a sink bias time due to epoch message structure and sample clocking scheme. From these details, we model delta-time as a function of resync-time for the sampled data and known *timing constants:*

$$\Delta^2 t^e_{m,k}(1_k) = 1_k + \delta t^e_{m,k}; \qquad k\in\{1,\ldots,K\}, \qquad (2a)$$

$$1_k \equiv T_{1pps,k} + T^e_{snk,k}(\text{measure time-unit}), \qquad (2b)$$

$$1_K \equiv T_{1pps} + T^e_{snk}. \qquad (2c)$$

Subscript "*m*" indexes the set of four modulation mnemonics: {BP4, QP4, QA4, A}; the superscript "*e*" indexes the set of three epoch mnemonics: {1 : 1, 4 : 16, 10 : 10}; and K designates the samples culled via our processing the time-series per modulation mode. Assuming a simple means-model for resync-time, we have in time-series *element form:*

$$\delta t^e_{m,k} = \mu(\delta \mathbf{t}^e_m) + \varepsilon^e_{m,k}; \qquad k\in\{1,\ldots,K\}. \qquad (3a)$$

In a time-series *array form,* for all *k,* we rewrite Equations 2 and 3a as follows:

$$\delta\mathbf{t}^e_m = \mu(\delta\mathbf{t}^e{}_m) + \boldsymbol{\varepsilon}^e_m, \qquad (3b)$$

$$\Delta^2\mathbf{t}^e_m(1_K) \equiv 1_K + \mu(\delta\mathbf{t}^e{}_m) + \boldsymbol{\varepsilon}^e_m. \qquad (3c)$$

Where data array $\delta\mathbf{t}^e_{m\mathrm{k}}$ designates the resync-times we seek, and $\boldsymbol{\varepsilon}^e_m$ is an array of scalar sampling errors, $\varepsilon^e_{m,k}$, normally distributed, zero mean of unknown variance: $\mathcal{N}(0,\sigma^2)$ – "white-noise" – residuals. The sample clock constant, T_{1pps}, is a time window on the epoch. Each clock half-period is trigger aligned on GPS-1pps edge controls switching "ON" and "OFF" times. The sample clock period is $T_{sample}=2$ seconds. So, the ON and OFF windows are $T_{1pps}=T_{sample}/2\approx 1$ second. The period drifts and jitters. This drift and jitter "error" is realized with offset *residue* modulo the epoch, $T_{epoch}=1$ second, for each sample. We also offset "time-bias," T^e_{bias}, which is sink slot-time duration. We represent these as arrays of scalar errors. To offset, we refer the epoch to the GPS-1pps: $(T_{1pps,k}-T_{epoch})\in \varepsilon^e_k$; and the bias to the sink slot-time: $(T^e_{snk,k}-T^e_{bias})\in \varepsilon^e_{b,k}$. We fold these errors and the sample measure error into one $\mathcal{N}(0,\sigma_T{}^2)$ *residual distribution*. We have an error array:

$$\varepsilon^e_{m,b,k} \equiv \varepsilon^e_k + \varepsilon^e_{b,k} + \varepsilon^e_{m,k} = 1_k - 1_M + \varepsilon^e_{m,k};$$
$$k\in\{1,\ldots,K\}. \qquad (4a)$$

In array form for all *k*, and from Equation 5c.3 below, we get total residual error:

$$\boldsymbol{\varepsilon}^e_{m,b} \equiv \boldsymbol{\varepsilon}^e + \boldsymbol{\varepsilon}^e_b + \boldsymbol{\varepsilon}^e_m = 1_K - 1_M + \boldsymbol{\varepsilon}^e_m. \qquad (4b)$$

This error includes trigger clock drift, slot bias jitter, and sampling errors.[6] Twice differencing the matrix $[\mathbf{c},\mathbf{t}]$ yields (Equation 3c), we filter with theoretical offsets via a *time-window filter:*

$$W[\mathbf{t}|T_L,T_U] \rightarrow \mathbf{t} : \{t\in\mathbf{t}|T_L\le t\le T_U\}. \qquad (5a)$$

This operation filters the resync-time plus sample-clock half-period and sink slot-time biases, which contribute to "total error" via this process. Using the definition from Equation 3c,

$$W\left[\Delta^2\mathbf{t}^e_m\middle|T_L,T_U\right] \rightarrow T_L \le 1_K + \mu(\delta\mathbf{t}^e_m) + \boldsymbol{\varepsilon}^e_m \le T_U, \qquad (5b)$$

where

$$T_L = 1_M + T_1$$
(resync-time lower bound offset by unit), (5c.1)

$$T_U = 1_M + T_u$$
(resync-time upper bound offset by unit), (5c.2)

and

$$1_M \equiv T_{epoch} + T^e_{bias}$$
$$= 1 + T^e_{bias}(\text{model time-unit}), \qquad (5c.3)$$

then offsetting the window filter by T_s yields

	BPSK4	QPSK4	QAM164	AUTO
	1.9983	2.9990	1.9999	2.0000
	1.9983	2.9992	1.9997	2.0003
	1.9983	2.0000	1.9997	1.9996
	2.0002	2.0000	1.9997	2.0000
	2.0014	1.9998	2.0018	2.0004
	2.0002	2.9901	1.9999	2.0002
	2.0000	1.9998	1.9987	2.0020
	2.0001	2.0003	2.0000	2.0011
	2.0000	2.0000	1.9999	3.0003
	2.0003	1.9983	1.9999	3.0000
	1.9989	1.9981	1.9998	2.9995
	2.0004	2.0011	2.0003	3.0003
	2.0000	2.0021	2.0136	2.9997
	2.0019	1.9999	1.9998	3.0001
	2.0021	2.0003	1.9997	2.9998
	2.0174	2.0001	1.9999	3.0002
	2.0018	2.0001	1.9996	3.0018
	2.0020	1.9986	1.9999	2.9991
	2.0019	2.0001	1.9998	2.9997
	2.0021	1.9999	1.9998	3.0000
	2.0000	2.0001	1.9999	2.9986
	2.0001	2.0022	1.9999	3.0003
	1.9980	2.0018	2.0017	2.0002
	1.9981	2.0000	2.0018	3.0003
	1.9983	1.9999	2.0000	2.0017
	2.0019	1.9999	1.9999	2.0019
	1.9998	2.0001	2.0011	2.0016
	2.0001	1.9981	1.9997	1.9999
	2.0002	1.9982	1.9997	1.9997
	2.0000	1.9980	2.0000	2.0002
	1.9985	1.9979	1.9975	2.0009
	1.9984	2.0001	1.9999	2.0002
μ(δt):	2.0006	2.0932	2.0004	2.4691
σ(δt):	0.0033	0.2950	0.0025	0.5067
95% cf:	0.0012	0.1039	0.0009	0.1784
95% CI				
lcl	1.9994	1.9894	1.9995	2.2907
ucl	2.0018	2.1971	2.0013	2.6474
|CI|	0.0024	0.2077	0.0018	0.3567

Anova: Single Factor

SUMMARY

Modulations	*Count*	*Sum*	*Average*	*Variance*
BPSK4	32	64.01898	2.000593	1.11597E-05
QPSK4	32	66.9831	2.093222	0.087053177
QAM164	32	64.01249	2.00039	6.44087E-06
AUTO	32	79.0096	2.46905	0.256747276

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Modulations	4.78031	3	1.593437	18.53813813	5.33544E-10	2.677699
Within Modulations	10.65836	124	0.085955			
Total	15.43867	127				

Figure 1. Data and 95% confidence interval for resync-time statistics for 1 : 1 epoch at 5 milliseconds.

$$W\left[\Delta^2 \mathbf{t}^e{}_m - 1_M \middle| T_L - 1_M, T_U - 1_M\right] \rightarrow T_1 \le \mu^e_m + \boldsymbol{\varepsilon}^e_{m,b} \le T_u. \quad (5d)^7$$

After processing, we have a matrix of 30-plus ($K = 30+$) samples. In Equation 5d, we simplified Equation 3c using Equations 2c, 4b, 5c.3, and

$$\mu^e_m \equiv \mu(\delta \mathbf{t}^e_m), \quad (6a)$$

to get the following:

$$\delta \mathbf{t}^e_m = \mu^e_m + 1_K - 1_M + \boldsymbol{\varepsilon}^e_m = \mu^e_m + \boldsymbol{\varepsilon}^e_{m,b}. \quad (6b)$$

These are the resync-times we seek as a distribution. Also note that delta-counts are counted to filter delta-times (Equation 3c). We use counts to confirm payload, which should approximate the MTU of the Ethernet traffic "tuned" to the modulation-epoch scheme: MTU^e_m. As it turns out, all

$$c \in \Delta^2 \mathbf{c} \approx h MTU^e_m; \qquad h \in \{1, \ldots, 4\}. \quad (7)^8$$

All time-tagged counts are stored in spreadsheet format for easy export for later analyses. We discuss analyses next and then summarize our results in the conclusion.

Data analyses and interpretation

We used Microsoft Excel™ and R (open source statistical program) for standard statistical analyses. We analyze via our *means-model* to derive a Confidence Interval (CI) statistic on $\delta \mathbf{t}^e_m$ for *each* modulation mode *within* an epoch. We then extend the means-model to an *effects-model* for an Analysis of Variance (ANOVA) on $\delta \mathbf{t}^e_m$ for *all* modulation modes *across* epochs. We employ a one- and two-factor ANOVA, respectively, for modulation schemes within an epoch, and modulation within and across epochs.

Confidence and one-way ANOVA

Figure 1 presents CI results of a two-node channel for epoch 1 : 1 at 5 milliseconds. From the table, we can list the resync-times (in milliseconds) with 95% confidence specified to 95% confidence of 31 degrees of freedom, using the simplified nomenclature yields:

$$CI_p(\mu^e_m) = \mu(\delta \mathbf{t}^e_m) \pm q \underline{\sigma}(\delta \mathbf{t}^e_m) / \sqrt{n}. \quad (8a)$$

$$CI_{.95}(\mu^e_m) = \mu^e_m \pm (2.0395) \underline{\sigma}^e_m / \sqrt{31}. \quad (8b)$$

Anova: Two-Factor With Replication

SUMMARY	BPSK4	QPSK4	QAM164	AUTO	Total
EPOCH1:1					
Count	32	32	32	32	128
Sum	64.01897812	66.983097	64.0124855	79.009602	274.02416
Average	2.000593066	2.0932218	2.000390172	2.46905	2.1408138
Variance	1.11597E-05	0.0870532	6.44087E-06	0.2567473	0.1215643
EPOCH10:10					
Count	32	32	32	32	128
Sum	102.0125046	66.984196	72.98419952	70.999996	312.9809
Average	3.187890768	2.0932561	2.280756235	2.2187499	2.4451632
Variance	0.674071263	0.0875402	0.208590696	0.1764699	0.469824
EPOCH4:16					
Count	32	32	32	32	128
Sum	103.8013115	80.018093	64.02531052	64.990597	312.83531
Average	3.243790984	2.5005654	2.000790954	2.0309561	2.4440259
Variance	0.1965016	0.320102	3.6768E-06	0.0311866	0.3881777
Total					
Count	96	96	96	96	
Sum	269.8327942	213.98539	201.0219955	215.00019	
Average	2.810758273	2.2290144	2.09397912	2.2395854	
Variance	0.616249998	0.198685	0.085696221	0.1840861	

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Epochs	7.874879241	2	3.937439621	23.180908	3.253E-10	3.0199871
Modulations	29.23162078	3	9.743873594	57.365156	1.683E-30	2.6289025
Interaction	31.98645619	6	5.331076032	31.385671	1.659E-30	2.1229643
Within	63.18680541	372	0.169857004			
Total	132.2797616	383				

Figure 2. Analysis of variance for resync-time across all epochs.

For our data in a two-node TDMA of 1:1 epoch ratio at 5-millisecond frames, these resolve as follows:

- $CI_{.95}(\mu^{1:1}{}_{BP4}) = 2.0006 \pm 0.0012$
- $CI_{.95}(\mu t^{1:1}{}_{QP4}) = 1.9894 \pm 0.1039$
- $CI_{.95}(\mu t^{1:1}{}_{QA4}) = 1.9995 \pm 0.0009$
- $CI_{.95}(\mu^{1:1}{}_{A}) = 2.4691 \pm 0.1784$

The $CI_{.95}$ results of 4:16 ratio at 5 milliseconds are as follows:

- $CI_{.95}(\mu^{4:16}{}_{BP4}) = 3.2438 \pm 0.1560$
- $CI_{.95}(\mu^{4:16}{}_{QP4}) = 2.5006 \pm 0.1992$
- $CI_{.95}(\mu^{4:16}{}_{QA4}) = 2.008 \pm 0.0007$
- $CI_{.95}(\mu^{4:16}{}_{A}) = 2.0310 \pm 0.0622$

The $CI_{.95}$ results of 10:10 ratio at 5 milliseconds areas follows:

- $CI_{.95}(\mu^{10:10}{}_{BP4}) = 3.1879 \pm 0.2890$
- $CI_{.95}(\mu^{10:10}{}_{QP4}) = 2.0933 \pm 0.1042$
- $CI_{.95}(\mu^{10:10}{}_{QA4}) = 2.808 \pm 0.1608$
- $CI_{.95}(\mu^{10:10}{}_{A}) = 2.2187 \pm 0.1479$

The length of the confidence interval is

$$\left|CI(\mu)_{.95}\right| \approx 2q\underline{\sigma}/\sqrt{n}. \qquad (8c)$$

This interval centers on sample mean μ. The *sample standard deviation*, $\underline{\sigma}$, is an estimate of *population standard deviation* for resync-times; $\underline{\sigma}/\sqrt{n}$ denotes the *standard error* for "n" degrees of freedom, and q is a quantile weighting factor of the residual-distribution.

ANOVA

We ran ANOVA to find any statistical significance between resync-times within and across modulations, and across epochs. *Figure 1* also shows the one-way ANOVA for epoch 1:1 alone. The model is based on Equation 3, which contains one dependent variable, resync-time, and one independent factor, modulation. Note the very small *P value*—an indicator of the probability of occurrence that the variation between modulations is significantly different than the variation within any one modulation scheme. The very small probability of this random event indicates there is a statistically significant difference between resync-time due to modulation. But for practical purposes, the differences are not important. For the two-way ANOVA, we model variation due to modulation and epoch structure, so we extend the mean of Equation 3 to an *effects-model* by extension of the mean to include effects due to modulation and epoch:

$$\mu_m^e \equiv \mu + \alpha_m + \beta^e. \qquad (9a)$$

If we assumed interaction between effects (Equation 9):

$$\delta t_{m,k}^e = \mu + \alpha_m + \beta^e + (\alpha\beta)_m^e + \varepsilon_{m,k}^e; \qquad k \in \{1, \ldots, K\}, \qquad (9b)^9$$

where μ is the *grand-mean*, α_m is the *modulation effect* and β^e is the *epoch effect*, and $(\alpha\beta)_m^e$ denotes the *interaction effect* of modulation and epoch, $\varepsilon_{m,k}^e$ is the error.

In this model, the "effects" cause variation around the *grand-mean,* μ. The two-factor ANOVA compares all modulations and epochs, as shown in *Figure 2*. The data table has been left out, as it is too long to fit legibly. All data are included in an Excel workbook that is available on request.

Conclusions

In the ANOVA results, small *P values* of the analyses indicate that there is a statistical significance with respect to modulation and epoch structure (see row "*P* value" in one- and two-factor tables); i.e., the variation across modulation schemes is statistically significant with respect to modulation within an epoch structure. If we examine resync-time over epochs, the average time is also significantly different with respect to epoch. But these differences have no practical effect.

The small CIs for each modulation scheme suggest our samples are a good estimate of the "unknown population mean" that represents the transceiver resync-time and epoch reacquire-time. Whatever the form of the ANOVA model, we are confident that resync-time never exceeds a duration of minimum slot-time, not to exceed 5 milliseconds.[3] ❑

DANIEL LAIRD received an associate of science from the University of Maryland, College Park, Maryland, in 1985, a bachelor of science in electrical engineering from Rutgers University, New Brunswick, New Jersey, in 1988, and a master of science in electrical engineering from California State University, Fresno, California, in 1992. He has worked over 30 years for the United States Air Force, more than 22 years at the Air Force Flight Test Center as an instrumentation engineer and research and development engineer in the Electronic Warfare Engineering Instrumentation development branch. He worked on the advanced range telemetry and integrated networked enhanced telemetry developments and presently works at the F-22 Combined Test Force. He has published and/or presented papers for International Telemetry Conference and ITEA since 1993. E-mail: daniel.laird@edwards.af.mil

Endnotes

[1]A modulation scheme entails a waveform (fixed or variable), coding, and power level.

[2]Universal Datagram Protocol (UDP) allows for dropped packets.

[3]For more detailed information, on algorithms and models, times bases, and detail of tuning payload, readers are referred to subsections of a 60-page supplementary paper to the author's initial test report. These supplementary documents are available upon request from the author. Supplements available include: (1) Algorithms and Models document, (2) Excel workbook with R results, and (3) R Data Processing and Analyses code.

[4]Ibid., "Time Bases."

[5]Ibid., "Models."

[6]Ibid., "Algorithms and Models," "Time Bases."

[7]Ibid., "Filtering Details."

[8]Ibid., "Details for Tuning Payload."

[9]Ibid., "Models."

ITEA Journal 2011; 32: 132–142

Interoperability Impacts the Warfighter: Organizing to Accomplish the Joint Interoperability Testing Mission

Randon (Randy) R. Herrin
Joint Interoperability Test Command, Fort Huachuca, Arizona
Chris Watson
Defense Information Systems Agency, Arlington, Virginia
With Contributions From
Gene Fenstermacher, Colonel, USAF, Retired

For many years, the Joint Interoperability Test Command (JITC) has directly contributed to the implementation and operations of critical Information Technology (IT) and National Security Systems (NSS) through the execution of complex test events and on-demand warfighter support efforts. From a technical standpoint, the military services and agencies view JITC as the preeminent evaluator of systems interoperability. This article discusses the rich history of the JITC organization and lessons learned gathered as a result of 4 decades of interoperability test support to the joint warfighter.

"The question should not be 'need to know,' it should be 'need to share.' If you approach this issue with the imperative of sharing, rather than restricting information, you literally change the way you approach the whole process." U.S. Army Gen. David Petraeus, Head of Central Command (Cavas 2010).

Army Gen. David Petraeus, the head of Central Command, also noted some of the technical challenges that make it difficult for coalition partners to communicate with one another. "We have located 18 separate command-and-control systems in Afghanistan alone," he said (Cavas 2010). Gen. Petraeous's statement highlights one of the most significant challenges facing the warfighter in Afghanistan today. This vision was realized in the spring of 2010 through the establishment of a single, fully integrated information domain known as the Afghanistan Mission Network. The Afghanistan Mission Network is the primary Coalition, Command, Control, Computer, Communications, Intelligence, Surveillance, and Reconnaissance (C5ISR) network in Afghanistan for all International Security Assistance Force (ISAF) forces. The environment is complex. A single Blue Force Track[1] traversing the network can pass through dozens of systems and be transformed into almost a dozen message formats and protocols. To minimize interoperability issues, the Coalition Interoperability Assurance and Validation (CIAV) working group was chartered to test new national C5ISR applications in an operationally realistic environment prior to fielding.

Joint Interoperability Test Command (JITC) plays two key joint interoperability testing roles in the Afghanistan Mission Network. The first is in leading the overall governance of the network until a formal North Atlantic Treaty Organization (NATO) program becomes operational. The other role is in cochairing the CIAV working group and leading interoperability testing and analyses. Currently, five nations and NATO comprise the CIAV membership, and five to seven new nations will join in 2011. This creates an incredibly diverse analysis environment that requires significant cooperation of up to 12 nations to complete system assurance and validation prior to fielding, resulting in significantly complex, joint, testing scenarios. A key component of testing the Afghanistan Mission Network lies with testing mission areas and mission threads that tell us whether or not a user can accomplish a particular mission and, if not, where the problem lies. In the Afghanistan Mission Network case, JITC developed the Coalition Mission Thread (CMT) test methodology, which focuses on the mission instead of a system or specific country's process. Emerging results have shown that the benefits far outweigh any previous attempts, thus setting the methodology as the way ahead for all allied interoperability and operational development. Through JITC's leadership and support, the Afghanistan Mission Network will change the execution of all future coalition operations. Clearly, JITC did not begin its interoperability testing mission by supporting a complex, coalition warfighter network like Afghanistan Mission Network. Let's now examine the austere evolution of JITC's interoperability testing mission.

Establishment of the Tri-Service Tactical (TRI-TAC)

The interoperability issue has its roots in a coalition problem. Coalition magnifies the already challenging interoperability issues we face in joint operations. The concept of achieving and ensuring communications interoperability began as a multi-national initiative between the United States, the United Kingdom, Canada, and Australia. In the mid 1960s, these four countries participated in a military conference called "Theater Equipment and Logistics" (TEAL), where they recommended the establishment of a program for common tactical communications for these combined military services. These four countries established the Mallard Project in the latter half of 1960s. Ultimately, this international program included the four countries that participated in the TEAL conference. The Federal Republic of Germany was invited to join but did not participate. The Mallard Project office was located in a commercial building in Shrewsbury, New Jersey, adjacent to Fort Monmouth. This was the location of the TRI-TAC office.

Over 130 of the best military and civilian experts from the four countries were assembled to choose an overall design for the system. It was a significant first step in the development of joint and combined operations interoperability, and the Information Systems Security Board (ISSB) chose one major feature that was not part of any of the three studies; integrated Communication Security (COMSEC) with electrical distribution of keying material that became a core capability of follow-on systems.

On October 3, 1970, the deputy secretary of defense announced that the U.S. participation in the Mallard Project would be terminated. At that time, however, he also noted that "the need for better tactical communications remains great." On February 23, 1971, he announced the formation of a field organization to be known as the TRI-TAC Office to help meet major tactical telecommunications objectives. Effective June 30, 1971, Headquarters Army Materiel Command (AMC) formally announced the termination of Mallard, and the transfer of selected personnel to the TRI-TAC Office. The TRI-TAC Office was established in the same facility that had been the home of the Mallard Project.

As an element of the Department of Defense (DoD) TRI-TAC Office, the TRI-TAC Systems Program was responsible for ensuring interoperability between the U.S. Army, Navy, Marine Corps, and Air Force. The initial TRI-TAC program found its home in the Mallard Project commercial building in Shrewsbury. A fundamental premise of TRI-TAC was that military services that tested together, procured together, and fielded together would be more successful when they went to war together with interoperable equipment. The original members of the various TRI-TAC testing teams (U.S. Military Services and the National Security Agency [NSA]) began arriving at Fort Huachuca, Arizona, in 1976.

The Joint Test Element (JTE) served as the testing headquarters for defense contractors, military elements from all Services, and the NSA. In addition, the Services provided their own respective testing groups. By 1982, TRI-TAC JTE personnel in Fort Huachuca and the individual Service groups totaled nearly 400 personnel. When personnel on temporary assignment were counted, the total peaked at over 550. Most test team members worked in the old "Splinter Village" area of Fort Huachuca (primarily World War II barracks), which was approximately 1 mile inside the fort's main gate. Much of Splinter Village was destroyed in a high wind–driven wild fire in July 1987. The U.S. Air Force (USAF) Tactical Communications Test Team owned 12 trailers that were located in close proximity to JITC's current main building on Fort Huachuca. Each of these trailers was used for particular test support activities. There was a Communications Nodal Control Element (CNCE) trailer, a TRC-170 trailer, and so forth—one for each major TRI-TAC test program. Each trailer had a test support noncommissioned Officer-In-Charge (OIC) who reported to the overall Test Control OIC.

The JTE operated the Central Test Facility (CTF), which is what the JITC now calls the Joint Test Facility (JTF) *(Photo 1)*. The JTE also managed a Joint Data Management Office (JDMO), responsible for all of the JTE data collection coding, data management, and analyses taskings.

Under the original TRI-TAC test strategy, the respective equipment procuring Service would act as "Lead Service" for Developmental Test and Evaluation (DT&E) and Operational Test and Evaluation (OT&E) of their respective Joint Service Command,

Photo 1. Tri-Service Tactical (TRI-TAC) TYC-39 Central Test Facility (CTF) with 12 U.S. Air Force (USAF) tactical communication test office trailers, circa 1979 (top); Joint Interoperability Test Command (JITC) headquarters at Fort Huachuca, circa 2010 (bottom). The CTF is now known as JITC's Joint Test Facility and can be seen in both photos above (center of top photo).

Control, Communications, and Intelligence (C3I) development systems. The flaw in this system was that typically the Lead Service would buy the equipment they tested *(Photo 2)*; but because there were generally no joint procurement agreements/requirements, the other Services were usually under no obligation to procure the same equipment and quite often did not do so, citing dissatisfaction with the testing results.

The individual Service testing units began to phase down in 1983. TRI-TAC was certainly a success; but because it could not enforce joint procurement or joint requirements, it could not go far enough to ensure the joint interoperability necessary for a modern joint military.

Designation of the Joint Tactical Command, Control, and Communications Agency

By 1984, there was an apparent need for an agency truly able to test to and implement interoperability

Photo 2. Tri-Service Tactical (TRI-TAC) TYC-39 nuclear, biological, chemical testing, circa 1981/1982.

requirements. Toward this end, the first action was to redesignate the TRI-TAC Office as the DoD Joint Tactical Command, Control, and Communications Agency (JTC3A) in July 1984. The JTC3A was initially made up of personnel from the TRI-TAC Office and was headquartered in Russell Hall near the east entrance to the main post at Fort Monmouth, where TRI-TAC had previously been relocated.

The JTE at Fort Huachuca and the Joint Interoperability Test Force (JITF) at Fort Monmouth became part of the new JTC3A. The JITF used the Joint Interoperability Test System (JITS) for independent testing that supported the development of and ensured compliance with the Tactical Digital Information Links A, B, and C (TADIL A, B, and C) and Message Text Formatting (MTF) message standards managed under the Joint Interoperability of Tactical Command and Control Systems (JINTACCS) program. During this time, extensive efforts were initiated to establish an advanced state-of-the-art JITS called the Distributed Tactical Test Network (DTTN) for testing and certifying the evolving TADIL J Joint Tactical Information Distribution System (JTIDS) message standards. The initial efforts of the JTC3A were to obtain Service agreement for accomplishing joint interoperability testing (what, where, how). Fort Huachuca and the Naval Ocean Systems Center

(NOSC) at San Diego were the leading contenders for the location of the Joint Interoperability Test Center.

In November 1984, the JTC3A provided direction to proceed with a plan to consolidate the agency test facilities at Fort Huachuca in order to reduce the number of DoD testbeds and to reduce overall JTC3A testbed overhead. Under JTC3A, the initial mission of the JTE was to conduct technical interface certification testing as well as provide continued support of other DoD test requirements such as Service/agency Developmental Testing/Operational Testing (DT/OT). The consolidation of TADIL A/B/J, MTF, and technical interface testing centers in one location would allow flexibility to pursue a more comprehensive evaluation of joint and combined C3I interoperability. The existing JTE CTF and a new facility to be built next to the CTF was initially referred to as "Joint *Interface* Test Center," and the total combined resource, personnel, and facilities was initially referred to as the JTC3A Test Center (JTC3ATC). The Joint Interface Test Center name changed to the "Joint Interoperability Test Center" in 1986. The design of the 93,000–square foot JTC3A JITC headquarters building included a 30,000–square foot TEMPEST-controlled facility to house the DTTN and selected equipment. Construction of the building began in June 1987 and was completed in June 1988.

The JTE continued managing DT/OT testing at the CTF throughout the design and construction phases of the new building. During construction, parallel efforts were underway to obtain agreement from the Unified and Specified Commands CINCs (Commander in Chiefs—now, of course, known as Combatant Commanders), Services, and DoD Agencies (C/S/A) regarding how technical and procedural interoperability would be certified. This was no small task as it involved 17 DoD components: 9 Unified and Specified Commands; the 4 military Services; and 4 Defense Agencies (Defense Communications Agency [DCA], NSA, Defense Intelligence Agency [DIA], and Defense Nuclear Agency [DNA]). It also involved the U.S. Coast Guard in the final coordination process.

Evolution of Interoperability Policy

DoD Directive 4630.5 *"Compatibility and Commonality of Equipment for Tactical Command and Control"* was originally established in 1967. It was overhauled, expanded, retitled, and finally republished on October 9, 1985, as DoD Directive 4630.5: *"Compatibility and Interoperability of Tactical Command, Control, Communications, and Intelligence Systems."* It rightfully added a specific new interoperability area, tactical and strategic intelligence system interfaces, to the JTC3A responsibilities. It charged JTC3A with the responsibility to develop and conduct a C3I systems interoperability certification program to verify proper implementation and maintenance of technical and procedural interface standards.

JTC3A Circular 9002, *"Interoperability Certification of Tactical C3I Systems and Equipment Interfaces,"* became the operative document. After exhaustive coordination with the 17 C/S/As, it was finally published on April 21, 1986. The program was established to verify compliance with technical standardization documents such as technical Interface Design Specifications (IDSs) and the MIL-STD-188 series (technical standards for tactical and long-haul communications). On the procedural side, the program verified compliance with U.S. Message Text Format (USMTF) IDSs, JCS PUB 25 (USMTF program), and TADIL procedural interface standards.

The Five Year Interoperability Assurance Plan (FYIAP), JTC3A Plan 3100, evolved as a document to be updated and jointly coordinated annually. Its purpose was to identify and schedule the joint and combined operations interfaces that must be certified as interoperable. It was to include requirements for multi-service DT&E, OT&E, Initial OT&E (IOT&E), Non-Developmental Item (NDI) testing, and Service and agency unique testing. It also included some combined operations issues. Key personnel from the 17 C/S/As and the Office of the Joint Chiefs of Staff (OJCS) convened in June 1986 at Fort Monmouth for the first JTC3A-chaired FYIAP planning group meeting. Results of this meeting initiated the publication of the first FYIAP, which was eventually published on December 17, 1986.

The Goldwater-Nichols (GN) Department of Defense Reorganization Act of 1986, sponsored by Sen. Barry Goldwater and Rep. Bill Nichols, caused a major reorganization of the DoD, the most significant since the National Security Act of 1947. Operational authority was centralized through the Chairman of the JCS as opposed to the Service chiefs. The GN also increased the JCS's role in the Planning, Programming and Budgeting System (PPBS), as well as the development and enforcement of joint interoperability policy and doctrine.

In 1992, the JCS established Chairman, Joint Chiefs of Staff Instruction (CJCSI) 6212.01, *"Compatibility, Interoperability, and Integration of Command, Control, Communications, Computer, and Intelligence (C4I) Systems,"* (July 30, 1993). This document has since been republished numerous times and is now titled *"Interoperability and Supportability of Information Technology and National Security Systems."* The current version, 6212.01E, dated December 15, 2008, establishes the policies and procedures for developing, coordinating, reviewing, and approving Information Technology (IT) and National Security Systems (NSS) Interoperability and Supportability (I&S) needs (CJCSI 2008). It also establishes a formal process for the joint interoperability testing and certification of IT and NSS.

Formation of the Joint Interoperability Test Command

In 1988, the DCA absorbed the TRI-TAC JTE and the JTC3A Joint Interoperability Test Center and kept the name "Joint Interoperability Test Center." DCA consolidated these organizations in 1989 and moved them to Fort Huachuca. The "Joint Interoperability Test Center" name was changed to the "Joint Interoperability Test Command (JITC)," and it has kept this name ever since. With this name change, JITC transitioned its top leadership position from a Senior Executive Service Director to an active duty O-6 Commander, the original intent being to rotate the O6 Commanders among the four Services. The rotational concept has worked well, although JITC has never had a U.S. Marine Corps (USMC) Commander.

JITC personnel began moving into the 93,000–square foot main building (affectionately referred to as the "Taj Mahal" building by Fort Huachuca Army tenants) in June 1989. The building was dedicated by John T. Myers, Lt. Gen., U.S. Army (retired). JITC was developed under DoD Directive 4630.5, JTC3A Circular 9002, and the FYIAP as the primary initial directives and mandates for operational policies and procedures. JITC was also assigned a new mission in 1989: to function as the Operational Test Agency (OTA) for DCA.

The primary mission of the new "JITC" was to provide interoperability compliance testing and certification. As the designated lead for DoD C3I support, DCA tasked JITC to perform interoperability tests of various systems including High Frequency (HF) radio systems, Military Satellite Communication (MILSATCOM) systems, and the Worldwide Military Command and Control System (WWMCCS).

Since JITC's inaugural year in 1989, the Command has gone through numerous phases of modernization and technological changes and enhancements. For example, DCA was renamed "Defense Information Systems Agency (DISA)" in 1992 to reflect its expanded role in managing the Defense Information Infrastructure (DII), now known as the Global Information Grid (GIG). As a result, JITC's responsibilities for ensuring joint interoperability of all military systems began to increase, causing the need for growth and expansion within the organization. In 1992, JITC was also designated a member of DoD's Major Range and Test Facility Base (MRTFB) to provide test and evaluation services to all of DoD, other federal agencies, state and local governments, and private industry.

In 1993, the Naval Computer and Telecommunications Command (NCTC) proposed an initiative to transfer the functions and resources of the Naval Telecommunications Systems Integration Center (NAVTELSYSIC) to JITC. Since 1976, the NAVTELSYSIC test facility had operated out of Cheltenham, Maryland, and it was the primary site for Quality Assurance (QA) and functional certification testing of all Navy messaging systems. DISA and the Chief of Naval Operations (CNO) agreed that the transfer of NAVTELSYSIC resources to JITC would improve the abilities of both agencies to enhance operational fleet support. Thus, JITC's East Coast arm, known as Washington Operations, was established. In 1998, the Washington Operations facility was moved to the Naval Surface Warfare Center (NSWC) at Indian Head, Maryland.

JITC is one of the key organizational elements of DISA and serves as DISA's developmental and operational test organization. As designated by the Joint Chiefs of Staff, JITC is also the authority that certifies that DoD IT and NSS meet interoperability requirements for joint military operations.

JITC facilities are strategically located at Fort Huachuca, Arizona; Indian Head, Maryland; and Fort Meade, Maryland. The diverse capabilities of each respective location allow the Armed Services to have access to a dynamic environment for laboratory tests and on-site field evaluations. DoD organizations around the globe have benefited from JITC's robust test environment and continue to leverage off of their vast resources and technical expertise.

In 2006, JITC established a test group at the DISA facility in Falls Church, Virginia, to support the fielding of global net-centric solutions by providing continuous and effective Test and Evaluation (T&E) services to DISA joint acquisition programs within the Enterprise Services construct. The Falls Church elements began moving to the new DISA headquarters at Fort Meade, Maryland, in April 2011 as part of the Base Realignment and Closure (BRAC).

Today, testers at all JITC facilities work in concert to provide a seamless and unparalleled test capability in support of the warfighter. JITC now consists of over 1,300 government and support contractor personnel located in three different geographical places *(Photo 3)*.

JITC Mission Support Structure

JITC's mission is supported by a blend of military, civilian, and contractor personnel including engineers,

Photo 3. Joint Interoperability Test Command (JITC) headquarters, Fort Huachuca, Arizona (left); JITC Washington operations, Indian Head, Maryland (center); Defense Information Systems Agency (DISA) headquarters, Fort Meade, Maryland (right).

computer scientists, and other technical and operational experts. These personnel provide technical direction, policy decisions, schedules, and program-cost controls in the management of JITC daily operations. JITC is organized to align with warfighting mission areas, provide consistent practices and processes across the entire organization, and support the implementation of a risk-based test strategy that enables fast and agile testing. The organization is primarily composed of divisions and portfolios. Divisions offer business and internal services to JITC to perform its mission, and the portfolios execute the JITC mission and address customer requirements. JITC elements are as follows:

- Operational Test and Evaluation Division,
- Business Management Division,
- Warfighter Support Division,
- Strategic Planning and Engineering Division,
- Testbed Operations Division,
- Facilities Division,
- Enterprise Services Portfolio,
- Focused Logistics and Business Portfolio,
- Force Application/Force Protection Portfolio,
- C2 and Battlespace Awareness Portfolio,
- Battlespace Communications Portfolio,
- National Intelligence Portfolio, and
- Homeland Security/Information Assurance Portfolio.

Achieving the goals of the DoD requires a fundamental change in the way IT is managed in the Department. Historically, IT resources have been managed and acquired as stand-alone systems rather than as integral parts of a net-centric capability. This has had the effect of allowing duplicative investment in systems or platforms that deliver the same or similar capabilities. Managing "portfolios" of capabilities aligns IT with the overall needs of the warfighter, as well as the intelligence and business activities that support the warfighter. In support of this overarching strategy, JITC's T&E methodology is migrating to a capabilities-based evaluation approach that aligns with the vision for IT portfolio management. JITC has aligned its organizational structure to follow the DoD capability portfolios as defined in DoD Directive 7045.20 (DOD 2008). The capability portfolio managers provide recommendations to appropriate DoD decision makers regarding capability investments within their respective functional areas. This alignment allows JITC to provide better direct support to the capability portfolio activities as their role in DoD acquisition continues to grow and evolve *(Figure 1)*.

JITC's "Full-Spectrum" IT Capabilities

JITC facilities serve as the "IT Testbed in the MRTFB" and provide capabilities and infrastructure for end-to-end system engineering and execution of distributed, net-centric testing of core technologies and mission-enabling applications supporting the joint warfighter. The DISA MRTFB is a national asset that provides full-spectrum T&E services in support of the DoD, other government agencies, and industry.

Through their unique expertise and technical resources, JITC is able to provide "Full Spectrum" IT

Photo 4. Panoramic view of the Joint Interoperability Test Command (JITC) Joint Test Facility (JTF), circa 2009.

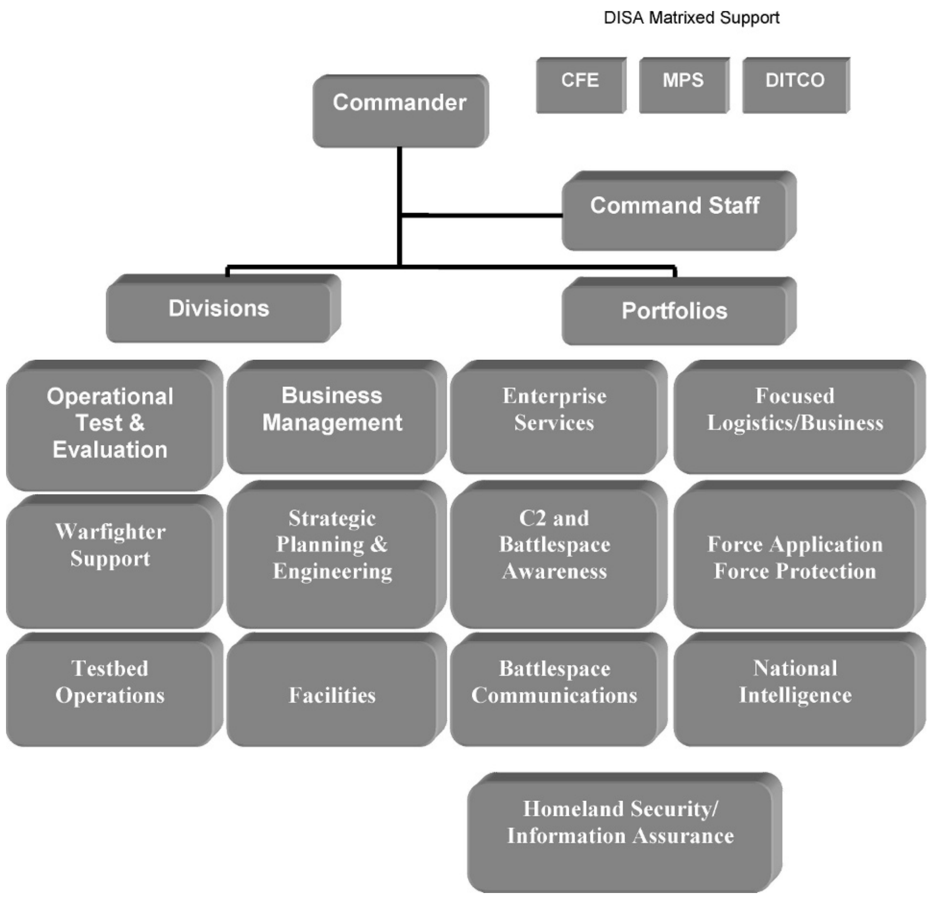


Figure 1. Joint Interoperability Test Command (JITC) organizational structure.

T&E capabilities to homeland defenders and warfighters across the globe. The diverse capabilities and resources associated with each respective location allow the armed services to have access to a dynamic environment for laboratory tests and on-site field evaluations. Over the past four decades, JITC has grown significantly and now manages and operates over 60 separate and distinct testbeds/laboratories at their three locations.

For example, the JITC JTF located at Fort Huachuca *(Photo 4)* is capable of supporting both strategic and tactical testing communities. It sits on approximately 16 acres and contains four test nodes labeled A, B, C, and D. The JTF provides a persistent emulation of DISA's strategic-to-tactical network interfaces with reconfigurable capability. The JTF maintains the systems and services necessary to provide tactical T&E support with strategic interfaces. These interfaces consist of the equipment necessary to emulate a Strategic Tactical Entry Point (STEP) and teleport system and services. The Teleport infrastructure has been upgraded with Synchronous Optical Network (SONET) Dense Wavelength Division Multiplexing (DWDM) capabilities representing the GIG. JTF capabilities are constantly upgraded to ensure they are capable of supporting CJCSM 6231, Operational Area Network (OAN) standards. All nodes are interconnected through the Systems Control (SYSCON) Patch and Test Facility (PTF) using approximately 75 miles of underground cable and fiber optics. The JTF is further augmented by its connection to the Distributed Network Control Center (DNCC), which serves as a gateway to GIG networks and services. In addition, the JTF has connectivity to the Gigabit Network Test Facility (GNTF) for testing Defense Switched Network (DSN) and Defense Red Switch Network (DRSN) switching systems, and to a high frequency test facility, video teleconferencing lab, and others.

The Global Information Grid Network Test Facility (GNTF) consists of several labs that can operate together or separately to test the interoperability of new and modified DSN and DRSN hardware and software. Supporting these primary networks, the GNTF integrates transport, crypto, Customer Premise Equipment (CPE), Voice and Video over Internet Protocol (VVoIP), and Video Teleconferencing (VTC) labs into an architecture closely replicating the

Photo 5. Joint Interoperability Test Command's (JITC's) Advanced Technologies Testbed (ATT) features All-Optical Switching, Dense Wavelength Division Multiplexing (DWDM), Gigabit Ethernet, Synchronous Optical Network (SONET), Gigabit Passive Optical Networking (GPON), and wireless.

Photo 6. Joint Interoperability Test Command's (JITC's) Radio Frequency Test Facility (RFTF) has a Local Area Network (LAN)-based test capability to perform simultaneous, manual, and automated Radio Frequency (RF) parametric testing on multiple radios.

operational DSN and DRSN. The test facility supports testing between strategic and tactical systems and provides voice and data switching equipment to emulate real-world equipment configurations. Occasionally, the GNTF is connected to "exercise" networks, but it is rarely connected to an operational network. The GNTF is also capable of connecting to other JITC laboratories and facilities, or to commercial facilities, and fully supports the Unified Capabilities (UC) testing and certification of numerous emerging technologies.

The Advanced Technologies Testbed (ATT) at Indian Head extends JITC's capabilities and testing infrastructure to include those products developed for increased network speed, primarily through exploitation of the optical realm. The ATT consists of a multivendor architecture that encompasses optical switching (both optical-electrical-optical and all-optical), DWDM, Gigabit Ethernet, Packet over SONET, and Gigabit Passive Optical Networking (GPON). The testbed employs a wide array of protocols and services such as Multi-Protocol Label Switching (MPLS), Internet Protocol (IP) telephony, video streaming over IP, network management tools, and wireless Local Area Networks (LANs) *(Photo 5)*. The ATT has positioned itself at the forefront of communication technology, and the laboratory keeps up with the latest communications breakthroughs, so JITC can "mitigate" the risk of introducing new technology within the Defense Information Systems Network (DISN) and the GIG. It also supports the testing and certification of UC products.

JITC's Radio Frequency Test Facility (RFTF) is used to perform standards conformance testing and radio Interoperability (IOP) testing and certification. The RFTF has a LAN-based test capability to perform simultaneous, manual, and automated Radio Frequency (RF) parametric testing on multiple radios. The facility is equipped with state-of-the art portable and fixed test equipment that allows it to perform all standard RF measurements *(Photo 6)*. Space has been allocated within the RFTF dedicated to the development of the test procedures and eventual conduct of the IW Standards Conformance program.

The Coalition Test and Evaluation Environment (CTE2) is a globally distributed, operationally relevant test environment utilizing the Combined Federated Battle Laboratory Network (CFBLNet) as transport. The JITC CTE2 laboratory supports the interoperability effort for the Afghanistan Mission Network under USCENTCOM. The CTE2 provides an operationally realistic Combined Enterprise Regional Information Exchange System-ISAF (CENTRIXS-ISAF) architecture for the U.S. side. This lab utilizes the transport capabilities of the CFBLNet to connect to other U.S. and Coalition test labs/sites for distributed test efforts. JITC is the U.S. Lead for Afghanistan Mission Network interoperability supporting USCENTCOM and is the CTE2 Manager.

The InterTEC Systems Integration Laboratory (SIL) contains the workstations necessary to support the use of the Joint Mission Environment Test Capability (JMETC) Virtual Private Network riding the Secure Defense Research and Engineering Network (SDREN) and the Defense Information System Network – Leading Edge Services (DISN-LES) *(Photo 7)*. The SIL is JITC's primary area for conducting developmental, unit, and system integration testing of applications in a classified environment.

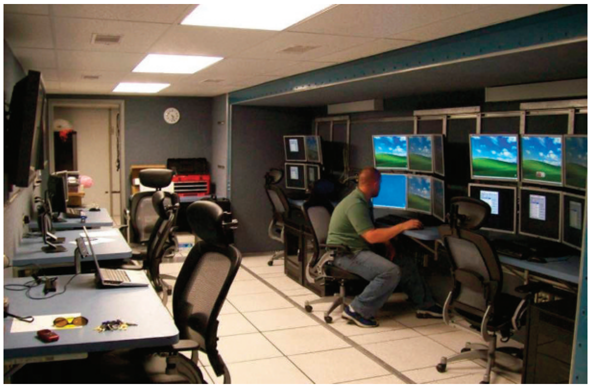


Photo 7. Joint Interoperability Test Command's (JITC's) InterTEC Systems Integration Laboratory (ISIL) is utilized during tests with Major Commands (MAJCOMs), battle labs, services, and Combatant Command.

It is utilized in support of development of an integrated system to test interoperability, while participating in tests with major commands, battle labs, Services, and combatant commanders with the purpose of performing system acceptance and accreditation for use by JITC as an interoperability certification test tool. The SIL also provides a facility where staging of systems prior to sending them to remote sites can be performed, and where system training and demonstrations can be held.

For over 2 decades, the DoD Interoperability Communications Exercise (DICE) has been one of the premier federal government test and training events. The JTC3A conducted the first DICE exercise in 1989 as a response to communications interoperability issues found during Operation Urgent Fury. The exercise focused on ensuring interoperability of mobile switch equipment, tri-service tactical switch, and unit-level circuit switch systems. This initial interoperability exercise has evolved to include tactical and deployed systems employing net-centric technology *(Photo 8)*.

Photo 8. Communications trailers and emergency response vehicles converge on Fort Huachuca to participate in Department of Defense Interoperability Communications Exercise (DICE).

Today, DICE offers an operationally realistic Joint Task Force (JTF) environment that provides the necessary opportunities to vigorously exercise, demonstrate, and evaluate new communications system capabilities. Because of the high demand, JITC now conducts DICE as a tri-annual event, with one major effort occurring in February (6 to 8 weeks) and two other efforts occurring in July and November (spanning 3 weeks each). DICE provides an excellent forum for testing emerging technologies, allied communications initiatives, regressive testing with legacy equipment, and realistic joint communications training. JITC uses multiple laboratory, testbed resources to connect to the DICE distributed environment.

Reform Through Mission-Focused T&E

The 2010 National Defense Authorization Act (NDAA), Section 804, is very clear in its description of IT acquisition reform and the need to adopt "Agile" processes (NDAA SEC. 804. 2010). Section 804 requires a new DoD IT acquisition process that includes early and continual involvement of the user; multiple, rapidly executed increments or releases of capability; early, successive prototyping to support an evolutionary approach; and a modular, open-systems approach. The March 2009 Defense Science Board Task Force Report (DOD 2009) recognized that the development process must be collaborative, iterative, incremental, and test driven. As the Department's IT programs become more agile, all testing and certification processes, including Joint Interoperability T&E, must become equally agile.

As the DoD evolves its acquisition process to deliver new capabilities more rapidly, a requirements generation process must be established that acknowledges and fosters evolving user priorities and an equally agile T&E process. In support of an agile T&E environment, JITC will assemble integrated test teams and Services to execute the "plan-test-report" cycle and provide objective assessments of key technical, operational, interoperability, and security metrics necessary for decision makers to understand capabilities and limitations.

In preparation for IT acquisition reform, over 150 DISA T&E and JITC personnel have been certified as "Scrum Masters." JITC is involved in numerous agile test programs and internal software developments efforts using both "Scrum" methods and Kanban processes and tools.

JITC will also lead the way in constructing a test infrastructure that will connect partner facilities and resources for more collaborative and structured test events. The future T&E environment must evolve from the current disjointed and duplicative test networks to a persistent and operationally realistic test environment. Continuous user involvement, combined with appropriate risk-based, mission-focused testing will ensure that T&E certainly becomes an enabler of rapid acquisition of enhanced information technologies. Ultimately, the new IT acquisition process will become more responsive and beneficial to the warfighter as we influence the future T&E environment. JITC is well prepared to conduct mission-focused testing across all of the IT portfolios and strives to continuously improve efficiencies across DoD's overall distributed joint test infrastructure.

Conclusion

The foregoing discussion covers close to 40 years of efforts that contributed tremendously to the joint and combined operations of tactical and strategic IT and NSS. It covers the digitization era and a resulting, rapidly advancing, technology development period that played significantly in addressing joint and combined operations interoperability and security issues. As you can see, it didn't happen without a lot of tedious high-level coordination and some service parochialisms. Many lessons were learned throughout the "organizing" process, and these lessons continue to affect JITC's support strategies and methodologies today.

JITC's reputation for providing unmatched technical support and expertise can be fully realized by the many success stories and impact reported by the warfighters over several decades. JITC's pursuit of excellence and its rich history will continue to serve as the foundation for conducting mission-focused testing and ensuring interoperability in the battle space for years to come. ❑

Randon (Randy) Herrin is a retired USAF Lt. Col. currently working as a project manager and action officer in the Operational Test and Evaluation Division of JITC at Fort Huachuca, Arizona. Herrin did some helicopter flying but spent most of his USAF career with Research, Development, Test and Evaluation (RDT&E), and systems acquisition Program Management Office (PMO) organizations of C4ISR systems. Herrin has an undergraduate degree in mathematics and three advanced degrees including being a distinguished graduate from the "C4I Systems Technology" master's degree curriculum at the Naval Postgraduate School (NPS). He has been an ACAT I PM for two different C4ISR programs and has possessed his DAWIA Level III certifications in T&E, program management, and systems engineering since the early 1990s. Herrin is widely published in a variety of professional journals and has served in various local chapter and national positions for Information Systems Test & Evaluation (ITEA), AFCEA, and IEEE. He has been the ITEA Publication-of-Year Chairman since 2004 and just this year became the ITEA Education Committee Chairman. In this capacity, he hopes to get NPS, the Air Force Institute of Technology, the Service Academies, and various universities more involved in publishing in ITEA, and more active with the development and instruction of ITEA tutorials at conferences and symposiums and various ITEA short courses. Herrin's first DoD T&E job was working as an Air Force Captain for Col. Gene Fenstermacher (major contributor to this article) at Fort Huachuca in the TRI-TAC era.

Chris Watson serves as the acting National Capital Region (NCR) liaison and chief of strategic communications for DISA T&E. In this capacity, Mr. Watson serves as the DISA T&E executive's principal advisor for defining overarching organizational planning and outreach strategies associated with all T&E elements within DISA. He also serves as acting executive agent to the Military Communications and Electronics Board (MCEB) Interoperability Certification Panel (ICP). He has supported the JITC and the DISA T&E organizations in various capacities for over 22 years and has published over 20 articles associated with T&E processes, capabilities, and events of interest. He serves as the DISA T&E liaison to the Joint Atlantic and Chesapeake Ranges Cooperative (JACRC) and the Federal Laboratories Consortium (FLC).

Gene Fenstermacher, Col., USAF, Retired, was the deputy director, interoperability assurance, at the DoD JTC3A at Fort Monmouth during the JITC concept, planning, and site selection process from 1984 to June 1987 when ground was broken for the facility at Fort Huachuca. He retired (effective July 1, 1987) with nearly 39 years of continuous active duty service. Prior to his JTC3A assignment he was the AFTEC/AFOTEC operational test director for Joint Service testing of Air Force–developed TRI-TAC program systems at Fort Huachuca for nearly 5 years. With a career that began in 1948 with radar maintenance training and instructing at Keesler Air Force Base (AFB), radar installation and maintenance assignments in Japan and Korea (1952/53), and a Radar Bomb Scoring (RBS) system maintenance assignment, he was commissioned through Officer Candidates School (OCS) in 1957. He then went on to a career that included ground radar guided Atlas-D ICBM system crew duty and a missile staff assignment at headquarters SAC, where he was selected for an AFIT electrical engineering degree program at Oklahoma State University, graduating in 1968. Following assignment as a tactical Communications and Electronics (C&E) system advisor to the Royal Thai Air Force in Thailand, he served a tour with the USAF Inspector General (IG), as commander of an engineering and installation squadron, and had a C&E advisor assignment with the Turkish Air Force. He was then assigned to the evolving TRI-TAC

program at AFTEC in 1975, where he was subsequently selected for promotion to Colonel in 1977 and for attendance at the National War College (Class of 1978). He then served as the DCS Engineering and Installation at headquarters Northern Communications Area (NCA) at Rome AFB prior to his assignment to Fort Huachuca in January 1980. Following retirement at Fort Monmouth, he worked at Fort Huachuca as a contractor test engineer at the U.S. Army Electronic Proving Ground (USAEPG) for nearly 15 years before retiring again in 2002.

Endnotes

[1]Blue force tracking is a United States military term used to denote a GPE-enabled system that provides military commanders and forces with location information about friendly military forces. For more information go to http://en.wikipedia.org/wiki/Blue_Force_Tracking.

References

Cavas, Christopher P. 2010. "Petraeus: U.S. must share more info with allies." *Army Times*. May 12, 2010. http://www.armytimes.com/news/2010/05/military_petraeus_sharing_intel_051210w/ (accessed March 15, 2011).

NDAA SEC. 804. 2010. Implementation of New Acquisition Process for Information Technology Systems. P.L. 111-84, The National Defense Authorization Act for Fiscal Year 2010. http://www.wifcon.com/dodauth_10.htm.

ITEA Journal 2011; 32: 143–145

AFOTEC:
Testers Must be Part of the Solution: Begin With the End in Mind and Leverage Ability to Influence and Inform

Maj. Gen. David J. Eichhorn
Commander
Air Force Operational Test and Evaluation Center (AFOTEC)
Kirtland Air Force Base, New Mexico

"The national debt is the single biggest threat to national security," this according to the Chairman of the Joint Chiefs of Staff Adm. Mike Mullen.

A strong military rests on the foundation of a strong economy. I've been working for this nation's defense for over 30 years and I've never seen anything like the current situation. All government workers have to be part of the solution but I think testers have a special role to play.

Testers can inform decision makers on weapon programs that are failing to live up to expectations and will require an inordinate infusion of cash to just be adequate. Given the insight into how a weapon system is really working, acquisition decision makers can restructure, cancel, and/or give more resources to programs. That is where the test and evaluation (T&E) community supplies vital insight.

When Former Secretary of Defense William J. Perry referred to the role Department of Defense test organizations play in acquisition, he said "testing is the conscience of acquisition." As the "conscience" of the system, testers provide an often unwelcomed opinion on the overall health of a weapon system. But that voice in the back of the national skull, warning us that maybe this isn't the smartest thing we could be doing with extremely scarce national resources, is more critical today than it's ever been.

The current state of testing is fractured. Too much dependence on the seller, industry, has led us to accept some really glossy brochures when the weapon itself, if it is ever to live up to that brochure, will take more time, and more money than the original contract. While industry is full of patriots, they are first and foremost, free market, capitalists, and they want their companies to not only survive but thrive. And since many of our pension funds depend on them, we want them to thrive as well. But, we need to re-energize the government's test enterprise.

In today's budget deficit environment, testing is more important than ever. The July 2000 Government Accounting Office (GAO) Report entitled, "A More Constructive Test Approach is Key to Better Weapons Systems," stated commercial firms, "achieve the type of outcomes DOD seeks: they develop more sophisticated products faster and less expensively than their predecessors. Commercial firms have found constructive ways of conducting testing and evaluation that help them avoid being surprised by problems late in a product's development."

The key to industry's success is motivation. Industry is motivated with their commercial products to test early and often to ensure profitability and to minimize liability. For government products, however, politics, not the bottom line, is what gets our weapon systems funded (*no bucks, no Buck Rogers*). The only way the government can match industry's motivation is with the right leaders in the right jobs with the right advocacy above them. Here's a true story to help make my point. Recently, a tester was called on the carpet by his company vice president. The VP wanted to know why the test office's productivity had fallen by 13 percent. The lead tester said, "It's fallen 13 percent in the face of a 700 percent increase in engineering changes to the vehicle. Pretty good, yes?" The tester was sent back to work and the program's lead engineer took his place on the carpet.

In the government test community, it is not about our size, it is not about our expense, it is about our influence and our ability to inform. As a government tester, we're always seen at the end of the process and it's easy to point to us as taking too long. "You're holding up fielding of this critical capability." Sad. We've seen this movie before. In World War II we rushed many systems to the field and forced them to work rather than sending operators what we knew already worked. After the war, Gen. Hap Arnold, Gen. Jimmy Doolittle, Dr. Theodore Von Karman and especially Gen. Bernard Schreiver insisted we won't do that to our fighting forces again.

That same 2000 GAO report quoted previously concluded that, "For [government] testing and evaluation to become part of a constructive effort to validate the maturity of new weapons systems in DOD, the role it plays and the incentives under which it operates must change. Currently, testing and testers are not seen as helping the product succeed but as potential obstacles for moving forward." Contrast that with the attitude after World War II, when, I believe, we had the right incentives.

Testers look at the promises being made about a system and ensure those promises are true or at least, true enough. In order to overcome the perception of being naysayers, early influence, early involvement, and early discovery are absolutely critical. The test community does not want to be seen as Black Hats. Testers have a vital role to play in the development community to get to the end game with a system that works instead of a system that's already bought and we'll make it work once it's fielded.

But also consider that there is far more money to be saved for the nation if failing programs are cancelled early, restructured early, or that other options are considered early (I hope you noticed, early is the key). The government test community provides the national decision makers with vital insight so they can exercise the wisdom of Solomon and bolster the foundation of the national defense.

In the book "The Seven Habits of Highly Effective People," author Stephen Covey describes Habit 2 as "Begin with the end in mind." For testers, the end of our work is the test report. In that report we should highlight when deficiencies were noted and their disposition. If we have influence, if we word it right, and if we ensure we have the big picture, then we ought to be able to follow a test program from start through the report in the end to see if it had the desired effect. If the proverbial "can" was simply kicked down the road with the attitude of we'll fix it later because the program has money and there is no desire to jeopardize the program by highlighting its warts, we've failed. The truth is most weapon systems we buy will not fully live up to their brochures. Only informed, enlightened leadership can decide if that's good enough.

Testers must continue to work early and often and use the tools they are given to inform decision makers in order to support and further the Air Force priority of recapturing acquisition excellence. Government and industry both need to insist on clear, executable specifications and standards. The KC-46A program seems firmly on this path. As defense budgets shrink, mission and business will nose dive unless we work together to execute what's really important with less available resources. In today's budget deficit environment, testing is more important than ever.

DOD can't afford everything it wants, but industry will be pushing for the sale. Everything is a trade off in the sense that if you want this then you can't have that (even at the national level). Additionally, everything is coalition warfare and by that I mean we all have to work this together. No matter what organization I've worked in, it has never had the corner on all goodness. There's lots of goodness throughout government, lots of good people trying to do a great job, and we need to work it together as much as possible and we'll get better products.

Several lessons learned by leading commercial firms were noted in the July 2000 GAO report. The report states, "Lessons were learned when managers ran into problems late in product development that cost the firms money, customers, or reputation. Essentially, they recognized that testing and other validation techniques, along with candor and realism, were instrumental to the success of that product. Thus, the improvement in their practices had more to do with a better appreciation for why testing is done versus how it is done." Industry realized the truth that <u>**testing doesn't cost, it pays!**</u>.

Again, "the national debt is the single biggest threat to national security." One might think that giving decision makers a reason to cancel or improve a program and ultimately save money would be welcomed. Well, maybe not. Testers are caught in the middle. Only a solid fact and value-based decision process will work. Test results have to stand the test of time and be enduring in their conclusions and sometimes they even have to stand up in court—a huge responsibility. Testers have to work well with acquisition decision makers as we uncover sometimes unflattering weapon system characteristics. Our goal as a test community is to be great stewards of the taxpayers' money and to give warfighters the best and most affordable weapons systems designed in the world today. ❑

MAJOR GENERAL DAVID J. EICHHORN is the Commander of the Air Force Operational Test and

Evaluation Center at Kirtland AFB, N.M. Major General Eichhorn reports to the Air Force Chief of Staff regarding the operational test and evaluation of more than 76 acquisition programs valued at more than $650 billion being assessed at 12 different locations. He directs the activities of more than 625 civilian and military people, as well as 225 contractors. As a member of the test and evaluation community, General Eichhorn coordinates directly with the offices of the Secretary of Defense and Headquarters U.S. Air Force while executing realistic, objective, and impartial operational testing and evaluation of Air Force, coalition, and joint warfighting capabilities. Major General Eichhorn has served as an experimental test pilot, and his commands include two flight test squadrons, a test group, a test wing, the Arnold Engineering Development Center, and the Air Force Flight Test Center. A certified acquisition professional, he served at the Electronic Systems Center as the Vice Commander, where he was previously assigned as Director of Advanced Command, Control and Communications Systems, as well as Director of Advanced Aircraft Systems. He has also served as Director of the Aeronautical Enterprise Program Office and Deputy Program Executive Officer for Aircraft at the Aeronautical Systems Center. At Headquarters Air Force Materiel Command he was Deputy Director of Plans and Programs, and Director of Air, Space, and Information Operations. He has flown the B-52D/H, B-1B, F-111 and T-38, serving as an instructor pilot and aircraft commander. He has accumulated more than 6,100 hours in more than 47 aircraft types. E-mail: David.Eichhorn@kirtland.af.mil.

References

Covey, S. R. 1989. *The seven habits of highly effective people*. New York, New York: Free Press Publishers.

United States General Accounting Office. 2000. *Best practices: A more constructive test approach is key to better weapon system outcomes* (GAO/NSIAD-00-199). Washington, DC: Government Printing Office.

ITEA Journal 2011; 32: 146–148

OPTEVFOR:
Organizing to Accomplish the OT&E Mission (Integrated Testing Through Mission Based Test Design)

RADM David A. Dunaway
Commander Operational Test and Evaluation Force (COMOPTEVFOR), Norfolk, Virginia

Test and evaluation (T&E) is a subject that has frequently generated both emotion and sometimes, confusion, especially within the context of shrinking budgets when the hot topic of the day is cutting T&E costs. Although OT&E costs for most programs are only 0.5 to 1.5% of total program costs, OT efficiencies have been and will continue to be pursued. At the same time, efficiencies must not come at the expense of valid and relevant assessments of operational effectiveness and suitability. There is most certainly a hidden cost associated with shortchanging up front work in developmental, integrated, or operational testing when a system fields that does not work. One need only imagine a multi-billion dollar warship tied up to the pier vice on deployment because of a software glitch in the combat system. In response to the need for operational test (OT) efficiencies and better early involvement, OPTEVFOR has incorporated several internal process improvements over the last few years. The most significant of these include the mandatory use of integrated testing (IT), mission based test design (MBTD), and use of design of experiments (DOE). The ultimate intent is to ensure that a "minimum but adequate" test is performed. The term "minimum" implies the use of all qualified data, elimination of redundancy, and combining efforts wherever possible. The term "adequate" implies that all factors critical to operational effectiveness and suitability are explored to a level of confidence that is practical and relevant.

Current COMOPTEVFOR policy mandates 100% participation in IT; this is incorporated into every program document that is staffed. Although there are still some previously planned programs executing suboptimal IT plans, every effort is made to find test efficiencies within the confines of signed, documented agreements as OT&E periods approach. While great strides have been made in the understanding of IT, the unfortunate fact is that although many programs do practice the proven methods of IT, many still do not. In the past, the aversion to IT may have been originated with the DT, OT, or Program communities. With the implementation of the recent COMOPTEVFOR policy, at least one of the three key entities has eliminated any ambiguity when it comes to the commitment to integrating test teams.

Maximizing IT is enabled via MBTD. In short, MBTD is a process by which every platform, weapon, network or sensor (PWNS) is related to the operational mission thread in which it must perform. MBTD provides a clear, traceable means of describing the necessary contribution of a PWNS within a system of systems (SoS) in an effort to create a warfighting capability. During the process, OPTEVFOR staff, along with Fleet and program subject matter experts (and DOT&E for oversight programs), systematically decompose every mission thread that is required per the particular acquisition program documentation. These mission threads are broken down into tasks, sub-tasks, attributes, and conditions to create a comprehensive integrated evaluation framework (IEF). This process includes an analysis of how the system under test operates within the SoS so that we can understand how to adequately test the required warfighting capability. The MBTD process includes the use of design of experiments (DOE). The latter techniques are used to reduce the comprehensive OT matrix to the "minimum, adequate" test possible. Ultimately, what constitutes a "minimal and adequate" test is further driven by balancing costs, mission requirements, test resources, and DOT&E's assessment of test adequacy. Once minimal and adequate testing is documented in the IEF, the plan is closely coordinated with DT to avoid any unnecessary, repetitive testing thereby optimizing the use of test resources. One of the fundamental principles of IT is early involvement. Having OT participation very early in a program is key to early identification of critical issues. The value of early involvement is only limited by the maturity of the system. One of the key tenets of OT, instantiated in law, is the use of production representative equipment. It is an unfortunate artifact that data gathered on early, pre-production systems, although valuable, cannot be used as the ultimate factor in determination of effectiveness or suitability. Although careful attention is paid to avoiding repeated testing, the determination of operational effectiveness and suitability requires final system maturity and configuration.

The Department of Defense (DoD) acquisition programs have been greatly influenced by current overseas operations as well as potential future threats. The current fiscal environment will most certainly limit the amount of funds available to DoD in the execution of our national security. The increasing sophistication of the threats we confront and the increasing complexity of our own systems mean the demands for greater integration and interoperability are higher than they have ever been. The mandate for affordable and capable systems demands that the test community work at the peak of efficiency. Although there are some examples of exemplary behavior, there are, unfortunately, examples of less than optimal behavior. In general, sub-optimal behavior is routinely characterized by programmatic cuts in DT, which ultimately means there will be major deficiency discovery in OT. The pressures on programs to meet schedule at promised cost can, and frequently does, equate to a decision to reduce early testing that almost inevitably leads to discovery in OT. The near-term argument that a complex test set up is too expensive is unfortunately dwarfed when a multi-billion dollar platform is procured but cannot be deployed to conduct the Nation's bidding because of an integration issue. When IT is performed with a clear and mutually understood construct of how the PWNS plays within the greater SoS, our warfighters get the greatest return on taxpayer investment.

Some recent examples of both good and bad behavior are included below:

- FA-18 Operational Flight Program (OFP) testing: The H-6E variant OFP, Airborne Electronically Scanned Array (AESA) Radar FOT&E 2, and 23X variant OFP testing were combined into a single FOT&E resulting in a 60% savings of nearly $8.5M over stand alone costs. In addition, four out of four required missile firings were conducted as IT resulting in improved data collection and representing an additional $4.8M of savings for the weapons alone. Although this effort is a top priority of the Integrated Test Team (ITT), these synergies are in jeopardy of being lost in the event that the current program schedules slide. Although a comprehensive IT strategy was carefully planned, the unfortunate truth is that if one of the participating programs loses funding, has a schedule slip, or fails to meet its goals, the entire IT plan could unravel and significantly increase cost.
- Both the Advanced Anti-Radiation Guided Missile (AARGM) and SM-6 development programs are examples of cases where financial and schedule pressures have driven program managers to push events that clearly belonged in the verification phase of DT into the validation phase of OT.
    - AARGM: OT had intended to use the last three AARGM DT shots to reduce the number of shots required for IOT&E by that number (saving $824K/shot). Unfortunately, developmental testing uncovered major deficiencies. Software and hardware changes implemented after the shots were taken precluded the use of the earlier events for OT purposes due to scope of the configuration changes.
    - SM-6: The SM-6 is a highly complex system that has a total test requirement of 17 shots agreed to by DT, OT, and the Program office. Because of schedule constraints, only 5 of the 17 shots will be conducted as DT leaving 12 shots in OT. This programmatic decision has greatly increased the likelihood that major issues will be discovered in OT. Although all parties agree that this solution is suboptimal, the program views the breach of schedule milestones as a higher risk to program success.
- Past FA-18 IT scenarios on H2E, H3E, and the Growler were ultimately successful in combining multiple test events and using DT data for OT resolution. When program maturity is adequate and configurations remain stable, up to 50% of an OT test event can and frequently is eliminated as redundant testing.

In too many areas today, we are not delivering effective and suitable systems to the Fleet. The root cause is often the late discovery of significant problems. The bottom line is that when we have the discipline to execute our known best practices we minimize the cost to the taxpayer and the adverse impact to warfighting capability. Good IT with clear communication between all stakeholders is essential to program success. Too often the pressures of the day force suboptimal decisions early in a program's development that ultimately land squarely on the backs of deploying

Sailors, Airmen, Soldiers, and Marines. We, in the test community, are both stewards of taxpayer dollars and warfighting capability. We must have the discipline to refine our best practices and execute them so that minimum resources are expended for maximum warfighting capability. Minimum, adequate, and early must be a mantra that we inculcate into our testing genome. ❑

REAR ADMIRAL DAVID DUNAWAY was born in El Paso, Texas. He received his wings in April 1984 and subsequently served as a selectively retained graduate flight instructor in Meridian, Mississippi. After completing FA-18 initial training, he served in VFA-151, aboard the USS Midway in Yokosuka, Japan, from 1986–1989, when he was selected for the U.S. Naval Test Pilot School Class 96, Patuxent River, Maryland.

Dunaway's test assignments include: VX-5 as the A-12 operational test director; F/A-18 branch head (during this tour, he was selected as an aerospace engineering duty officer); F/A-18 Weapon System Support Activity as the deputy for Test and Evaluation; and, VX-9 as the F/A-18E/F operational test director. In this position, Dunaway flew more than 200 developmental test missions and was selected as the Test Pilot of the Year.

His program management assignments include: PMA-265 as the F/A-18 Radar IPT lead for the APG-79 Active Electronically Scanned Array radar, for which he and his team received the 2003 Aviation Week and Space Technology Laureate Award in developing this state-of-the art radar; PMA-201 as the program manager for the Precision Strike Weapons program office, for which the JSOW program received the David Packard Award for innovative business practices; and, most recently, as the deputy program executive officer, Air Anti-Submarine Warfare, Assault and Special Mission Programs.

Dunaway served as the Commander of the Naval Air Warfare Center, Weapons Division China Lake and Point Mugu, California, and the Naval Air Systems Command Deputy for Test and Evaluation from September 2007 until January 2009. He currently serves as the Commander, Operational Test and Evaluation Force in Norfolk, Virginia.

Dunaway is a Class of 82 graduate of the U.S. Naval Academy and holds a bachelor of science degree in mechanical engineering, a master of science degree in aviation systems management from the University of Tennessee and a master of science degree in aerospace engineering from the Naval Postgraduate School. His personal decorations include the Legion of Merit, Meritorious Service Medal, Navy Commendation Medal, and the Navy Achievement Medal. He has accrued more than 2,900 flight hours and 290 arrested carrier landings.

ITEA Journal 2011; 32: 149–158

JITC: Accelerating the Nation's IT Dominance

Colonel Joe Puett, United States Army
Commander
Joint Interoperability Test Command, Fort Huachuca, Arizona

Unfortunately, the gap continues to grow between the pace at which information technology (IT) speeds to market in the commercial industry and the slower pace by which IT is fielded within Department of Defense (DoD). To meet this challenge and close this gap, it is not enough to just keep pace with the delivery of commercial IT; DoD must strive to accelerate the pace at which value-added warfighting "capability" (the combination of technology, doctrine, tactics, and training) is placed in the hands of warfighters. All DoD organizations involved in acquisition are seeking and implementing new or improved organizing principles in their procedures, their structure, and their training. The Joint Interoperability Test Command (JITC) is one such DoD Agency that is executing a number of lines of effort designed to improve the delivery of capability and accelerate the Nation's IT dominance.

As designated by the Joint Chiefs of Staff, JITC, a major field activity of the Defense Information Systems Agency (DISA), the DoD organization that certifies DoD IT and national security systems (NSS) meet interoperability requirements for joint military operations. JITC also serves as the DoD's Joint Operational Test Agency (OTA), and the responsible test organization (RTO) for various DoD program offices. JITC also plans, directs, and executes a variety of test activities outside the bounds of formal developmental testing and operational testing. For instance, JITC works closely with the warfighting combatant commanders (COCOMs) during exercises and contingency operations to provide them on-the-spot evaluations of problem areas and viable mission-oriented solutions. JITC has deployed teams to Afghanistan and Iraq on multiple occasions to resolve IT/NSS issues in theater. JITC's core competency is our expertise in IT/NSS interoperability and application of that expertise wherever it is needed, inside or outside the test lab.

In addition to the support provided to the military services, JITC supports organizations such as the National Geospatial Intelligence Agency (NGA), the National Security Agency (NSA), and the Federal Emergency Management Agency (FEMA).

JITC's lab facilities form the central element of DISA's Major Range and Test Facility Base (MRTFB). The DISA MRTFB is DoD's only non-Service and IT-focused MRTFB. The organization's global reach extends to the entire spectrum of DoD, Federal government, private industry, and allies in support of command and control, intelligence, and defense reform initiatives. As an MRTFB, JITC can deal directly with vendors to provide critical pre-acquisition test results. This early involvement in development enables faster acquisition, better systems, and lower cost.

JITC emulates IT/NSS operational architectures in their test facilities, ensuring interoperability issues around the globe can be reconstructed and addressed remotely. The command's facilities are strategically located at Fort Huachuca, Arizona; Indian Head, Maryland; and Fort Meade, Maryland (Figure 1). The diverse capabilities of each respective location allow the Armed Services to have access to a dynamic environment for laboratory tests and on-site field evaluations. DoD organizations around the world have benefited from JITC's robust test environment and continue to leverage their vast resources and technical expertise. Today, testers at all JITC facilities work in concert to provide a seamless test capability in support of the warfighter.

JITC'S value-added operational test & evaluation

JITC serves as one of the five Operational Test Agencies (OTAs) within the Department of Defense.

Figure 1. JITC's Fort Huachuca Laboratories form a portion of DISA's Major Range and Test Facilities Base (MRTFB).

In this role, JITC conducts operational testing of IT/NSS acquired by DISA or other DoD/non-DoD Federal organizations to determine the operational effectiveness, suitability, interoperability, and security of a particular IT system under realistic operational conditions. JITC independently determines if thresholds in the approved capability production document (CPD) and critical operational issues are satisfied; and assesses the operational impact to the warfighter. JITC is solely focused on IT operational assessment (unlike the Service OTAs which evaluate equipment, weapons, vehicles, aircraft, etc., as well as conduct live fire testing).

With the migration towards more agile development and acquisition of DoD IT capabilities, JITC is continually evolving our operational test policies and processes to proactively support evolving IT acquisition reforms. The Office of the Secretary of Defense in a November 2010 report to Congress entitled "A New Approach for Delivering Information Capabilities in the Department of Defense," stated that "the traditional acquisition process used to develop and acquire military technology is not aligned with the speed, agility, and adaptability at which new IT capabilities are introduced in today's information age."[1] The report outlined new principles to the departments approach as:

- Deliver Early and Often
- Incremental and Iterative Development and Testing
- Rationalized Requirements
- Flexible/Tailored Processes
- Knowledgeable and Experienced IT Workforce

Although being agile truly starts with the programs acquisition strategy and development, JITC has continued to refine test methodologies to help assist the acquisition community in meeting these objectives. We have adopted the Capability Test Methodology (CTM)[2] to create mission based evaluation strategies which facilitate more relevant evaluations with traceability to mission impact. This also facilitates a capabilities based risk assessment in which events can be tailored to the most critical missions and only those affected by the software changes being proposed. One significant change is applying scientific test design principles, such as design of experiments (DOE) to identify major factors and utilizing those to drive the conditions in which testing needs to be conducted. This allows for truly integrated tests in which data can be collected once and used by all; significantly reducing schedule and costs.

Interoperability

Interoperability means different things to different people. Joint Publication 1-02 defines interoperability as the ability to operate in synergy in the execution of assigned tasks and a condition achieved among communications-electronics systems when information or services can be exchanged directly and satisfactorily between them and/or their users. For IT, interoperability means that warfighters can publish/subscribe and send/receive usable, accurate, timely, and secure information to make informed decisions.

While "interoperability" might be easy to describe, it is not necessarily easy to achieve given DoD's acquisition model. Current acquisition, development, integration, and testing processes are funded, structured, and designed by individual system and program. Ideally, interoperability is designed into the system at its earliest stages. A particular system is envisioned operating within a set of capabilities and is designed and built to operate within that context. Unfortunately, too often interoperability is an afterthought to system design. While the reasons for this vary, the fact remains that the further downstream you are in the development of a particular system when you realize that it must interoperate with one or more others, the more costly (in terms cost, schedule, and performance) are the changes needed to achieve that interoperability.

JITC is involved in a number of lines of effort associated with ensuring that systems are interoperable. These lines of effort include: joint interoperability certification of classes of IT products based on Joint Staff approved Unified Capabilities Requirements (UCR); influencing DoD policy by which systems are envisioned and acquired to ensure that "capabilities-based" approaches are the norm; getting involved earlier in the acquisition of particular technologies so that interoperability is a forethought, not an afterthought; and establishing exercise and test venues so that systems at various levels of maturity can be tested for interoperability.

Figure 2. JITC conducts unique interoperability tests of advanced communications technologies during DICE.

One example of an organizing construct that JITC has undertaken to establish synergy in the delivery of interoperable IT capability is JITC's tri-annual DoD Interoperability Communications Exercise (DICE). DICE provides both a testing and a training opportunity for Services, Agencies, COCOMs, allied nations, and vendors. DICE is a major JITC initiative supported by the Joint Forces Command and the Joint Staff and replicates a geographically dispersed Joint Task Force (JTF) environment for the purpose of Joint Interoperability Certification or assessment (Figure 2). DICE also provides an excellent forum for testing emerging DoD technologies, allied communications initiatives, and regression testing with legacy equipment. The continuing integration of emerging technology into DoD interoperability certification testing is critical to enhancing JITC's overall operationally realistic testing capability.

A second example of JITC's interoperability efforts is the mission-based testing approach we are undertaking associated with the multi-national effort of the Afghanistan Mission Network (AMN). AMN stemmed from an unanticipated requirement in which mature systems and applications, never envisioned or designed to work together, were suddenly brought together by circumstance. Here, "after-the-fact" interoperability is the only viable approach, yet there still needs to be an organizing framework around which that interoperability is achieved. The key component of testing the AMN lies with testing mission areas and mission threads. It is through this type of testing where we can tell whether the system as a whole can accomplish a particular mission and if not, where the problem lies. In AMN's case, we've developed the Coalition Mission Thread (CMT) testing, which focuses on the mission instead of a system or specific country process.

Figure 3. JITC participates in Global Interoperability events such as Combined Endeavor to assist in streamlining communications.

Warfighter support

A third core area in which JITC adds value to DoD is in its direct warfighter support mission. In the execution of its interoperability certification and operational test missions, JITC has gathered, cultivated, and retained a workforce of highly trained IT and NSS interoperability experts. The command leverages this team of experts in direct support of the warfighter in two distinct ways. Initial support is provided remotely, via JITC's 24/7 Interoperability Hotline (literally a 1-800 number), where warfighters directly contact experts on JITC's staff to resolve their interoperability problems. JITC is able to leverage its configurable and distributed laboratory facilities (or in partnership with other Service labs remotely access additional facilities) to replicate, identify, and resolve even the most complex interoperability issues. Advanced follow-on support is provided around the world when required, through deployment of highly trained and specialized JITC Tiger Teams.

JITC has seen an increase in the number of coalition related interoperability issues being experienced by the COCOMs. JITC has long supported the COCOM's multinational interoperability events such as the Endeavor series of exercises hosted by EUCOM, AFRICOM, and PACOM, and is developing the means to share the results of these events with the rest of the DoD community (Figure 3). The Endeavor exercises allow COCOMs to determine the capabilities and limitations of specific IT systems of nations within their areas of responsibility (AORs). JITC provides detailed test reports (in the form of a searchable database) following each exercise to provide COCOM planners an easy to use and efficient tool to conduct theater communications planning for contingency operations. JITC is expanding this capability by linking all three of these Endeavor databases. The Global

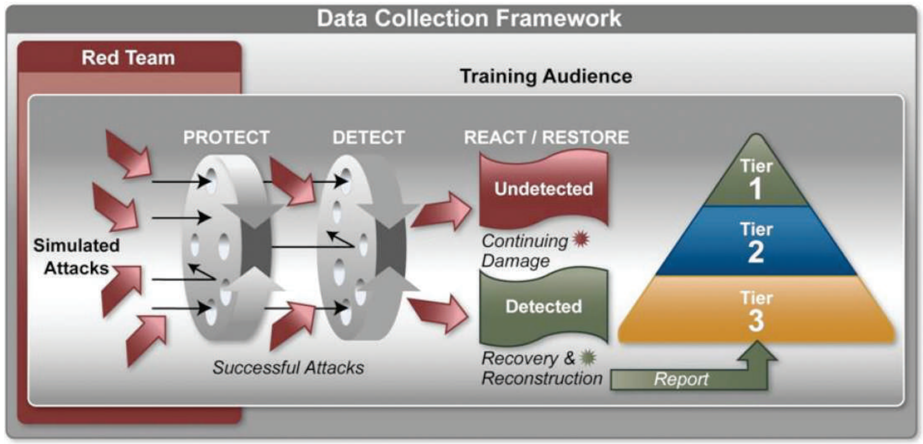


Figure 4. JITC's Information Assurance Data Collection Framework.

Communications Interoperability Program (GCIP) is a fully searchable, online database that contains results from all previous Endeavor events encompassing equipment from over 80 countries. The intent in sharing this information is to provide a tool that will improve the ability of U.S. communications to plan and implement an interoperable coalition C2 network on demand.

JITC continues to provide direct technical and operational related interoperability guidance to the COCOMs, Services and Agencies is support of their day-to-day operations, exercises and real-world contingency missions, generally supporting 15–20 events and exercises each year. Today, testers at all JITC facilities work in concert with the Services and program managers to improve functionality and ensure interoperability of the IT systems being fielded to the warfighter.

Information assurance

A fourth core area in which JITC adds value within DoD is in the area of information assurance (IA) testing. JITC ensures IT and NSS acquired, assigned, or managed by DOD are able to protect and defend the warfighter critical information and information systems. Central to this is the test and evaluation (T&E) of these systems within a Joint Network Operations (NetOps) context. JITC conducts realistic cyber performance assessments during acquisition developmental testing and operational T&E (OT&E). JITC constructs real-world performance based assessments to ensure comprehensive IA capabilities of a system (to include the people, processes and technologies of the system) are able to detect and respond to cyber threats (Figure 4).

JITC, in support of DOD Chief Information Office (CIO) Unified Capabilities (UC), assesses the IA posture of commercial capabilities to support their placement on the DOD UC Approved Product List (APL). Pre-selection testing reduces overall test time and resources while it ensures the warfighter can rapidly purchase and implement commercial capabilities with the assurance they already meet DOD IA requirements and protect critical missions. JITC developed the common test plans, procedures and report formats for the IA testing under the UC program. All DOD Components use these common test products to ensure a common set of measures and results. Through the Principal Accrediting Authority (PAA) Reciprocity Memorandum, these results can be shared among Designated Approving Authorities (DAA) in the DoD Information Assurance Certification and Accreditation Process (DIACAP).

JITC, in support of United States Strategic Command (USSTRATCOM) and DISA's Program Executive Office-Mission Assurance (PEO-MA), ensures the performance of capabilities critical to NIPRNET/SIPRNET hardening. This suite of capabilities is designed and implemented as the first layer of defense to critical DOD mission information. The command's assessments ensure a cohesive and effective boundary layer detection and protection capability across the DOD enterprise.

In partnership with Joint Staff, DISA Field Security Operations (FSO), and the National Security Agency (NSA), JITC developed the Joint Common Vulnerability Assessment Methodology (JCVAM) to ensure vulnerability assessments conducted as part of the connection approval process contain common measures and metrics, thereby improving information and assessment sharing to facilitate a 'test once use by many' construct. The Joint Staff has required use of the JCVAM in its recent update to joint staff policy.

JITC serves as the IA T&E subject matter expert to the DOD IA Range and extensively uses the IA Range infrastructure in support its evaluation mission. The DOD IA Range incorporates DOD Defense in Depth (DiD) design principles which provide us with a methodical, repeatable, and verifiable Cyber T&E capability to measure (quantify and qualify) network defense capabilities to integrate people, operations, and technology, to protect, monitor, detect, analyze, diagnose, and respond to cyber security attacks. The IA Range provides a realistic global network defense environment and architecture that is segregated from the operational environment. JITC uses the IA Range to measure and evaluate critical DISA global network defense programs prior to implementation within the DOD enterprise.

Finally, JITC IA testing also supports COCOM readiness assessments and operational exercises. At the request of a COCOM, JITC undertakes operational performance assessments of the COCOM's complete network defense capabilities to conduct successful cyber defense operations. These assessments build upon the JITC performance-based assessment methodology to put the network defenders into war gaming exercises with active threat teams to measure and quantify their ability to detect, diagnose, react and respond.

JITC's organizing initiatives

The exigencies of the current DoD warfighting and acquisition environment require JITC to seek new organizing principles for their structure, processes, and infrastructure. JITC's initiatives center around our efforts to converge our T&E network infrastructure, apply agile methodologies to testing within the IT acquisition process, improve workforce training, and apply risk-based, mission-focused testing approaches within a capabilities oriented testing framework.

Converged T&E network and virtualization services

Central to achieving efficiencies within DoD's testing infrastructure are efforts to combine and consolidate multiple testing networks. JITC is supporting T&E network convergence to consolidate the multiple circuits required to transport various research, development, certification and coalition networks operating on the DISN and commercial leased lines. The initial effort will converge the DISN-Leading Edge Services (LES) with the Combined Federated Battle Lab Network (CFBLNet) into a common transport called the DISN T&E Transport Network (DTEN) (Figure 5). This convergence will achieve economy of scale by eliminating duplicate circuits and transport equipment at sites currently operating with both DISN-LES and CFBLNet. Additionally, the DTEN will employ agile provisioning and connection approval processes which will support agile IT acquisition processes. Once the DTEN is established, services will be offered to non-DISA T&E networks such as the Joint Mission Environment Test Capability (JMETC), Army Federation of Accredited Network Sites (FANS) and Navy Distributed Engineering Plant (N-DEP).

The DTEN will operate on DISN Internet Protocol (IP) independently from NIPERNET and SIPRNET to provide transport services to multiple test communities. The figure below illustrates the overall proposed topology of the DTEN. The DTEN will use operationally expired data to test systems and capabilities prior to deploying a particular system or program onto live operational networks. Separation of the test enclaves (known as communities of interest (COIs)) from the DISN operational traffic will be accomplished through the use of NSA approved Type 1 IP encryption devices and Multiprotocol Label Packet Switching (MLPS) technology. MLPS will allow the DISN IP Core to bring quality-of-service features available in a circuit based network to the connectionless IP world.

The DTEN network is being established to improve transport services, connection approval, and provisioning of testing COIs. The T&E Customer Support Center (TE-CSC) will provide the customer support necessary to provision distributed test capabilities using the DISN in an agile fashion. Effective monitoring and utilization of the network will be managed through the use of service level agreements (SLAs) and performance reports provided to the COIs. The connection approval process will be tailored to support the non-operational dynamic nature of IT system development with maximum use of reciprocity between Information Assurance Managers within the COI.

Security enclaves will be established on the DTEN to meet the network requirements of the COI being supported. These enclaves can range from unclassified up to Top Secret with both Coalition and US Only data releaseability. Depending on the requirement, cross domain solutions (CDSs) can be used to permit the exchange of data between enclaves. Additionally, approved security solutions are available to allow controlled access to operational networks to allow the transfer of data that may not be available within the COI required to test capabilities.

The DTEN Distributed Development Test Enterprise (DDTE) COI provides virtualized core capabilities for defense common ground/surface systems (DCGS) enterprise testing. Key DCGS services are virtualized by installing DCGS integration backbone (DIB) components on virtual machines and operating systems. JITC recently virtualized DIB-to-DIB web service transactions for content discovery and retrieval

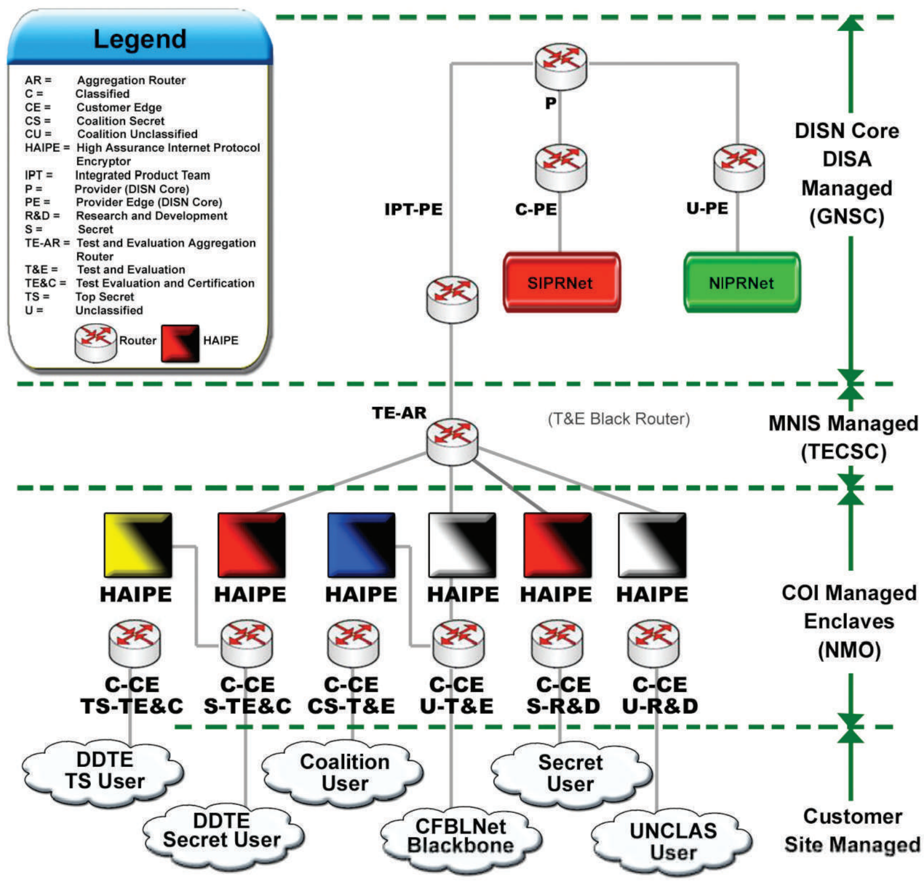


Figure 5. Proposed DTEN Topology.

(CD&R). Deployment of virtual DIB players in a distributed test network that federates DCGS functionality is ongoing. These virtual DIB systems can be federated as if they were operational players, allowing for more realistic operational environments to be simulated during test and evaluation.

Agile approaches and their impact on T&E

For over ten years commercial industry has successfully leveraged Agile development processes in support of the delivery of IT. Smaller, more frequent delivery "sprints" have ensured that the highest value, most important requirements (in the eyes of the customer) are delivered in weeks instead of years. Acquisition risk (risk that the development effort will not produce its expected result) has been more rationally reduced and proportioned over smaller periods of time. Current DoD IT acquisition reform initiatives seek to achieve similar objectives. Development and adoption of standardized methodologies that support rapid IT deployment cycles is clearly a required foundation of the DoD move toward Agile Development. JITC is leading the effort in defining new DoD testing concepts (based on established commercial testing concepts) that keep pace with rapid IT development.

Testing by way of Agile concepts is an accepted software testing practice that follows the principles of the Agile Manifesto being used as commercial best practices. While it may not always fit the DoD environment, implementing these concepts in DoD acquisition has the potential to drastically improve quality and decrease time to the battlefield (Figure 6). To do this within DoD, we must develop, socialize and endorse formal guidance that removes the separation between current test teams (development testing, operational testing, interoperability testing, security testing) and instead endorses a single team concept with a single goal: rapid deployment of enhanced military capabilities.

Enabling this single team approach are improved collaboration and information exchanges such as: informal chat-and-develop modes, short, time-boxed daily standup meetings (for asking questions and resolving problems), and weekly iterative review meetings to track the progress for each iterative cycle.

Another enabler is the promotion of light-weight "user stories" in place of ponderous acquisition

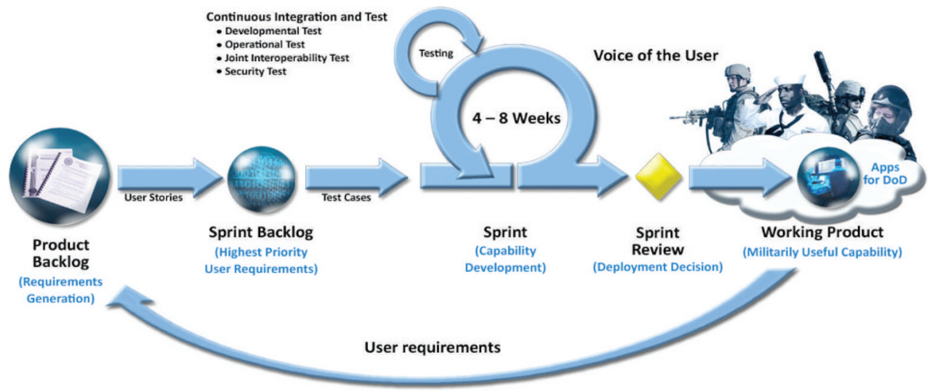


Figure 6. An Agile IT Development Model for DoD.

documentation. JITC has embraced the Agile methodology by piloting the user story concept in development of some of its test cases, test cases written before the code is even written. This is known as "Test Driven Development" a highly successful practice in the commercial sector that truly achieves "early involvement" of testers. These stories become the foundation of repeatable test cases that can be re-used for multiple test events across a number of programs and capabilities. Another critical element of agile testing is the necessity to provide immediate, prioritized feedback so that the developer and integrator can make changes during the sprint. This ensures delivery of working software, with fewer defects, at the "Speed of Need."

Workforce training

Methodologies and processes involved in adopting a more Agile T&E model within the DoD require a corresponding investment in the human capital. DISA and JITC are undertaking an effort to evaluate the skills that exist in today's T&E environment, how those skills must change to allow for DoD application of Agile IT acquisition, and what the appropriate path is to achieve that training (Figure 7).

The first major step in adopting these skills is to break down the developer-tester wall. While there are contractual limitations to how far we can "reach back" in the cycle when industry is developing IT, it is in our collective best interest to increase knowledge and understanding across all disciplines. The individual responsible for development of one small component of a system must fully understand the operational environment in which the system will be fielded. This will allow them to overlay the "effectiveness, suitability, interoperability, and security" perspectives into each decision. Likewise, an increased engineering and scientific background in all facets of the T&E community will allow a more thorough understanding of the capabilities, limitations, and potential test scenarios as the system or component is developed.

The second major step is to infuse the T&E community with a greater instrumentation capability. Any steps toward reducing the man-in-the-loop process on repeatable tests increases the accuracy of that test. By focusing our limited human capital resources on development of test scripts for use across multiple programs, and the analysis of the data collected during the test, we present decision-makers with not only accurate *data*, but useful *information*.

In order to accomplish these two steps, DISA/JITC is embarking on a partnership with both academia and industry. Our efforts include in-depth training on commercial best practices, hands-on internships with industry, and both academic and commercial certifications of our test community. This will be followed by an adaptation of those training efforts into the DoD

Figure 7. DISA and JITC employees take advantage of educational opportunities associated with "Agile" practices and concepts.

training pipeline through the Defense Acquisition University and internal training programs. The partnership will allow us to quickly insert industry lessons-learned into our formal training structure.

An important by-product of this process is that industry and academia, as our partners, gain a further understanding of the DoD environment, that focuses not only on performance but also on operational effectiveness, suitability, interoperability, and security. This provides the DoD with its engineering and test professionals of the future.

JITC is working towards becoming an Agile Command. We have established a JITC Agile Center of Excellence (JACE) with Agile coaches and over 100 Certified Scrum Masters (testers certified in Agile Development methodology) to support three different customer bases: operations within JITC, DISA and its product lines, and DoD systems undergoing interoperability and operational testing. Agility at JITC is based upon a JITC Agile Architecture (JAA) approach. The JAA, along with the Scrum methodology, helps to capture and catalog a system's intended performance (via well-organized user stories).

In Agile, testing becomes even more important because it becomes fully integrated into the development sprint in coordination with systems' users, developers, and product owners. Testing (developmental and operational, information assurance and interoperability) is no longer managed as separate types by separate organizations, but rather fits together in a single framework occurring in concert. This change in the testing "stove-pipe" paradigm requires a change in the way our testers are taught and certified—a change that JITC is fully embracing.

Mission-based testing within a capabilities-oriented framework

Responding to the demands for more responsive and cost effective OT&E support to programs of record (PoRs) that are experiencing ever-increasing budget constraints, JITC has adopted several proactive and innovative approaches. Through early program involvement, JITC influences and ensures the testability of technical and operational requirements by developing a comprehensive analysis structure from which we can evaluate operational effectiveness, suitability, interoperability, and security from a mission/task perspective, throughout the development lifecycle. Applying the Capability Test Methodology (CTM)[2] JITC works with program managers, system engineers, and users to develop Critical Operational Issues and test measures tied to DoD doctrine-based Joint Capability Areas (JCAs) and Universal Joint Task Lists (UJTLs) (Figure 8). Based on the applicable

Figure 8. JITC applies the Capabilities Test Methodology (CMT) during onsite laboratory tests and field assessments.

JCAs and UJTLs, we work with the user community to develop operational mission threads (e.g., scenarios) which exercise the system-provided capabilities in an operational context. The early establishment of an overall analysis structure enables the development of a sound test and evaluation strategy. JITC uses our in-house-developed Joint Capability Analysis Structure Tool (JCAST) to aid test directors in the creation of mission-based analysis structures. JCAST provides a structured framework to systematically guide evaluators through the analysis structure development process with prompts for defining the data elements associated with each test measure. Efficient test events can be designed to gather the specific data elements needed under the specific conditions which the data must be generated and collected. Correctly designed test events generate the right data, under the appropriate operational conditions, and are usable by all testing stakeholders with reduced time and resources.

Applicable portions of the analysis structure can be applied to incrementally developed program capabilities as well. Since the test measures are mission-based, they are solution agnostic. This aspect greatly enhances re-use and can be applied to capabilities developed with Agile processes. This provides cost and schedule efficiencies because only a single analysis structure has to be developed and maintained. JITC performs risk assessments to "right-size" the scope of testing required for incrementally or agile developed systems. By mapping back to the analysis structure, risk assessments focus on the mission areas and associated tasks potentially impacted by additions, changes, and/or enhancements to previously tested system capabilities. When mission critical tasks are potentially affected by system changes, the risk assessment dictates more rigor in testing. For those mission areas and tasks

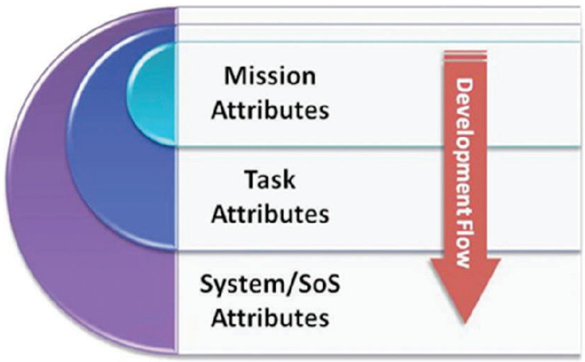


Figure 9. Joint Mission Thread Attribute Levels.

not affected or minimally affected, the risk assessment dictates significantly less testing. Cost and schedule efficiencies are realized, while maintaining the necessary test adequacy.

Based on assessed risk, the resulting "right-sized" evaluation analysis structure for the increment is ported into the JITC Data Management Tool (JDMT). Also developed in-house, the JDMT is a web-based capability that helps JITC test directors efficiently perform test execution, data collection, analysis and reporting functions. During test execution, the web-accessible JDMT provides data collectors a mechanism for test data entry, and provides stakeholders with near real-time visibility of data collection status and incident reports for authentication and scoring. We intend to enhance the JDMT with automated data analysis capabilities which will help to significantly reduce post-test data correlation, analysis, and reporting time.

These products provide an authoritative Joint framework to streamline programs' requirements definition processes, facilitating fast and agile testing, and improving assessment of operational impact when shortfalls are identified (the 'so what' factor). JITC is pursuing several activities in support of CPM efforts to provide Joint governance, as well as evolving a standard methodology to identify the systems and capabilities required to support Joint Mission Threads (JMTs) (Figure 9).

JITC's efforts to assess systems within the context of mission threads and Joint architectures supports the principles outlined in a recent Office of the Secretary of Defense Section 804 report to Congress:

- Incremental and Iterative Development and Testing
- Rationalized Requirements
- Flexible/Tailored Processes

Bottom line: The early establishment of an agreed-upon, mission-based analysis structure is key to enabling successful integrated testing, efficient re-use of test data and scenarios, and credible mission/task-based risk assessments. JITC's JCAST facilitates analysis structure development, and JDMT reduces the time to formally report test results. Collectively, these JITC efforts give us the ability to be more responsive to Agile-developed programs and more efficient in our testing support.

Summary

Within DoD, we find ourselves in an incredibly challenging environment. Warfighters are understandably demanding increased IT capabilities that they see being rapidly and routinely provided in commercial industry; at the same time, DoD is seeking to achieve budget efficiencies while also strengthening our national security capability. All DoD organizations, to include the JITC, are looking for better ways of meeting this challenge. We see the way ahead as a combination of several lines of effort in which we ensure a proper focus on investments and expenditures that are directly tied to the ever-increasing needs of the warfighter. In JITC we continue to focus on and improve our core capabilities: operational T&E, interoperability, warfighter support, and information assurance. Secondly, we have adopted a number of organizing principles related to our structure and procedures: converge T&E network infrastructure, apply agile methodologies to testing within the IT acquisition process, improve workforce training, and apply mission-thread based testing approaches within a capabilities oriented testing framework. The imperative as we look to future conflicts…. JITC: Experts in testing and certification, accelerating the Nation's IT dominance! ❑

COL Joe Puett is the commander of the Joint Interoperability Test Command and an information systems management officer (FA53). Upon graduation from West Point in 1982, COL Puett was originally commissioned as a field artillery officer and from '83–'90 he served with D Battery 4/325th Airborne Bn Combat Team, Vicenza, Italy and with the 319th Airborne Field Artillery Regiment, 82nd Airborne Division, Fort Bragg. During '92–'95, he taught mechanical engineering in the Department of Civil and Mechanical Engineering at the U.S. Military Academy. Next, he served as a NATO computer security and accreditation officer on the staff of the Supreme Allied Commander Atlantic (SACLANT), Norfolk, Virginia. From '98-'00, as an exchange officer to the United Kingdom, he helped develop the British Artillery's next generation field artillery command and control system. From '03–'05 COL Puett taught informa-

tion technology in the Department of Electrical Engineering and Computer Science, again at the U.S. Military Academy. From '05–'10, he served as the Inspector General (IG) and later as the ACofS G5 of NETCOM/9th SC (A) and as the Chief of the DISA Support Element-Afghanistan during Operation Enduring Freedom (OEF). He is a graduate of Ranger School, Jumpmaster School, the Armed Forces Staff College, and the Army War College. COL Puett holds a doctor of philosophy degree in software engineering from the Naval Postgraduate School, a masters degree in mechanical engineering from CalTech, a masters degree in business administration from Long Island University, and a masters degree in strategic studies from the Army War College. He is a licensed professional engineer (P.E.) and a certified information systems security professional (CISSP).

Endnotes

[1]Office of the Secretary of Defense, "A New Approach for Delivering Information Technology Capabilities in the Department of Defense", Report to Congress, November 2010

[2]The Office of the Secretary of Defense chartered the Joint Test and Evaluation Methodology (JTEM), Joint Test and Evaluation (JT&E) in 2006 to develop methods and processes for testing in a Joint Mission Environment. The Capability Test Methodology (CTM) and its associated products represent the outcome of this effort towards developing improved methods and processes for assessing the contribution of a system or System of Systems before it is fielded

ITEA Journal 2011; 32: 159–166

Development of a Torsional Thrust Balance for the Performance Evaluation of 5-N Class Thrusters

Marjorie A. Ingle, Jesus R. Flores, Nathaniel Robinson, and Ahsan Choudhuri, Ph.D.
Center for Space Exploration Technology Research, University of Texas–El Paso, El Paso, Texas

Development of high performance micropropulsion technologies will greatly benefit next-generation interceptors and microsatellite missions. Miniature interceptors (multiple kill vehicles, RPG countermeasures, etc.) and microsatellites (space-based radar formation) are constrained in mass, envelope, and electrical power. Therefore, all spacecraft subsystems, including the propulsion systems for orbital maneuvering (divert propulsion), attitude control, and station keeping, have to be miniaturized. However, accurate performance evaluation of N-class thrusters requires the development of low-thrust measurement technologies. This article presents design, development, and preliminary testing of a Torsional Thrust Balance (TTB) for performance evaluation of miniature thrusters. The thrust balance allows a measurement range between 25 mN to 4.5 N and can be configured for both steady-state and pulsing operations. Additionally, the TTB can accommodate a wide range of thruster weight by compensating the moment arm length with a mass counterbalance system. Displacement measurements are made using a triangulated laser-positioning sensor, the output of which is converted to thrust via a LabView interface. The use of frictionless pivots, a moment arm, and counterbalanced-stand design offset the influence of measurement complications such as environmental interference, natural vibration frequencies of the system, and tare forces. In future studies conducted with this system, uniformity of the reference positioning of the thruster for each firing will be ensured with the use of a one-sided damping system used in conjunction with a limit sensor to both reposition the thruster arm and trigger the thruster firing. The accuracy and precision of the TTB was tested using a 0.5-N reference cold gas thruster. Thrust measurements were shown to be accurate and repeatable.

Key words: Development; experimental setup; low-thrust measurement; micropropulsion technologies; miniaturization; preliminary testing; tare; thrust measurement testing; vibration.

The measurement of low-thrust levels is a challenging task for a variety of reasons spanning from the fact that the thrust levels themselves are so minute that the measurement devices required for such analysis must possess great resolution to the fact that such measurement systems require the incorporation of components that are highly responsive to environmental sources of error. Although it could be considered virtually impossible to completely eliminate sources of tare and inconsistency, experimental devices and setups can be designed to reduce the impact of such occurrences. The focus of this study is to develop a single component Torsional Thrust Balance (TTB) system that characterizes the thrust capabilities of a 5-N class thruster by determining axial force production due to cold gas pressurization of said thruster. Research has been conducted extensively in this area; however, most of the test stands employed for this purpose fall into several main structural categories, all with advantages and disadvantages. In this article, a variant of the uniaxial horizontal configuration will be discussed. Similar to the relatively simplistic linear uniaxial design, the TTB measurement system developed at The Center for Space Exploration Technology

Best Undergraduate Student Paper from the ITEA Live Virtual Constructive Conference in El Paso, Texas, January 2011.

Table 1. Requirement for the torsional thrust stand.

Thrust	25 mN to 4.5 N
Chamber pressure	5 psi to 200 psi
Mass flow	1e-5 kg/s to 4.5e-5 kg/s
Volume flow	2 L/min to 15 L/min
Measurements	Thrust, chamber pressure, and mass flow
Frequency	Steady state to 8 Hz
Thruster weight	50 g to 0.8 kg
Resolution	5% or less of full scale

Research will only have one degree of freedom but will minimize the influence of tare forces through its construction features.

Runyan, Rynd, and Seely (1992) describe the advantage of a horizontal thrust stand, namely the reduction in gravitationally produced tare. Not only is a TTB capable of reducing the influence of gravitational tare through the use of a counterweighted system, the fixation method and housing structure of the beam supporting the thruster also eliminate a phenomenon known as the pendulum effect that is common to horizontally-suspended axial thrust measurement systems. The authors further state that there are three main categories under which tare forces can be classified and either be accounted for or mitigated. These classifications include tare due to deflection, interaction, and gravitational forces acting on the measurement system. The TTB system utilized in this investigation addresses factors that fall within each of these three categories. To be specific, methods were employed to reduce deflection tare produced in the gas feed lines, interaction tare was reduced by limiting the contact between the torsional support beam housing and the propellant feed line, and balanced fixation method of the support beam limited the influence of gravitational forces acting on the thruster during the testing process. These methods will be described in greater detail in the following sections, and their validity and effectiveness will be evaluated by the quality of the thrust measurements obtained through the TTB measurement system.

Requirements overview

The design and development of the torsional thrust stand system was centered on meeting the criteria detailed in *Table 1*. The system is capable of measuring thrust ranging from 25 mN to 4.5 N without necessitating significant alteration of the structure aside from counterweight adjustments. The parameters measured include displacement, which was correlated with a thrust force value, chamber pressure, and mass flow in postprocessing data. In addition to these requirements, it was determined that the data acquisition sampling rate needed to be 10 times faster than the natural frequency of the thrust stand/thruster assembly to completely capture the entire thrust stand response. The chamber pressure and volumetric and mass flow are considerations for a 0.75-mm throat cold gas thruster, and its performance will serve as the benchmark for future testing of the thrust stand.

Technical approach

Mechanism description

The thrust measuring system consists of two frictionless pivots coupled to a housing attachment for the thruster load transfer beam, as shown in *Figure 1*. A load transfer beam to the pivots serves as both a moment arm and weight counterbalance. The entire assembly is Aluminum 6061 except for the frictionless pivots, which are constructed from stainless steel. Cup-point set screws attached to the beam housing, along with flats machined on the side of the flex pivots, are the standard locking method and prevent slipping of the pivot inside their housing. The moment created about the center of rotation because of the weight of the thruster on the measuring side of the beam will be offset by a counterweight to achieve static balance (Stark 1969). In future studies, the moment arm will be dampened on the thrust direction side to minimize backlash about the reference position when the thruster is shut off. The dynamic response of the moment arm will be monitored and recorded by means of a laser triangulation sensor; however, the scope of these preliminary tests will only include recording the distance travelled during firing. These measurements were then correlated to a thrust force through weighted calibration of the stand prior to experimentation. The calibration process involved the correlation of thruster displacement measurements to a series of applied loads ranging from 0 to 0.5 N with 5-gram increments.

This iteration of the thrust measuring system was assembled with a total torsional spring constant of 15.02 pounds-in/deg as a starting point. Initial calculations of total moment of inertia, natural frequency, and displacement under different loads were compared with a computer model using ADAMS (Automatic Dynamic Analysis of Mechanical Systems) (MSC Software Corporation, Santa Ana, California). In order to meet the requirements with the greatest accuracy possible, this comparison was made to ensure that the simulation was in fact capturing the actual response of the assembly. Analysis of this information will be used in future studies to optimize the spring constant, moment arm, and damping coefficient. The natural frequency determines the stiffness of the system and establishes the maximum-forced frequency that the TTB can accurately measure. A very stiff system can measure higher-forced frequencies; however, the

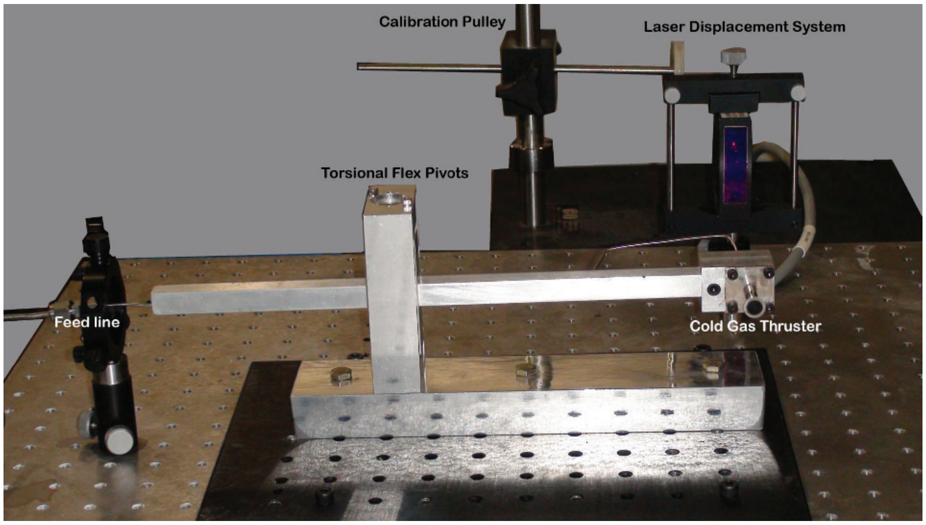


Figure 1. Torsional balance thrust stand description.

accuracy is negatively influenced because of the resolution of the laser triangulation system. Given its large moment arm displacement, a more lenient system measures force more accurately, however, it will oscillate about an arbitrary value at high thrust profile frequencies. It appears that this value is without behavioral significance (i.e., it is not correlated to a specific kind of dynamic response). As explained in following paragraphs, damping the return of the moment arm will be crucial for faster zero value return and therefore will provide more accurate results in continued experimentation. The damping constant value will be determined from the optimization of the system by simulation and implemented into the experimentation with a miniature linear damper.

The cold gas flow was delivered to the thruster through a feed line system that extended from the CO_2 cylinder through a control valve, pressure gauge, 0–30 L/min mass flow meter, anchored pipe diameter reduction interface, thruster/feed line interface, and finally to the thruster chamber and nozzle. To limit the influence of line expansion and/or oscillation during the initial pressurization of the cold gas feed lines on the laser displacement measurements, a solid 1/16-inch stainless steel line was fed through the load transfer beam and affixed to the thruster interface at one end and to a feed line of 1/4-inch diameter by means of a tubing reduction connection on the other end. Mass flow readings were obtained through a National Instruments NI SCC-68 DAQ system at a rate of 375 Hz. Displacement readings were recorded in LabView and measured using a Micro-Epsilon optoNCDT 1402 laser-positioning sensor. The TTB stand, measurement devices, valves, gauges, and portions of the feed line were fixed to an optical table to minimize measurement errors caused by table or component vibrations. For the safety of the investigators performing the experimentation, a plastic shield and fan were used to direct and disperse the CO_2 flow exiting the thruster. A weighted pulley system was used during the calibration process and will be described in greater detail in the latter section describing this process.

Theory and governing equations

To describe the response of the thrust stand for preliminary calculations, the equation of motion for a free vibration of an undamped spring-mass system was assumed (Rao 2004), i.e.,

$$I\bar{\theta} + K_t\theta = 0. \quad (1)$$

The moment of inertia I can be approximated for each of the components on the stand. The summation of these calculations is then used to find the total inertia as follows:

$$I_a = \frac{1}{12}m\left(w^2 + d^2\right) + md_a^2, \quad (2)$$

$$I_b = \frac{1}{12}m\left(w_0^2 + d_0^2\right) - \frac{1}{12}m\left(w_i^2 + d_i^2\right) + md_b^2, \quad (3)$$

$$l_c = \frac{mr^2}{2} + md_c^2, \quad (4)$$

and

$$l_t = l_a + l_b + l_C. \quad (5)$$

With this approximation of the moment of inertia, the natural frequency of the system can be computed as

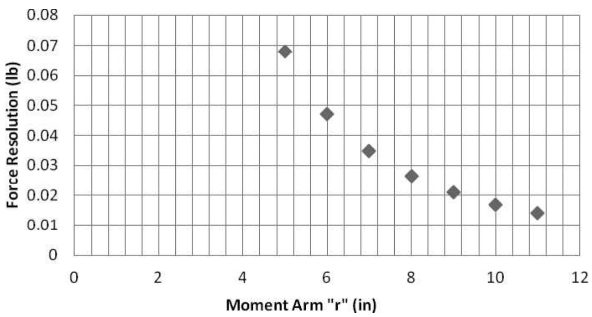


Figure 2. Resolution and moment arm.

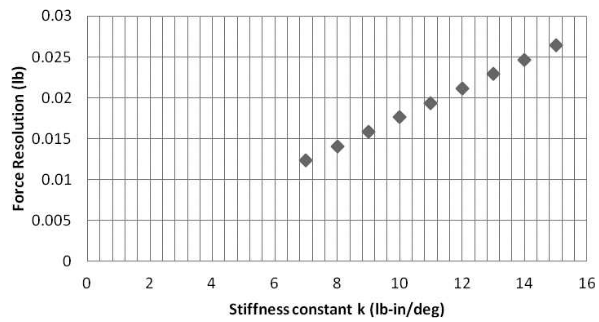


Figure 3. Resolution compared with stiffness factor.

$$\omega_n = \frac{1}{2\pi}^{2}\sqrt{\frac{K_s}{l_t}}. \tag{6}$$

Projected performance

As stated earlier, the dynamic response of the thrust stand under the load of the thruster was recorded with a laser-triangulation system. The recorded linear displacement was correlated to thrust through calibration previous to testing. The laser triangulation sensor used is the OptoNCDT Model ILD 1402-100, and it has a dynamic resolution of 12 to 50 μm. Using the worst resolution scenario (50 μm), the current force resolution of the thrust stand for a static thrust is projected to be 0.0265 pounds. This number depends on the stiffness factor constant and the moment arm of the thrust stand and was calculated using an 8-inch moment arm and a torsional stiffness constant of 15.04 pounds-in/deg.

A longer moment arm allows for increased deflection of the system which increases the force resolution, the consequence being a greater moment of inertia and lower natural frequency. The relationship between force resolution and moment arm length is not linear *(Figure 2)*. The minimum distance that the laser can detect corresponds to an angle deflection depending on the moment arm length. This is then used to compute a minimum force detected, referred to as the force resolution. The numbers shown here are for a constant torsional stiffness constant of 15.04 pounds-in/deg with the moment arm varying in length from 5 to 11 inches.

The stiffness factor can also be used to control the resolution of the torsional thrust stand. Less stiffness allows for greater angular deflection of the stand which results in greater resolution. Keeping the moment arm constant at 8 inches and varying the stiffness factor from 7 to 15 pounds-in/deg, it is evident that a softer pivot yields greater resolution with the disadvantage, again, of smaller natural frequency. The relationship between the stiffness constant and force resolution is linear *(Figure 3)*, and it can be seen that the force resolution is more sensitive to the moment arm than the stiffness constant. The numbers shown are for a constant moment arm length of 8 inches. A trade needs to be made between the stiffness factor and moment arm length to yield the required accuracy of 5 percent resolution while keeping the natural frequency large enough to precisely describe the dynamic response of pulsing thrust.

Pretest calibration process and verification

In order for the displacement measurements to be converted to thrust values it is necessary to perform a calibration involving the application of a force such as a measured weight and then measure the amount of displacement such a force produces. During the course of this study five calibrations were performed to verify the accuracy of the displacement/thrust correlations. The calibration method employed was fairly simplistic, but as shown in the following data, proved to be quite accurate. A pulley system was first placed on the opposite side of the TTB system from the thruster nozzle and facilitated the connection of a lightweight nylon fiber from the circumference of the thruster chamber to a small basket carrying the intended applied loading. An initial displacement measurement was taken to simply show the distance between the laser sensor and the thruster interface at the equilibrium (no applied load or thrust) position. Then a displacement measurement was taken after applying the load of the empty basket. The measurements included the displacement resulting from the addition of mass in 5-gram increments that produced force increments of 49.05 mN.

One might easily make the mistake of overlooking the applicability of calibration measurements to the evaluation of the performance of TTB or similarly purposed devices; however, accuracy and consistency in calibration measurements are important verifications of the sensitivity and accuracy of displacement, and therefore thrust measurement. For the purpose of this

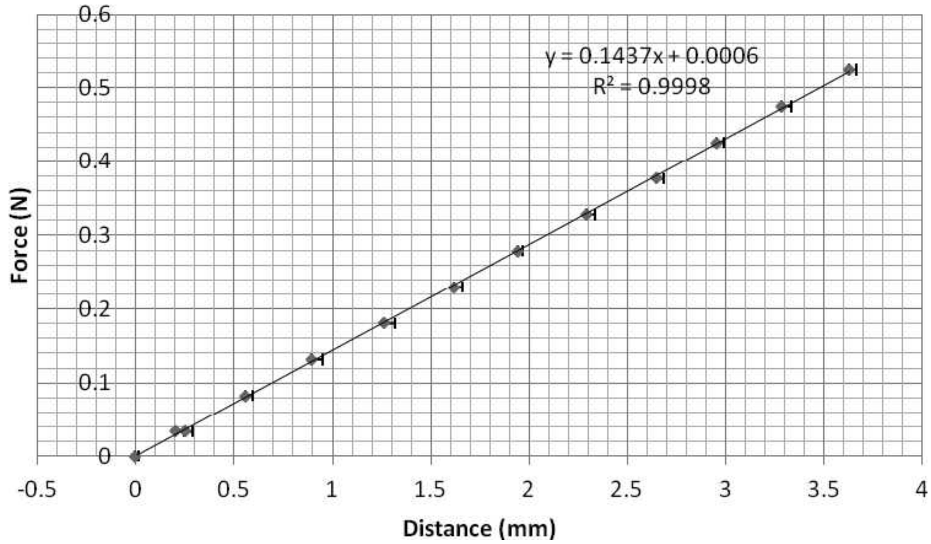


Figure 4. Calibration results.

study, five calibration measurements were taken in the aforementioned manner and produced readings that were highly accurate. In fact, the uncertainty analysis reveals very small uncertainty for the thrust stand considering a confidence interval of 95 percent with an α value of $\alpha = \pm 2.5$ percent (a total of 5 percent when considering both sides of the Gaussian plot). The absence of hysteretic behavior is another indication within the calibration data that the TTB measurement system possesses high accuracy. The displacement caused by the application of the empty basket load was measured before the graduated application of force. It was measured again after all the increments had been applied, recorded, and removed, the result being that the displacement produced was exactly the same (within the decimal range of our measurement devices) as the first time. This is significant in that it proves that any hysteresis behavior that may be observed in an actual cold gas firing of the thruster is due to sources independent of the structural behavior of the TTB test stand itself, therefore proving the integrity of the TTB system design. *Figure 4* is the graph of the calibration results. It shows the range of displacement measurements produced for each level of the graduated force application and also describes how much the results from each calibration differed from the others. It was from graphing the average of these results that a linear relationship was developed to convert the displacement measurements into thrust. As shown in *Figure 4*, the linear equation used to obtain the thrust values reported in the following section of this work is

$$y = 0.1437x + 0.0006, \quad (7)$$

where y is the output thrust value in N, and x is the input displacement measurement.

Thrust measurement testing

The experimental scope reported on includes thrust measurement values obtained through the firing of the TTB system with cold gas CO_2 at standard atmospheric pressure. The displacement values produced by the thruster with chamber pressures of 50 psi, 75 psi, 100 psi, and 125 psi were measured using the laser positioning sensor and were recorded in the LabView application. The mass flow rate of the CO_2 was measured with the Omega 0–30 L/min mass flow meter interfaced with the computer using the National Instruments DAQ system. The procedural sequence involved first opening all valves within the feed line system to adjust the gas cylinder pressure regulator to the required pressure. The line pressure was verified with a more sensitive gauge further down the line, and the regulator was adjusted to match the required pressure on the more accurate gauge. The control valve in between the two pressure gauges was then closed, and the loading beam/moment arm was manually steadied before commencing the test firing. The DAQ system and LabView recording mechanisms were started, and then the flow valve was manually opened. After it was observed that the displacement, and therefore the thrust, reached a steady state for a measureable amount of time, the flow valve was closed and the feed line to the thruster was allowed to depressurize. The moment arm was then returned to an unexcited state before the subsequent test was performed. This process was repeated for each of the aforementioned chamber pressures in the said sequence. The testing sequence was repeated three times with the obtained measurements being recorded and then imported into Microsoft Excel for further analysis.

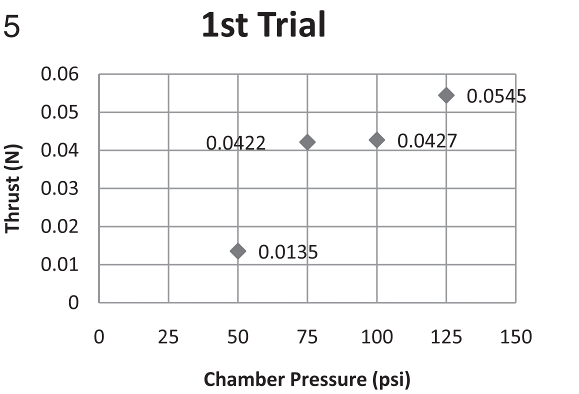
Preliminary thrust measurement results

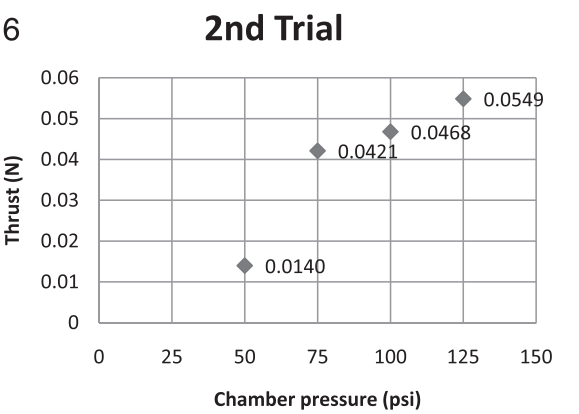
The postprocessing of the data involved a series of steps to factor out values obtained through the natural vibrational frequency of the system, thrust overshoot produced at the commencement of each test, and the thrust produced during the depressurization of the cold gas supply lines. By making a series of graphs, the steady state thrust region was identified by locating the section where the median values of the oscillations formed a fairly linear, horizontal line. A starting point and an ending point for the region was determined by locating them on the graph and were then located within the data set, and the median value of the thrust was calculated in Microsoft Excel. From this process *Figures 5–7* were produced, detailing the individual results for each trial.

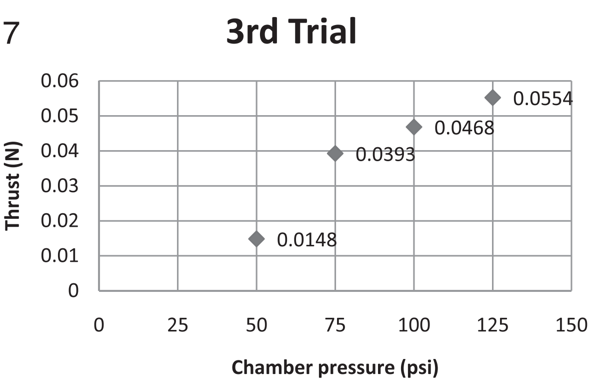
Figure 8 shows the graph of the average values obtained from the thrust measurements. As can be seen, the graphs are nearly linear, with the exception of the thrust measurement for the 75-psi chamber pressure in all three of the trials. Since the other pressures seem to have produced data with a more linear relationship to the other measurements, it is presumed that an error occurred while taking these measurements. As all of the factors within our test procedure remained constant for the duration of the experimentation process, it is believed that inconsistencies were caused by the human error factor introduced while manually opening the cold gas flow valve at the commencement of this series of trials. Other factors such as the flow of the ventilation system of the laboratory could have also changed during this test period.

The depiction of average thrust measurements in *Figure 8* shows the range over which the values varied. The two pressure ranges that exhibited the greatest differences were the 75-psi and 100-psi trials. At this time it is unclear as to why these variations are present, but one of the possible explanations for this behavior is that there is error in the pressure measurement. The pressure was measured upstream of the chamber, not at the chamber, so there might have been variations in the actual chamber pressure. A pressure transducer may need to be installed to measure the chamber pressure.

Figure 8 details the averaged thrust results. It is notable that the confidence in these measurements is significantly lower than that of the calibration values. This could be due to many factors including the presence of an exhaust fan to evacuate the test area of the released CO_2, the speed at which the gas valves were manually opened, and feed-line pressure variations due to losses at bends in the tubing. However, it is also important to mention that, as stated before, these error-inducing factors were proven by the calibration to be factors other than the structural characteristics of the thrust stand itself. Therefore, such influences can be minimized through modifications to the testing setup. Consideration for the next round of trials could include

5


6


7


Figures 5–7. Thrust versus chamber pressure.

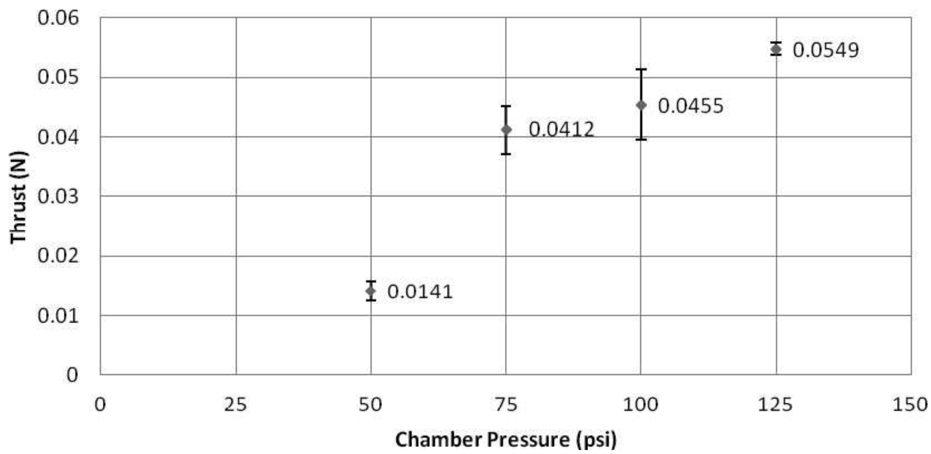


Figure 8. Average thrust measurements.

adding an automated valve system. Preparations for inclusion of a one-sided damper are currently being made.

Conclusions

With some preliminary results from the design prototype, most of the requirements were met. Under the static thrust profile of a cold gas thruster, the thrust stand was able to evaluate thrust levels with satisfactory repeatability. However, the natural frequency measured was in the range of 6–6.5 Hz and therefore does not meet the pulsed-frequency requirements. Several currently unexplained irregularities in postprocessed thrust measurement data reveals that there are still tare sources to be mitigated. However through the high accuracy of the calibration measurements, it can be determined that the sources of error are most likely related to either human error or the conditions within the environment surrounding the test setup. None of the data collected indicates that the problems lie within the structural design of the TTB system itself. This implies that the design is suitable and a full-scale stand can be built, but further examination of the setup as a whole is necessary for optimization. A computer simulation model is useful to optimize and refine the critical design parameters (e.g., torsional spring constant, damping constant, and moment arm length) of the thrust stand without the financial and time burdens of only experimentation. The result is a more accurate thrust measurement, fine tuned for a variety of thrust profiles.

Future work

To increase the accuracy of the measurements taken using the TTB system, further investigation will be made into the possible sources of tare/irregularities within the experimental setup. A one-sided linear damping system will be utilized to stabilize the natural vibrational frequency of the system as well as reduce the vibrational intensity of the overshoot phenomena. Further computer simulation modeling will be conducted to serve as a comparison to the observed dynamic responses/characteristics of the TTB system. With the collection of more data, it will be possible to conduct more extensive error analysis operations as well. Upon further optimization of the system, more studies will be conducted utilizing other common propellants. The system will also be modified for use in vacuum to simulate working altitude conditions. ❑

Marjorie A. Ingle is an undergraduate research assistant at the University of Texas at El Paso's Center for Space Exploration Technology Research (cSETR). Marjorie is an undergraduate in the Mechanical Engineering Department at the University of Texas at El Paso and is currently researching low thrust measurement techniques for 5N class propulsion systems. Marjorie has also participated in the University of Texas (UT) Alliance Louis Stokes Alliance for Minority Participation Program (LS-AMP) from 2007 through 2009, working on projects involving additive manufacturing technology, failure analysis of biomedical technology, satellite signal transmission, and miniature space propulsion systems. In 2008 and 2009, Marjorie was also a contributor to the Wohlers Report: State of the Industry. She has interned with the Community Scholars Program in El Paso, Texas where she co-authored and published two papers about public policy in Texas. Upon graduating with her bachelor of science degree in mechanical engineering, Marjorie plans to pursue a masters degree in mechanical engineering, with aspirations to earn a doctorate.

References

Rao, S. S. 2004. *Mechanical vibrations*. Upper Saddle, New Jersey: Pearson Prentice Hall.

Runyan, R. B., J. P. Rynd, and J. F. Seely. 1992. *Thrust stand design principles.* AIAA 17th Aerospace Ground Testing Conference. Nashville, TN, AIAA-92-3976: American Institute of Aeronautics and Astronautics.

Stark, K. W. 1969. *Patent No. 3572104*. Hyattsville, MD.

Acknowledgments

The material is based upon work supported by NASA under award No NNX09AV09A.

ITEA Journal 2011; 32: 167–179

Sustainability via Energy Harvesting and Scavenging

Jose Barajas, Anand Sagar Bokka, Isaac Fernandez, Juan Gonzalez, Ahmad Hammad, Javier Hernandez, Sughosh Kuber, Daniel Marquis, Ravi Chandra Mikkilineni, Chaitanya Motupalli, Abdi Musse, Vineet Nair, Gur Notani, Javitt Higmar Padilla-Franco, Balwinder Singh, Prasant Sinha, Murali Sithuraj, Paari Sivaprakasam, Vanishree Narayan Venkatesan, and Nadipuram (Ram) R. Prasad, Ph.D.

Rio Grande Institute for Soft Computing (*RioSoft*) & *Rio*RoboLab
New Mexico State University, Las Cruces, New Mexico

This article presents three novel concepts for energy harvesting and scavenging, demonstrating the outcome effectiveness of a program for energy harvesting at New Mexico State University. The design concepts cover a wide range of harvested power outputs from microwatts to kilowatts and illustrate a sampling of the breadth of potential energy harvesting and scavenging capabilities needed to create sustainable environments. The concepts have wide-ranging applications and demonstrate how off-the-shelf components can be integrated into reliable, cost-effective, and environmentally benign implementations for meeting essential energy needs of human society. The outcome clearly demonstrates the value of an experiment-based approach to train graduate and undergraduate students in conceiving, designing, building, testing, validating, and evaluating environmentally benign systems for harvesting and scavenging energy from ambient energy sources that can lead to sustainability.

Key words: Build; design; sufficiency; systems engineering; test; thermal energy scavenging system.

We begin with a clear understanding that harvesting energy is not necessarily limited to the transformation of available energy resources into electricity. This transformation can occur in any useful form required to serve the needs of humans and other biological systems and act as a means to providing a level of sufficiency. Achieving sufficiency could be interpreted to mean the satisfaction of all fundamental societal needs by energy-bearing resources that are natural and freely available and that provide sustenance in all basic requirements by humans, machines, and biological systems. This brings about broader awareness of all forms of energy that prevail in the environment and can be used effectively. In this context, sustainability is the most frequently used word to describe the benefits derived by adopting a strategy towards achieving sufficiency.

Best Graduate Student Paper from the ITEA Live Virtual Constructive Conference, January 2011, El Paso, Texas.

A general strategy towards building energy sufficiency is to offset conventional nonrenewable resources by the energy from renewable resources. These renewable or regenerative sources have the ability to serve fundamental energy needs for indefinite amounts of time, including needs that are periodic and have long time-constants. The needs are indeed multitude in nature, ranging from charging a cell phone battery to providing power to life-support systems, or to serve the needs of any other application using the energy harvested from ambient sound, light, heat, water, and wind energy. Meeting the imminent energy needs with technologies can sustain growth and is therefore the means to sustainability. Such a strategy holds the potential to transform nonrenewable sources into backup sources and make renewable ambient sources the primary mechanism towards achieving sustainability. Under this strategy for example, a car battery and gasoline would become supplemental (backup) energy sources, with the energy harvested from the heat generated by the engine and exhaust system, the motion of the wheels, the pressure inside the tires, the shock-absorber actions, and the air-flow around the vehicle providing the power needs of critical operational functions. It is well known that the use of an air-conditioning system in an automobile significantly reduces gas mileage, and a means for using harvested electrical power would offset the power produced by gasoline consumption.

It is significant in the design and development of any new technologies to keep pace with ongoing world events so that the design includes the fundamental mechanisms of energy transformation (as observed from these events) to meet the energy needs. The events serve as worst-case scenarios of both natural and spontaneous events that place humans in harm's way. This information is required to understand the possible range of scalable technologies, and for developing approaches to harvest the maximum amount of energy possible. To date, very little research had been performed in the U.S. and worldwide into harvesting energy from extreme weather. This includes hurricanes, typhoons, earthquakes, and other natural disasters that have the potential to provide much needed electric power for rapid recovery following catastrophic events. Recent natural disasters in Haiti, Chile, Indonesia, and the Philippines stand out as examples where the lack of electric power following earthquakes and typhoons made recovery considerably longer, causing much pain and suffering to humans in these affected regions of the world. The April 2010 BP oil spill in the Gulf of Mexico is yet another example for the need to have technologies in place to harvest energy and provide resources on demand for emergency cleanup operations. Such resources minimize the overall cost of recovery while simultaneously providing the means for mitigating environmental damage.

Perhaps the most striking example of where technologies for natural energy harvesting could have been most useful is the November 11, 2010, incident in which a cruise ship that lost all power due to a fire in the engine room left 4,500 passengers with no air conditioning, no hot food, and no hot water. It was also reported that the swimming pool was off limits because there was no way to pump chlorine (Drachen Foundation 2004). With the abundance of solar, wind, and buoyant energy available at sea, it is easily perceived that such a scenario could have been averted.

A program for energy harvesting was initiated in fall 2010 in the Klipsch School of Electrical and Computer Engineering with a course offering, to stimulate research and development in this very important area. The course provides graduate students and seniors opportunities to conceive, design, develop, test, and validate revolutionary ideas for harvesting ambient energy through experimentation and researching new approaches. A poster entitled "Natural Energy Harvesting – A physics-based art form" presented at the University Research Council Poster Fair in October 2010 won the first-place award (KiWiGen Project 2003).

Energy from wind, water, light, heat, and sound, just to name a few of the most obvious source types, is abundant as are the natural energy sources in the environment. In southwestern New Mexico, the wind during March and April each year is a significant periodic event that has enormous potential for novel approaches to wind energy harvesting (Drachen Foundation 2004; KiWiGen Project 2003). *Figure 1* illustrates possible sources that can contribute to grid-connectible and distributed harvested energy as a means to achieve energy sustainability.

It is a natural human instinct to harvest the energy available from all sources to satisfy a majority of the energy needs using technologies that are environmentally benign. This attribute has immense value not only to national security and defense, but also to the entire society. The range of application for such an energy-harvesting strategy spans from the smallest to the largest amount of energy that is freely available that meets the application requirements. There are no limits, therefore, to how much energy one can harvest from both natural renewable sources and human-generated sources. The fact remains that conservation of energy principles apply.

Overview of course and design projects

Energy Harvesting I is a three-credit-hour course allowing students to apply their knowledge in *fluid dynamics*, *power electronics*, *machine design*, *control systems*, *structural design*, *computer control*, *embedded systems*, *system dynamics*, and many other areas and to combine this knowledge with strong *systems engineering* practices to design and develop revolutionary energy harvesting and scavenging systems. Experiments provide hands-on experience for developing innovative methods for energy-harvesting applications. Operating principles of several harvesting techniques such as solar, tidal, thermal, vibration, linear motion, and passive and active human power generation methods are discussed in class along with experiments that help confirm these concepts as viable means of energy harvesting. Student groups engage in experiment-based projects for energy harvesting, conduct research on harvesting technologies, and select off-the-shelf technologies to integrate into novel, simple, and cost-effective solutions aimed at building energy and environmental sustainability. Students perform simulation studies through modeling, submit project reports, make oral presentations, and demonstrate

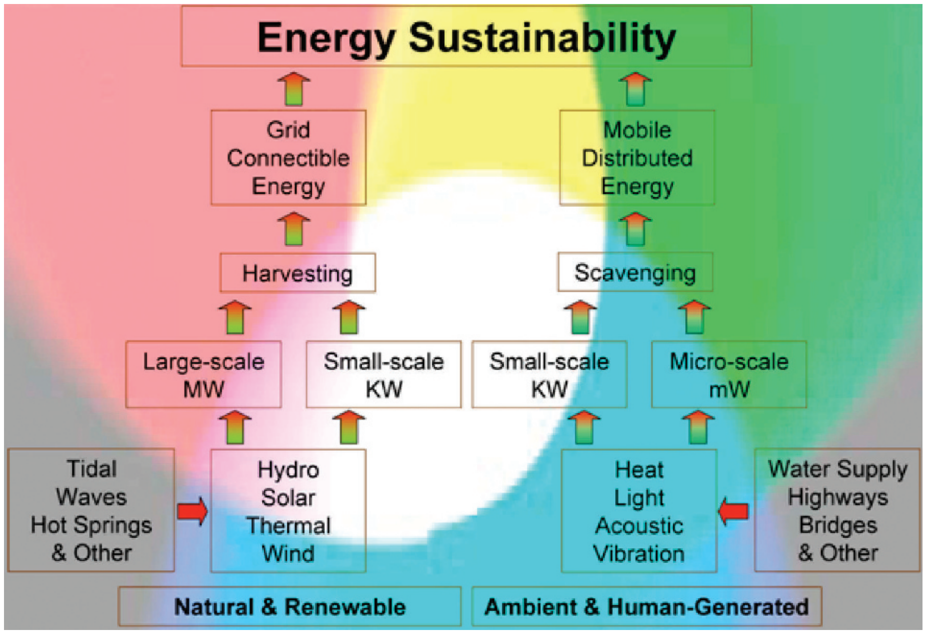


Figure 1. Constituent sources for energy sustainability.

their ideas to the extent possible to show project feasibility. Three projects are described in this paper.

Thermal Energy Scavenging System (TESS)

Thermocouples are widely used in all applications that require thermal measurements for purposes of control. Applications range from home heating and cooling to space applications. TESS is an energy scavenging system that extracts heat energy from any radiating heat source, using thermocouples. The type K thermocouple, which is the most widely recommended, is used to demonstrate a novel heat energy scavenging system. Type K thermocouple is composed of Chromel (+), an alloy that is 90 percent nickel and 10 percent chromium, and Alumel (−) consisting of 95 percent nickel, 2 percent manganese, 2 percent aluminum, and 1 percent silicon. The resulting p-n junction formed makes it the most common general purpose thermocouple, with a sensitivity of approximately 41 μV/°C. It is inexpensive, and a wide variety of probes are available in its −200°C to +1350°C (−328°F to +2462°F) range. One of the constituent metals, nickel, is magnetic; a characteristic of thermocouples made with magnetic material is that they undergo a step change in output when the magnetic material reaches its Curie point, around 354°C for type K. Rooftops and walls of large buildings, copper-coated glass panes covering multistoried buildings, metal structures, and virtually anything that radiates or reflects heat are viable applications for thermal energy harvesting and scavenging. An embedded controller monitors the hot and cold junctions and maintains an optimum thermal gradient to provide the best electrical output from the thermal environment. An architecture for optimal heat farming applicable to distributed power generation is discussed below.

Heat farming – a concept for distributed energy harvesting. Technologies for thermal harvesting require a temperature differential in order to generate voltage. This is described by the well-known Seebeck effect (Halliday, Resnick, and Walker 1997). The Seebeck effect describes a phenomenon wherein a voltage is produced by a temperature gradient across a junction of two dissimilar conducting elements. *Figure 2* illustrates this phenomenon very clearly.

TESS exploits this type of sensing to harvest useable energy. The amount of heat produced as a result of normal and abnormal operation of systems can be harvested in the form of charge.

Noting that charge Q is proportional to voltage V, the constant of proportionality C represents the

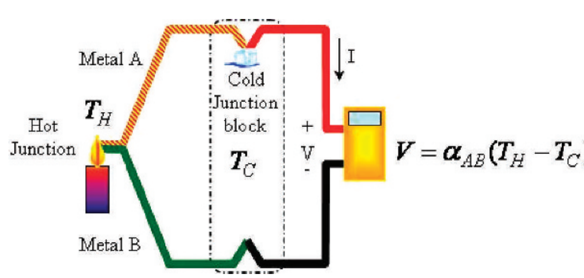


Figure 2. Voltage generated by thermal gradient.

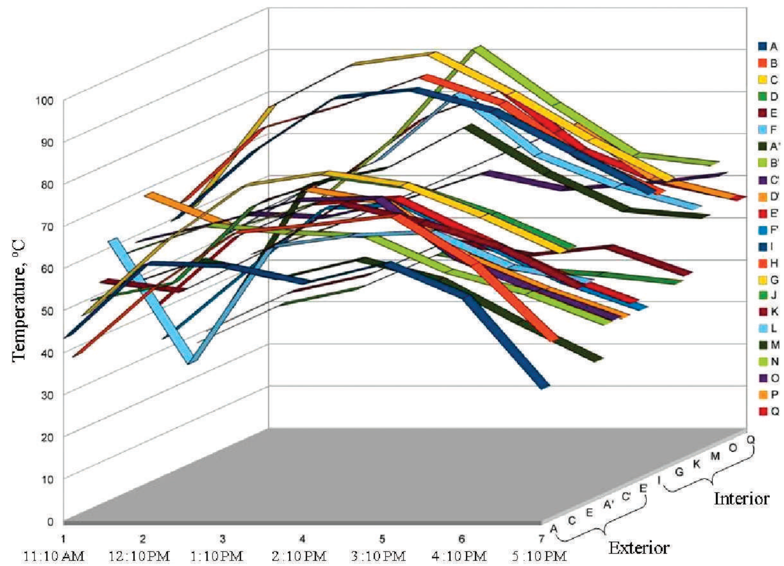


Figure 3. Temperature measurements outside and inside a vehicle over 7-hour duration.

number of coulombs per volt. Therefore, to produce 1 volt, 1 coulomb is exactly equal to 6.241 509 629 152 65 $\times$ 10^{18} positive elementary charges. According to Coulomb's Law, two point charges of +1 C, 1 meter apart, would experience a repulsive force of 9 $\times$ 10^9 N, a force roughly equal to the weight of 900,000 metric tons of mass. Given this fundamental knowledge, it is easy to visualize how much power can be harvested from thermoelectric effect and the numerous applications it can serve.

Figure 3 illustrates a set of temperature measurements made inside and on the outside of an automobile, using a Radio Shack infrared thermometer to record the temperature over a 7-hour period, on September 11, 2010. It is clear that the interior temperature is higher than the exterior by nearly 35°C. This thermal gradient is sufficient to produce power from a thermopile, a collection of thermocouples connected in series and in parallel configurations to act as voltage and current generating sources.

The important concept here is that a large temperature differential between the hot and cold junctions will produce a large voltage across the terminals. The voltage differential is time-of-day dependent, thus this generated voltage varies significantly. A means to maximize the thermal gradient is therefore needed. The embedded control system will choose among a set of references to dynamically adjust the thermocouple connection in order to maximize the amount of generated voltage.

Figure 4 illustrates a design concept that maximizes the thermal energy harvesting from automobiles. The scalability of this approach extends the possibilities for large-scale thermal energy harvesting. This may include harvesting energy from furnaces, jet engines, and other heat generating sources. Utilizing the fundamental behavior of a thermocouple is obviously the most efficient means to recover electrical energy from radiated heat. It provides a technology that is environmentally benign. The applications clearly demonstrate the possibilities for developing heat farms in regions of the U.S. where heat is a natural resource and to contribute towards achieving sustainability.

Design architecture. The principal design objective is to maximize the thermal gradient for maximum energy-harvesting efficiency. This is achieved by proper selection of a "cold" reference in real time. Solar irradiation on surfaces varies over the day as the sun rises and sets. As such, when one side of the automobile is hot, the opposite side is relatively cooler. To accommodate this change, a set of metallic plates placed in the cooler regions of the object act as a reference to provide the largest thermal gradient. A low power Microcontroller Unit (MCU) will switch the references according to the gradient changes. The MCU is programmed to execute a "Greedy Algorithm" based on the concept that if a local maximum is chosen, then the global maximum will be achieved. A low power MCU was used to test this concept. A schematic is illustrated in *Figure 5.*

In the operating mode, the microcontroller initially chooses the sensing mode through the digital output D_0. Once in the sensing mode, the microcontroller will cyclically switch among all the references in order to find the maximum voltage. Following this switching, when a maximum gradient is found, the microcontroller will change to generating mode using the digital

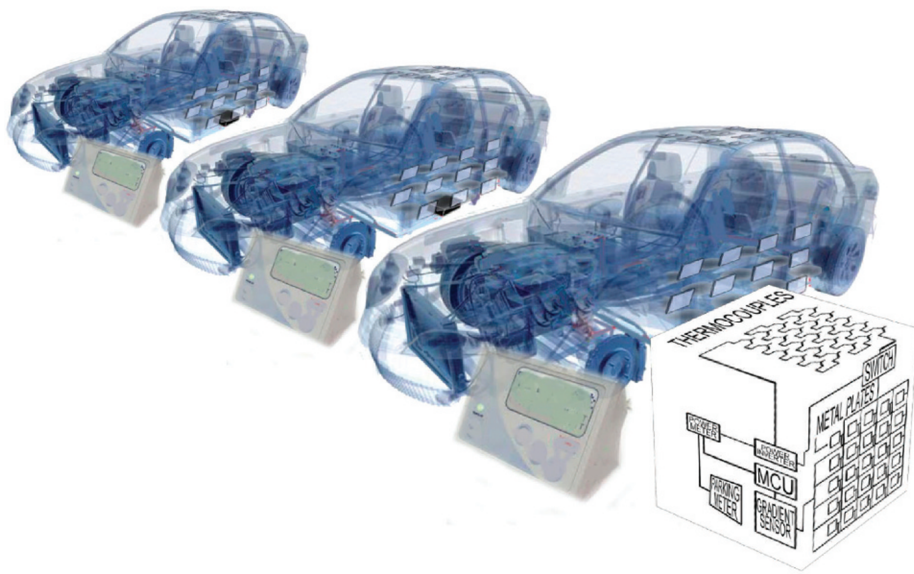


Figure 4. Parking lot thermal energy farm.

output D_0. This process of locating a set of references that yield the largest gradient will be executed throughout the day to achieve maximum output.

Energy harvesting Survival raft **(E-Surv)**

Survival at sea is epitomized by the abundance of freely available natural energy sources that provide life-sustaining capabilities under extremely adverse conditions. The objective of *E-Surv* is to provide humans lost at sea a means for survival and the resources needed to increase survivability through harvesting both solar energy (thermal and light) and the energy from buoyant wave pressure, and through an innovative idea for producing drinkable water.

E-Surv is equipped with power harvesters that meet the needs for several critical life-support functions. These functions include power for transmitting and receiving signals via satellite beacon, power to charge batteries, power for a miniaturized distillation system to provide drinking water, power for thermal blankets, and power for Light Emitting Diodes (LEDs) to provide search and rescue teams greater visibility of the raft in pitch darkness. Periods of bright moonlight require special consideration in controlling the wavelengths of light emission from LEDs, its amplitude and frequency. *Figure 6* illustrates an artist's rendition of a single-person *E-Surv* life raft (left) and an array configuration (right) showing the possibility for harvesting greater amounts of solar and wave energy that contribute to survivability.

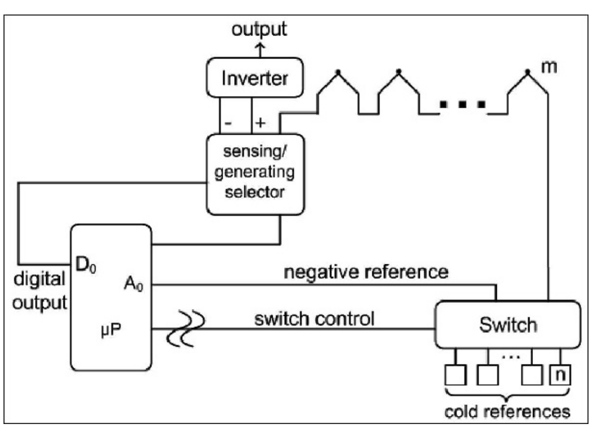


Figure 5. Microcontroller unit application to maximize temperature gradient across thermocouples.

E-Surv is a foldable inflatable raft, which opens into a floatation device large enough to fit a human. In the event of multiple life rafts being deployed, it is envisioned that each raft can be electrically connected and tethered. This configuration emulates an energy farm in the high seas by providing a larger surface area for solar energy harvesting; it improves raft stability and provides greater reliability to harvest sufficient power needed for surviving extreme weather conditions for prolonged periods of time. *Figure 7* illustrates a cross-section of the life raft showing details of a critical life-support function, namely, the distillation of sea water, along with locations where solar, thermoelectric, and piezoelectric power harvesters are most effective.

Referring to *Figure 7,* the raft consists of two chambers: one that is inflatable, and the other, which carries sea water pumped in either manually or by a miniature pump. This water will be allowed to evaporate by means of solar irradiation and then condensed to provide clean drinking water. The ribs,

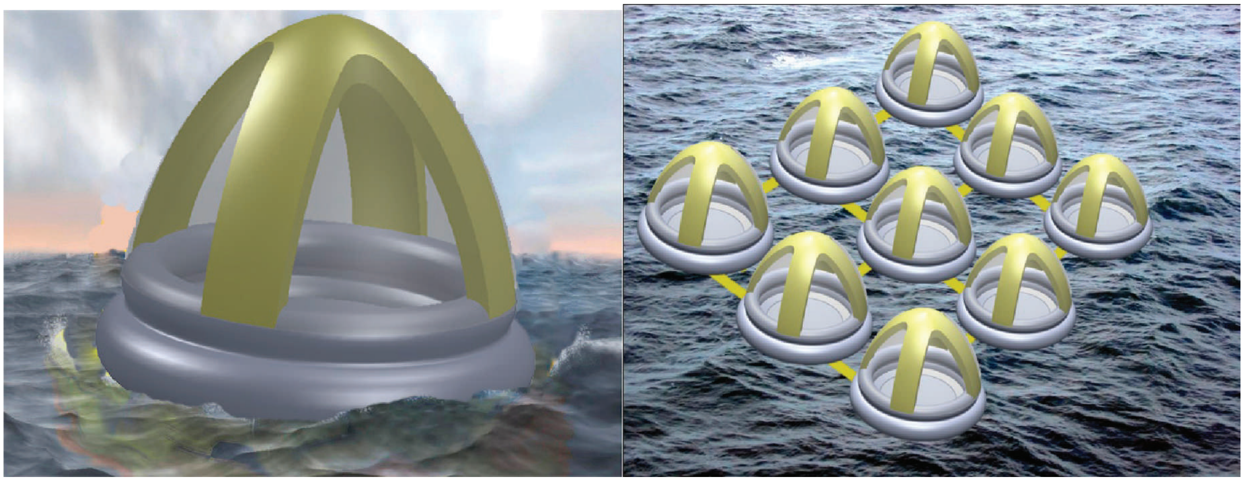


Figure 6. Energy harvesting Survival raft (E-Surv) *conceptual illustration – Single raft (left) and array (right).*

which are made out of durable and flexible plastic, provide a column for evaporated water to condense and be collected at the bottom of the rib structure through a reverse flow path.

The chamber holding salt water is coated with graphene, a highly processed form of graphite and among the best thermal conducting materials known. This allows salt water to evaporate naturally by solar heat and minimizes the use of harvested energy. In the absence of sufficient solar heat for evaporation, a temperature sensor provides feedback to the micro-controller to enable energizing a heating element with power harvested from the piezoelectric array and thermopiles.

The rib structure additionally provides support to a plastic canopy, laced with ultra-bright LED lights, which can be zipped closed. This serves to protect survivors from adverse weather conditions. Solar power allows the transmitting and receiving of Global Positioning System (GPS) and communication signals through a distress beacon to enable search and rescue operations (NOAH 2003). Piezoelectric arrays provide

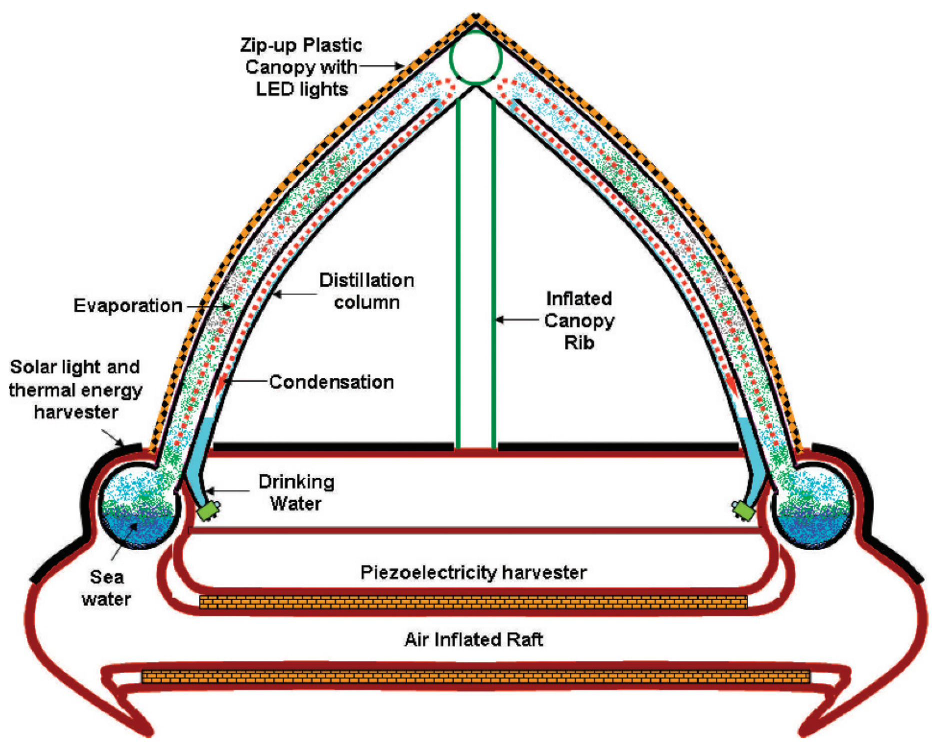


Figure 7. Cross-section of Energy harvesting Survival raft (E-Surv).

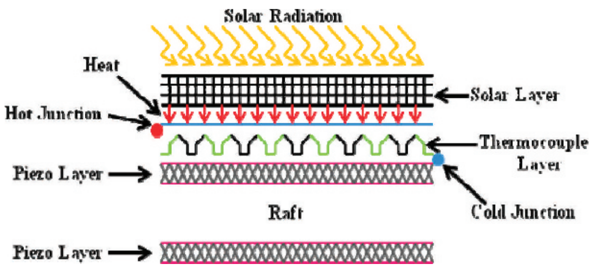


Figure 8. Raft energy-harvesting systems.

power for maintaining the operation of low-power electronic devices (Sunghwan 2002). Lightweight, highly portable, and with no moving parts, *E-Surv* is a viable technology for providing life support.

Figure 8 shows a novel configuration of solar panels, thermopiles, and piezoelectric strips to provide the maximum continuous power-harvesting capability for a selected size of life raft.

Design concept. The power required to operate *E-Surv*'s life-support systems is harvested by utilizing photovoltaic, thermoelectric, and piezoelectric effect. Flexible solar arrays fixed along the periphery of the raft harvest solar power during daytime. Heat generated by the solar array and the incident solar radiation can be harvested using thermopiles. Temperature gradient is created by using sea water as the cold reference. Buoyant pressure energy can be harvested by using piezoelectric strip arrays placed both on the floor and underneath the raft, as illustrated in *Figure 7*.

A low-power MCU is used in *E-Surv* for efficient usage of the limited power available. *Figure 9* shows the schematic for combining power harvested from various sources. Harvested energy is regulated and stored in a battery. An ultra-low-power microcontroller is used to manage power utilized by the GPS transmitter and receiver, distress beacon, LED lights, and to supply the heating-element power required by the distillation process and thermal blankets. The LED lights should be turned on intermittently only during low light conditions. Based on the battery power level and pre-assigned priorities, the microcontroller prioritizes the operation of the devices.

A voltage regulator is used to maintain a fixed input voltage to a rechargeable battery. The energy harvested from thermopiles and piezoelectric arrays is stored in capacitors. Alternating current (AC)–direct current (DC) converters are employed to get the required input to the LED lights, distillation process, and distress beacon. The light-dependent resistor is used to determine the ambient brightness. This information is used by the microcontroller to switch on LEDs during periods of darkness.

A 406-MHz distress signal beacon with an integrated GPS system for tracking is provided on the raft. It accesses the Cospas-Sarsat satellite for the transmission of the distress signal and for GPS tracking.

***Venturi-flow submarine-type Turbine Generator for low-head hydro* (VT-Gen)**

Conventional hydroelectric power generation requires dams and reservoirs to create the required potential head and is, therefore, considered environmentally unfriendly. Power harvested from free-flowing water in rivers, canals, and streams and human-generated energy supply systems including urban water supply systems and waste-water flow systems constitute a large portion of the hydropower that is otherwise unused. *VT-Gen* is a revolutionary axial-flow turbine-generator technology that paves the way to harvesting sufficient power to bring sustainability to agricultural and irrigation communities in rural and remote mountainous areas where there is an abundance of water resources but where conventional practices for generation, transmission, and distribution are impractical. The usefulness and practicality of *VT-Gen* lies in its ability to provide power-harvesting capability in high mountainous regions where there is an abundance of water resources. Areas such as in the mountains of the Hindukush and Kashmir, power

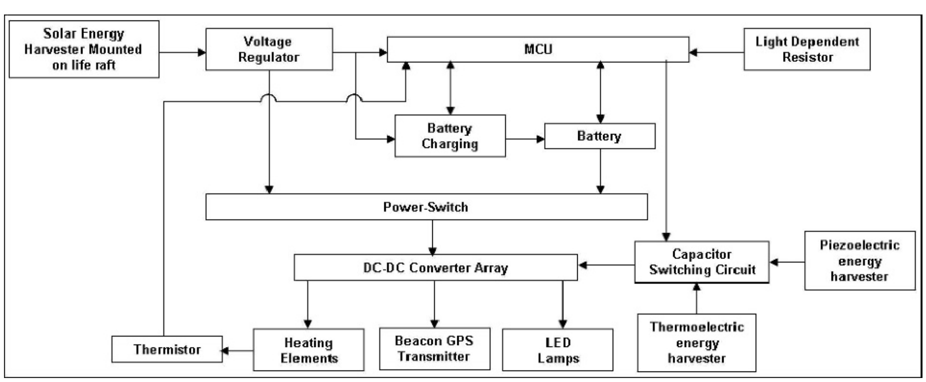


Figure 9. Schematic of power harvesting on raft.

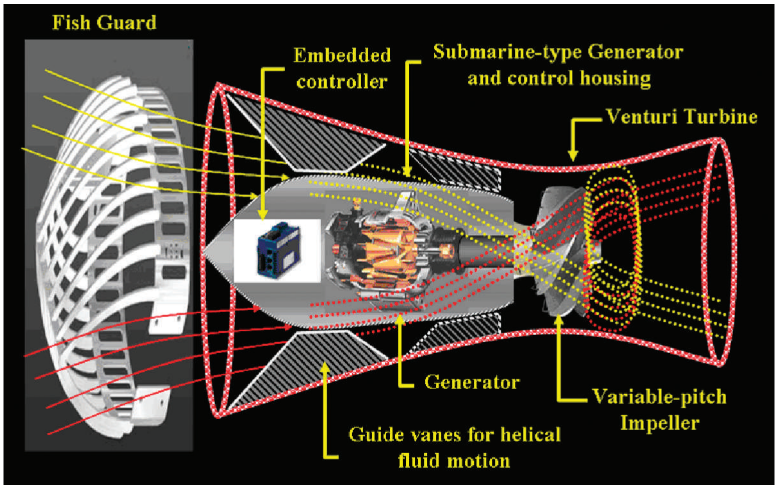


Figure 10. Venturi-flow submarine-type Turbine–Generator (VT-Gen).

generation using portable hydroelectric generators appears to be a feasible alternative to fossil fuels. In order to provide energy resources to these remote areas, diesel and gasoline fuels are generally transported by convoys of tanker trucks. In Afghanistan, for example, North Atlantic Treaty Organization (NATO) trucks carrying fuel have been routinely attacked by insurgents and set ablaze. From a strategic defense perspective, therefore, there is an imminent need to displace conventional fuels (diesel and other carbon-based fuels) with technologies that can harvest as much power needed to meet at least the minimum needs of the Armed Forces. There is no doubt that the introduction of new and novel technologies leads towards achieving energy sufficiency. *VT-Gen* is a novel concept for power generation that has the potential for application in any fluid-flow system. It is a scalable technology that can be useful over a wide range of flow-rate conditions.

Operating principle of VT-Gen. *Figure 10* illustrates the components of *VT-Gen* suitable for harvesting hydro energy. *VT-Gen* comprises a venturi-shaped shell acting as a conduit for fluid flow. A generator inside a submarine-shaped casing is mounted axially inside the venturi shell. The generator shaft is connected to an impeller, which acts as the prime mover. As water enters the venturi tube, the water accelerates, resulting in an increased velocity at the throat of the venturi. Guide vanes in the inlet of the venturi induce helical (swirling) fluid motion. At the throat of the venturi, water impinges upon the turbine blades and transfers a portion of its energy to the turbine. The water is then extracted through the remainder of the venturi structure due to the expanding cross-sectional area and the accelerating pull of water passing by the venturi exit. A fish-guard acts to prevent both fish and debris from entering the venturi shell.

Water flow rates in rivers, canals, and streams vary based upon seasonal conditions with high flow rates during monsoons and rainy periods, and low flow rates during off season. Irrigation canals, for example, generally have a constant flow rate in order to provide agricultural communities with a constant supply of water. For such applications, it is sufficient to have a fixed-blade impeller. However, in applications with highly variable water flow rates, a microcontroller is required to vary the pitch angle of the blades to ensure constant generator speed. The microcontroller makes pitch-angle adjustments based on the turbine speed.

Performance equations. The venturi flow describes the relation between volume flow rate, pressure difference, and the ratio of throat to inlet cross-sectional area. In Equation 1, Q is the volumetric flow rate, C is the discharge coefficient (usually between 0.9 and 0.98 for a venturi shape), A_1 is the cross-sectional area at the inlet, A_2 is the cross-sectional area at the throat, ρ is the fluid density, $(P_1 - P_2)$ is the pressure differential.

$$Q = C \cdot A_2 \cdot \sqrt{\frac{2\cdot(P_1 - P_2)}{\rho \cdot \left[1 - \left(\frac{A_2}{A_1}\right)^2\right]}} = C \cdot A_1 \cdot \sqrt{\frac{2\cdot(P_1 - P_2)}{\rho \cdot \left[\left(\frac{A_1}{A_2}\right)^2 - 1\right]}} \tag{1}$$

Kinetic power (kW) in the fluid passing through a fixed cross-sectional area can be calculated according to Equation 2. "A" is the cross-sectional area through which the fluid is flowing, ρ is the fluid density, and v is the flow velocity. Notice that power is proportional to the cube of the velocity. This is a well known fact in wind power that applies to hydro power also.

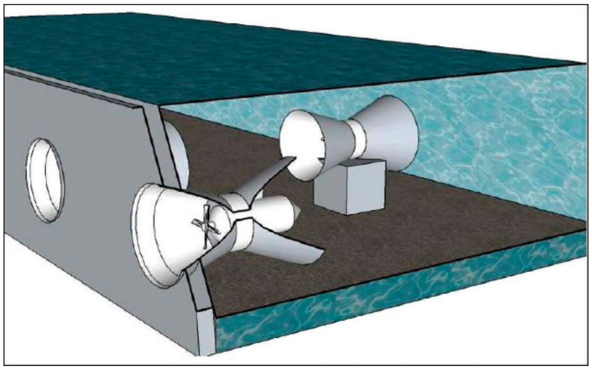


Figure 11. Run-of-river and weir wall application of Venturi-flow submarine-type Turbine–Generator (VT-Gen).

$$\text{Kinetic\_Based\_Power} = \frac{1}{2} \cdot (\rho \cdot A \cdot v) \cdot (v)^2 \qquad (2)$$

Potential power (kW) in water created by a dam, or weir, or a reservoir can be calculated according to Equation 3. “Anet” is the net cross-sectional area through the turbine at velocity, v; ρ is the fluid density; h is the effective net head of the fluid; and g is acceleration due to gravity.

$$\text{Potential\_Based\_Power} = \rho \cdot \text{Anet} \cdot v \cdot g \cdot h \qquad (3)$$

Power harvested from water by the turbine can be calculated using torque, τ, and angular velocity, ω, as shown in Equation 4.

$$\text{Mechanical\_Power} = \tau \cdot \omega \qquad (4)$$

Equations 1–4 establish the basis for the performance of *VT-Gen*. While these equations provide a basis for computing the ideal power-harvesting potential, the actual power that can be harvested depends on flow rates, impeller design, venturi shape, and of course generator design.

Some unique applications. Two unique possibilities for harvesting energy from hydro-potential sources are illustrated in *Figures 11* and *12*. *Figure 11* shows a concept for a run-of-river power generation in which *VT-Gen* is fixed onto a sturdy structure and placed along the axial flow of the river. In a weir application, *VT-Gen* is placed (embedded) in the weir wall. *Figure 12* illustrates a concept where the use of *VT-Gen* is ideally suited to harvest the unused hydro potential developed by urban water tanks. It is quite clear from these example applications that *VT-Gen* is widely applicable to both open-flow as well as closed-flow systems. As a scalable technology, it provides virtually unlimited possibilities to harvest energy from residential water supply, urban water supply, and waste-water systems with power output ranging from milliwatts to kilowatts. In any application, however,

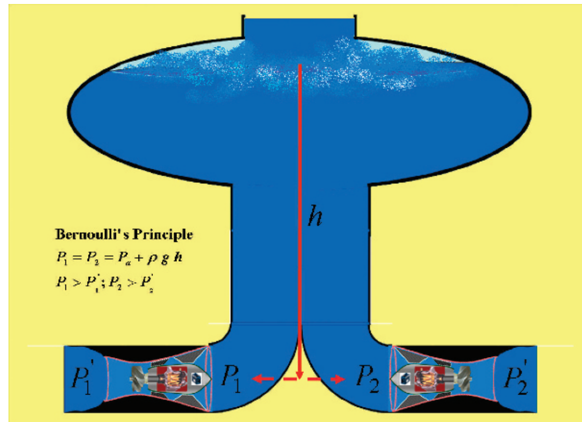


Figure 12. Urban water supply application of Venturi-flow submarine-type Turbine–Generator (VT-Gen).

accessibility is a principal design requirement to ensure serviceability towards reliable operation.

Design issues. There are three principal components in *VT-Gen* that are critical to the power-harvesting capability. First and foremost is the venturi shell, second is the impeller, and third is the generator. In the following sections, a few specifics regarding the design of the turbine generator are discussed.

Venturi shell. A preliminary computational fluid dynamics analysis of the turbine design has shown that an efficiency of approximately 90 percent or more is possible. This is in sharp contrast to conventional hydropower generation, where the turbine efficiencies range between 60 and 70 percent. The ratio of throat and inlet diameter and outlet and throat diameter are the main design specifications, and their effect is determined by the previously mentioned venturi flow (Equation 1).

Generator. The generator could be a DC or an AC unit. The unit will be encased inside a submarine shaped shell, which also houses the speed control mechanism. Since permanent magnet based generators have no rotor winding electrical losses, they are more efficient and will be used.

Proposed run-of-river prototype. The run-of-river application shown in *Figure 11* is currently being prototyped. The prototype is being designed with an outer diameter of 2 feet and a tube length of between 4 and 6 feet. These dimensions were chosen to generate at least 100 watts, to stay within the project's financial budget, and to allow use of off-the-shelf components. The energy production in a river with a flow rate of 1.90 miles/s (4 mph) is estimated to be around 150 watts. This approximation has been made based on general principals regarding energy, momentum, and fluid-flow continuity.

Table 1. Maximum, minimum, and most probable power levels for an example run-of-river Venturi-Turbine Generator (VT-Gen) Design (estimates only)

	Power harvested (watts)	Net effective area assumed (m^2)	Overall efficiency assumed (%)
Maximum (absolute) power limit	850	0.3	100
Most probable power estimate	170	0.15	40
Minimum power estimate	80	0.1	30

In designing the prototype, three power levels (energy per unit time) were calculated to help size the equipment: an upper bound, a lower bound, and an expected value. For example, the power producible by an x-foot diameter *VT-Gen* is upper bounded by the total power available in the water that would normally flow through an empty (and turbine-less) x-foot diameter tube. On the other hand, the power in said *VT-Gen* is perhaps lower bounded by the total power available in the small amount of water that would normally flow through just the net area of the venturi throat, minus 70 percent for inefficiencies. The true amount of power that will be produced lies somewhere between these two estimated bounds. In the above estimation process, the equation for kinetic power (Equation 2), an equation for net area (Equation 5) to exclude the generator shell, and an equation for expected power (Equation 6) are required.

$$\mathrm{NetArea} = \mathrm{Anet} = \pi\left(\mathrm{OuterRadius}^2 - \mathrm{InnerRadius}^2\right) \quad (5)$$

$$\mathrm{ExpectedPower} = \mathrm{PowerAvailable} \cdot \mathrm{OverallEfficiencyOfConversion} \quad (6)$$

Having chosen a venturi tube entrance diameter of 2 feet, a venturi throat diameter of 1.5 feet, a generator shell diameter of 1 foot, and a brisk river velocity of 1.79 miles/s, a maximum theoretical available power of almost 900 watts is obtained. Also, a lower bound estimate of actual power output is obtained at around 80 watts. Finally, an estimate of the mostly likely (expected) power output is perhaps closer to 150 watts. These estimates are summarized in *Table 1.* Note that (similar to a wind turbine) the power available varies by the cube of the river speed and by the square of the venturi radius. Moderate changes in venturi size and small changes in river speed have a very large impact on estimated power output.

Figure 13 illustrates a micro-grid concept, where power generated by water flowing through irrigation canals can be effectively integrated into the distribution system to serve residential and agricultural loads.

Embedded control

Intelligent embedded control brings a high level of autonomy to a system and also helps improve the efficiency of a system. This is useful in both the

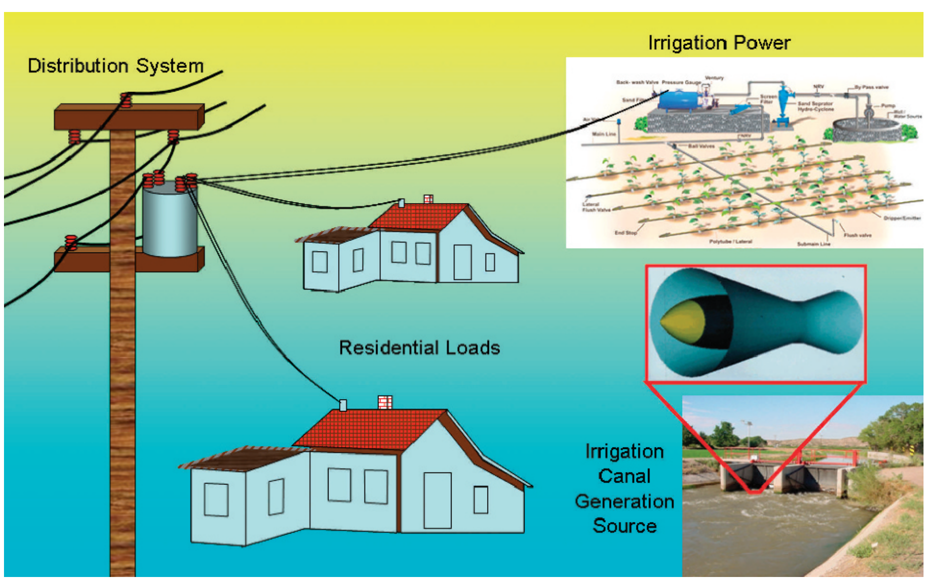


Figure 13. Micro-grid and energy farm.

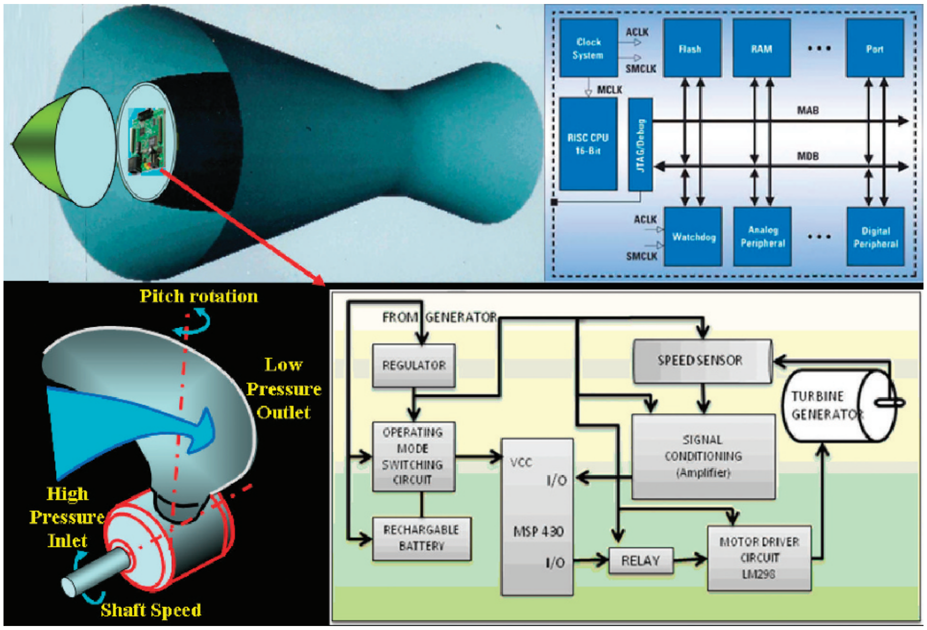


Figure 14. Embedded control block diagram for turbine blade pitch angle.

turbine-generator and *E-Surv* life-raft energy-harvesting applications.

There are a number of embedded processor choices in the range of 8-bit to 32-bit microcontrollers or even high-end processors, which provide key features that benefit the energy-harvesting applications discussed in this paper. They include the following:

- low cost,
- low supply voltages,
- single or dual core,
- specialized peripherals,
- power aware tools,
- on-chip MAC, FPU, math co-processors, etc., if required,
- wide temperature range, and
- ready-to-use software libraries.

The above-mentioned characteristics make embedded processors a viable choice to control the turbine-blade pitch angle of the *VT-Gen*. *Figure 14* illustrates a block diagram of the essential components required for pitch-angle control.

A signal conditioning circuit to amplify the signal from the analog speed sensor is required, as shown in *Figure 14*. Relays provide isolation between the microcontroller and the high voltage control circuitry for pitch control of the turbine blades. The motor driver circuit comprises the LM298 IC and is controlled by the microcontroller. Additional relays are used to switch the power source between generator power and the battery.

Under normal conditions, the MCU is powered by the output of a voltage regulator connected to the generator and is able to control the pitch angle of the turbine according to the speed of rotation. If the output of the voltage regulator falls below an operating threshold of the MCU, due to slowing down of the turbine speed, the microcontroller's power supply is switched to a rechargeable battery source to enable continuous operation. This switchover mode has the potential for voltage recovery caused by sudden changes in water flow dynamics. Immediately following the recovery, switchover occurs again to restore power supply to the MCU through the voltage regulator.

Figure 15 illustrates a conceptual closed-loop block diagram. The MCU acquires input from the analog speed sensor and computes the speed deviation from nominal. To implement a pitch-angle controller, a fuzzy logic–based approach could be adopted. *Figure 16* illustrates a control surface defined by nine fuzzy IF–THEN rules relating speed deviation and change in acceleration (deceleration) to the blade pitch angle of the turbine impellor.

Conclusions

Specific design details have been intentionally omitted to protect the intellectual value and property rights of all the contributors. However, the functionalities of the technologies have been clearly discussed.

The most significant aspect of design is to give form through a rigorous systems engineering process.

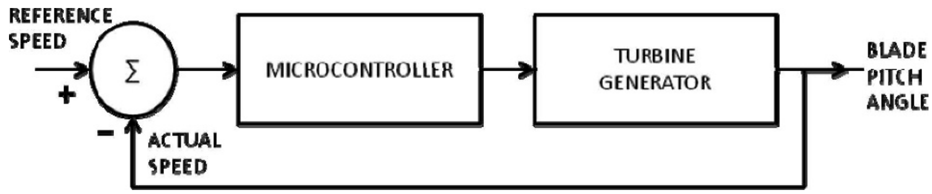


Figure 15. Closed-loop feedback for blade pitch-angle control.

During the 16-week fall semester, students engaged in concept design, prepared functionality specifications, and conducted experiments to confirm the energy-harvesting possibilities from ambient energy sources. As an outcome of this course, the paper presents three technologies that hold promise towards building cost-effective and highly reliable approaches for both autonomous and semi-autonomous energy harvesting and scavenging systems.

The approach has focused on innovative methods as a means for developing environmentally benign energy harvesting and scavenging systems. It is clear that such systems will play a significant role in advancing the state-of-the-art in energy-harvesting technologies so that our dependence on conventional power generation can be minimized. There is no doubt that novel energy-harvesting technologies are needed for deep-space travel and planetary exploration.

Embedded control is undoubtedly a major design component in energy harvesting and scavenging. Students will explore the performance of several low-power microcontrollers and choose those which are best suited both in cost and performance.

It is extremely beneficial for students to understand the meaning of sustainability and how sustainable energy-efficient environments can be developed. Experiment-based courses naturally stimulate curiosity and inquisitiveness towards exploration. The physics-based nature of experiments allows a better understanding of the laws governing physical behavior, thereby, sharply improving learning. Students engaged in projects demonstrate their innovative thinking and creativity through a deep sense for innovation. The motivation to build systems that are bio-inspired helps develop a clear understanding of the principles of systems science that allow research and development of robust, autonomous and semi-autonomous, and reliable energy harvesting and scavenging systems.

Student participation in local engineering chapters that sponsor conferences and poster presentations brings greater awareness among students to the dissemination of technology and the merits of preserving intellectual property rights.

Working with industry partners gives greater credibility towards developing technologies that may impact consumers. Students gain hands-on experience by working in teams, which takes dedicated and coordinated effort, to design usable energy harvesting and scavenging systems.

The fall course has helped to develop a research agenda aimed at integrating mostly off-the-shelf components, including reusable material from junkyards, to conceive, design, build, test, and validate novel environmentally benign technologies for energy

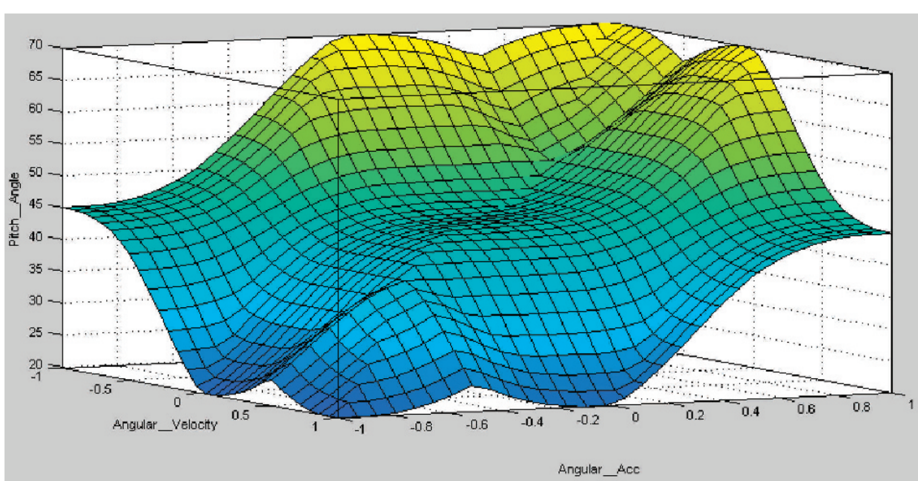


Figure 16. Fuzzy control surface for pitch-angle correction.

harvesting and scavenging systems. In a follow-up course during spring 2011 students will build prototypes of the systems described in this paper. ❑

STUDENTS: JOSE BARAJAS is a doctor of philosophy candidate with interests in control systems. ANAND SAGAR BOKKA is a master's degree candidate with interests in power systems, control systems, and energy harvesting. ISAAC FERNANDEZ is an undergraduate in electrical engineering. JUAN GONZALEZ is a master's degree candidate with interests in control systems and energy harvesting. AHMAD HAMMAD is an undergraduate student in electrical engineering. JAVIER HERNANDEZ is a doctor of philosophy candidate with interests in power systems and VLSI. SUGHOSH KUBER is a master's degree candidate with interests in power systems. DANIEL MARQUIS is a doctor of philosophy candidate with interests in power systems. RAVI CHANDRA MIKKILINENI is a master's degree candidate. CHAITANYA MOTUPALLI is a master's degree candidate with interests in control systems and energy harvesting. ABDI MUSSE is an undergraduate student in electrical engineering. VINEET NAIR is a master's degree candidate with interests in embedded control systems. GUR NOTANI is a doctor of philosophy candidate with interests in photovoltaic systems. JAVITT HIGMAR PADILLA-FRANCO is a doctor of philosophy candidate with interests in control systems, soft computing, and software engineering. BALWINDER SINGH is a doctor of philosophy candidate with interests in power systems and energy harvesting. PRASANT SINHA is a master's degree candidate with interests in computer engineering. MURALI SITHURAJ is a master's degree candidate with interests in power systems. PAARI SIVAPRAKASAM is a master's degree candidate with interests in power systems. VANISHREE NARAYAN VENKATESAN is a master's degree candidate with interests in power systems.

*FACULTY ADVISOR: DR. NADIPURAM (RAM) R. PRASAD is an associate professor in the Klipsch School of Electrical and Computer Engineering Department at New Mexico State University. He is also the Director of the Rio Grande Institute for Soft Computing (*Rio*Soft) and* Rio*RoboLab, a National Aeronautics and Space Administration (NASA) Ames–funded advanced robotics laboratory. Dr. Prasad's teaching and research interests are in soft computing, energy harvesting, and systems science. Dr. Prasad has authored and co-authored over 150 publications in journals and conference proceedings and is the co-author of two books on fuzzy and neural control. E-mail: naprasad@nmsu.edu*

References

Drachen Foundation. 2004. Using kites to generate electricity: Plodding, low tech approach wins. *Drachen Foundation Journal.* Fall: 14–16. http://www.drachen.org/journals/a16/Using-kites-to-generate-electricity.pdf (accessed March 11, 2011).

Halliday, D., R. Resnick, and J. Walker. 1997. *Fundamentals of Physics*. New York: John Wiley & Sons.

Kim, Sunghwan. 2002. Low power energy harvesting with piezoelectric generators. Doctoral dissertation, University of Pittsburg. http://etd.library.pitt.edu/ETD/available/etd-12092002-153537/unrestricted/SHKIMA.pdf (accessed March 11, 2011).

KiWiGen Project. 2003. The KiWiGen project kite wind generator: A new concept in wind power generation. http://ec.europa.eu/research/energy/pdf/other_res02_ippolito.pdf (accessed March 11, 2011).

NOAH. 2003. Emergency Position Indicating Radio Beacon (EPIRB) personal locator system. Noah Satellite and Information Service. http://www.sarsat.noaa.gov/emerbcns.html (accessed March 11, 2011).

ITEA Journal 2011; 32: 181–190

Background Discrimination of Bioaerosols

James J. Valdes, Ph.D.
CTNSP/National Defense University, Washington, D.C.
Raymond Mackay, Ph.D.
Independent Contractor, Bel Air, Maryland
Joshua L. Santarpia, Ph.D.
Johns Hopkins University Applied Physics Laboratory,
National Security Technology Department, Laurel, Maryland
Stephen Adams, Ph.D.
Department of Pharmacology, 310 George Palades Laboratories, La Jolla, California
Ewelina Tunia
National Defense University, Fort McNair, Washington, D.C.
Andrew Mara, Ph.D.
Center for Technology and National Security Policy/National Defense University, Washington, D.C.

The accurate testing and evaluation (T&E) of bioaerosol detectors requires a detailed knowledge of both the disseminated bioaerosol of interest and of any background particulates present during the testing window. Along these lines, there is a growing consensus among the bioaerosol T&E community that, as detector technologies continue to improve, knowledge of the "ground truth" state during detector testing is becoming increasingly important. In order to more accurately assess biodetector technologies, we identified currently or nearly available technologies and methodologies that could serve as referee systems during chamber, breeze-tunnel, and field testing. At the minimum, these technologies should allow testers to determine the particle size distribution of distributed bioaerosols, a key component of the T&E of point and standoff detection systems.

Key words: Background particulates; bioaerosol detection systems; biological viability; culturability; fluorescent tagging; Rapid Agent Aerosol Detector (RAAD); referee systems; Single Particle Aerosol Mass Spectrometry (SPAMS); Ultraviolet Aerodynamic Particle Sizer (UV-APS).

A number of variables are important when characterizing a disseminated bioaerosol. Most important, a referee detector needs to determine the particle size distribution (PSD) of the bioaerosol. Current backscatter technology used in most standoff detection systems is very sensitive to the number and size of particles. In addition, knowledge of the PSD, when combined with information on the composition of the aerosol, can provide an estimate of the total agent per liter of air (TALA) or the number of agent containing particles per liter of air (ACPLA) (Valdes et al. 2010). TALA contains information on the total amount of agent presented to a detector, the key variable in point detector testing. Thus, accurate measurement of the PSD of the distributed aerosol is likely the most important characteristic of a referee detector.

Additional details about the PSD are also desirable, though not necessarily required. For example, tests with some currently available technologies (i.e., Ultraviolet Aerodynamic Particle Sizer [UV-APS]) indicate that, while a distributed aerosol's PSD may extend to the submicron range, the majority of particles that actually contain agent are more likely to fall in the 2–3 μm+ range. This is because particles generated in the field, rather than in the laboratory, are normally greater than 1 μm regardless of the size of the agent in the solution or slurry. Data on which particles actually contain agent and therefore pose a larger threat would be very useful. Current standoff detector technologies do not test for the presence of agent, so in theory a disseminated aerosol containing only media and salt crystals could return a signal that is very similar to an aerosol that actually contains agent. Moreover, when translating testing and evaluation (T&E) results for operational decision making, the PSD of *agent containing particles* will affect rates of infectivity, hence operational considerations. In the case of many aerosols, media may dominate the observed fluorescence. For example, the fluorescence of an MS-2 aerosol at 355 nm excitation must be due solely to the media since tryptophan, its only likely biological fluorophore, is unlikely to contribute to fluorescence at this wavelength. The UV-APS may therefore not confidently provide separation between agent/simulant fluorescence at 355 nm and that of other contributors such as media.

The vast majority of particles present during a chamber test will be from the disseminated aerosol, with few background particles, as air that enters the chamber passes through multiple high efficiency particulate air (HEPA) filters. However, during breeze-tunnel or field testing, background particulates can be as high as 1,000 particles per liter (ppl), and possibly higher. Moreover, background levels are quite dynamic and can vary greatly over short periods of time. In this case, the presence or absence of background particulates can have a large effect on detector performance. In order to assess detectors under different ambient background conditions, testers need the ability to not only measure the disseminated aerosol, but also the background particulates present at each time point. This ability becomes even more important when testing detectors in the presence of deliberately disseminated interferents such as road dust, smoke, diesel exhaust, and so forth, which could be encountered in urban or battlefield environments. During these tests, the level of background particulates will be substantial and could greatly impact detector performance. A new referee system should be able to accurately measure both disseminated bioaerosols and ambient/intentionally disseminated background aerosols.

Ideally, a referee sensor would also determine the viability of organisms or lethality of toxin. As discussed in an earlier publication, while viability information would certainly be useful from an operational perspective, it is not necessary for the proper T&E of current detector technologies, which cannot determine viability (Valdes et al. 2010). There currently exists no technique that can accurately determine the viability of an organism after it has been disseminated. In fact, this is further complicated by the fact that viability is not synonymous with culturability. Disseminated organisms can become quiescent (i.e., nonculturable) after aerosolization and then subsequently revive; depending on conditions under which they were grown, and ambient conditions. This calls into question the traditional definition of viability as being based on culturability. Thus, while viability measurements are a characteristic of an ideal referee sensor, they will almost certainly be absent from a referee system employing light scattering, fluorescence, or other identification sensors such as laser-induced breakdown spectroscopy (LIBS), polymerase chain reaction (PCR), and immunoassay (Miziolek and Palleschi 2006; Hindson et al. 2005). An *estimate* of viability may be obtained simultaneously using existing techniques (e.g., slit to agar samplers).

Particle sizes of interest for a bioaerosol referee sensor

Ideally, a referee sensor would record information on all particles of all sizes during a bioaerosol challenge. However, practical and technical limitations make it exceedingly difficult to collect and process data for all particles. Generally speaking, respirable particles lie within the 1–10 μm range, making these particles of particular interest to the T&E community. Below 1 μm, most particles will not contain a significant amount of agent; hence they will not have an effect on point detector testing. However, these particles can still make a significant contribution to the detection signatures for standoff detectors, particularly technologies using shorter wavelengths to interrogate aerosol clouds. Likewise, particles above 10 μm tend to settle out fairly quickly and therefore make up a minority of particles within a distributed bioaerosol. Despite their low numbers, large particles produce signatures for standoff detectors and can make a disproportionately large contribution to detector sensitivity tests.

From a traditional threat viewpoint, most particles that contain biological material (and agent) are at least 2–3 μm in size, except for single-spore dispersions, which can be 1 μm in size. Since agent containing particles are of the most interest to point detection systems (and to estimations of operational implications), a strong argument can be made that particles between 2 and 10 μm are of particular interest to the T&E community. However, there are two caveats: (a) future threats may be engineered in smaller sizes, and (b) particles in the 10–15 μm range, while fewer in number, likely settle out more rapidly, and have difficulty penetrating the lungs as deeply as 1-μm particles.

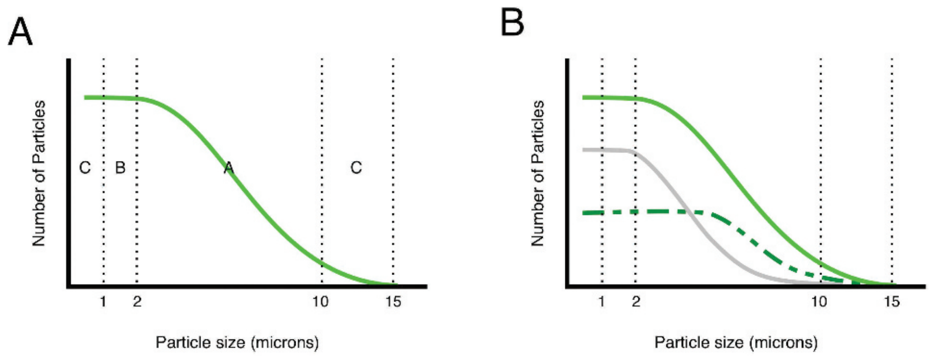


Figure 1. Desired capabilities of a referee system.

However, due to the large number of the former, and the ability of the latter to carry a large amount of agent, they could have a non-negligible effect since the number of organisms that an aerosol particle could contain increases as the cube of the particle's radius.

An ideal referee sensor would accurately assess particles from 0.5 µm to 15 µm in size. However, not all sizes in this range are of equal importance. The highest priority size range is likely 2–10 µm, followed by 1–2 µm. The third tier of size ranges includes the submicron and 10 µm+ particles, which are unlikely to make significant contributions to bioaerosol threat but may significantly affect standoff detector performance *(Figure 1)*.

Current particle sizing option – APS

PSDs are currently measured using a standard APS. While these systems manage to count particle sizes in the ranges required for T&E, they have no ability to discriminate between disseminated aerosols and background particulates. Thus, the current workhorse of the T&E community does not collect sufficient data to accurately test detectors when background particulates are present, as they are in breeze tunnels or field tests *(Figure 2)*. For these reasons, we conclude that the standard APS is not sufficient for T&E of biodetectors. Four potential technologies that could partially fulfill the disseminated and background aerosol measurement requirements are described below.

UV-APS

Standard APSs are currently routinely employed by the T&E community to gather PSD information during detector testing. While these APSs do an admirable job of counting particles within the size ranges of interest, they offer no discrimination between background particles and particles that are part of a disseminated bioaerosol. The UV-APS (TSI, Inc.) improves upon the standard APS by measuring both the size and the autofluorescence of particles (Agranovski and Ristovski 2005). Each particle is excited by UV light (355 nm) and particles containing biological materials then fluoresce. Disseminated particles that contain biological material can thus be discriminated from nonbiological (e.g., dust and sand) background particles.

In breeze-tunnel tests at Dugway Proving Grounds (DPG), the UV-APS appears to be able to discriminate disseminated aerosol from ambient background. Before a bioaerosol release, the UV-APS detects negligible numbers of biological background particles, effectively subtracting the ambient background from measurements of an intentional release. Furthermore, the UV-APS can measure particles from 0.5 to 15 µm, effectively covering all of the particles of interest (though counting efficiency drops significantly for particles under 1 µm).

The UV-APS has several drawbacks that limit its ability to be the sole referee instrument used to determine PSDs of background and disseminated aerosols. First, the UV-APS has difficulty discriminating between agent- or simulant-containing aerosol, and aerosol composed of biological media and some interferents including burning hydrocarbons and re-aerosolized dust that often contain fluorescent material from previous tests. Many interferents contain compounds that will fluoresce when excited with UV light. The UV-APS records this fluorescence and counts those particles as biological particles. In the presence of specific interferents, the UV-APS alone is unable to discern a disseminated aerosol from the interferent and cannot determine an accurate PSD for the disseminated agent or simulant aerosol. In addition, the UV-APS picks up a substantial number of biological particles when dust from a testing facility is re-aerosolized. Data from UV-APS (and other technology) testing at DPG suggest that re-aerosolization is not a large concern during standard breeze-tunnel tests; however, depending on the testing scenario, the UV-APS may not be able to discriminate background biological particles from disseminated biological parti-

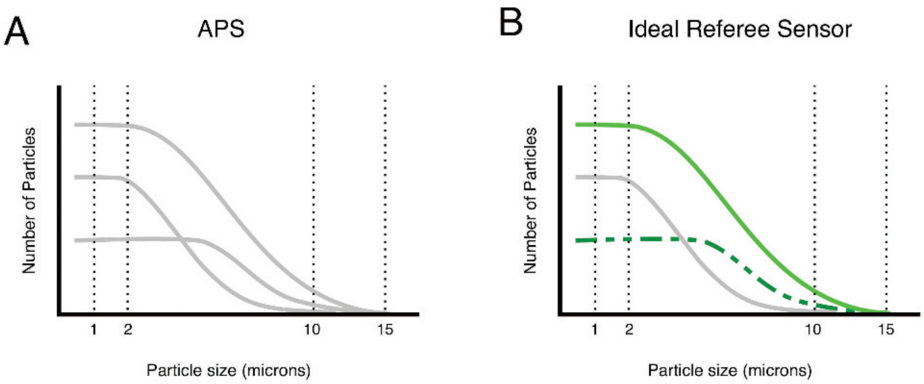


Figure 2. Standard aerodynamic particle sizer has no background discrimination ability.

cles. Furthermore, the UV-APS cannot discriminate disseminated nonbiological particles (e.g., disseminated particles that do not contain agent) from nonbiological background. While this measurement is not required for the T&E of point detectors, standoff detectors will include nonbiological disseminated particles when interrogating an aerosol cloud *(Figure 3)*.

Despite these limitations, the UV-APS certainly increases the ability of the T&E community to accurately measure a disseminated bioaerosol. Perhaps its strongest asset is the fact that it is relatively affordable (~\$120,000), commercially available, and already in limited use. For these reasons, use of the UV-APS as a referee system in breeze-tunnel and field tests could be implemented immediately.

Rapid Agent Aerosol Detector (RAAD)

The RAAD is a new aerosol detector that has been developed by MIT Lincoln Laboratories to serve as the next generation biological aerosol warning system for use in point detectors. The RAAD makes a number of measurements on single particles (Gonser, Christensen, and Guicheteau 2004). First, the RAAD uses a patterned laser beam to determine the size of incoming particles, as well as their relative position and velocity within the RAAD sample path. This patterned laser beam should actually offer higher fidelity measurements of particle size than a standard APS, as knowledge of the position of the particles can be used to correct location-specific limitations of the device. More specifically, particle sizing sensitivity decreases at the edges of the sample path, but this problem can be mathematically corrected if the particle location is known.

The second step measures autofluorescence in a manner similar to a UV-APS. However, the RAAD excites each particle with UV light of two different wavelengths (266 nm and 355 nm), as opposed to one in a UV-APS, and collects four emission wavelengths (the UV-APS collects one); using these data to determine which of the particles contain biological material. Thus, the first two stages of the RAAD could be described as more advanced versions of the UV-APS.

The unique feature of the RAAD is the third step, Spark-Induced Breakdown Spectroscopy (SIBS) (Cremers and Radziemski 2006). Plasma created by high-voltage spark breaks down particles that the RAAD has determined are biological in nature. The broken

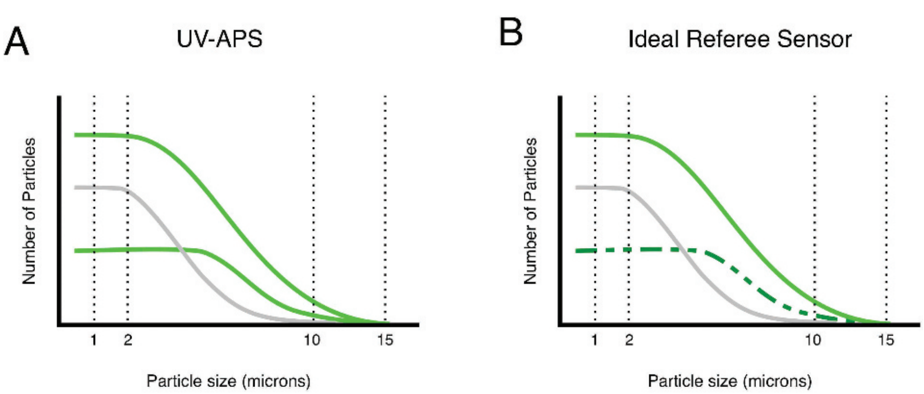


Figure 3. Ultraviolet aerodynamic particle sizer can discriminate between biological and nonbiological particles.

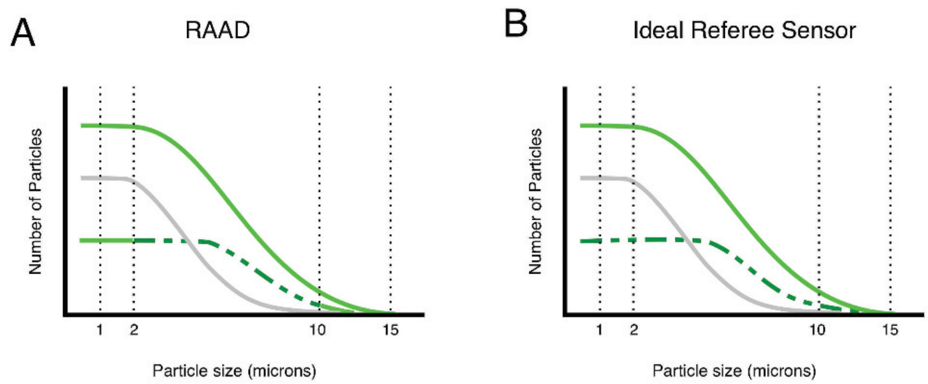


Figure 4. The Rapid Agent Aerosol Detector provides ideal discrimination in the 2–10 μm range.

down particle is then subjected to an elemental analysis in which four elements (i.e., potassium, magnesium, phosphorus, and sodium) are detected. The elemental analysis of the SIBS, when combined with the fluorescent information that was collected earlier, allows for robust determination of the content of each particle *(Figure 4)*. The RAAD not only determines if a particle is of biological origin but also identifies what materials (or media, salts, etc.) are present in each particle. The RAAD is able to accomplish this feat because the results of the fluorescent and elemental analysis create two highly orthogonal data sets. While many types of particles may have similar fluorescent spectra, very few types of particles have similar *fluorescent and elemental* spectra.

Thus, the RAAD is well positioned to provide much of the information needed to be a valuable referee sensor. In particular, the RAAD has the capability to differentiate background aerosols from disseminated aerosols and can further differentiate agent containing aerosol particles from non-agent containing particles. Moreover, the RAAD has undergone extensive field testing and is currently slated for commercial production in 2011. Production model RAADs, or variants modified for T&E purposes, should be available in the short term.

The RAAD has limitations. It is currently configured to measure particle sizes between 2 and 10 μm. While this population will likely contain the vast majority of agent containing particles, other disseminated and background particles are certain to exist outside this range. As mentioned above, these particles can significantly affect standoff detector testing. Second, the rate-limiting step is SIBS. Not all particles that enter the RAAD can undergo SIBS (SIBS has a maximum rate of 20 particles/s), thus only a subset of particles can truly be identified. However, the RAAD does collect data on the proportion of particles it subjects to SIBS, and reasonably accurate estimates of the identity of all particles are possible in most conceivable scenarios. Indeed, it is relatively straightforward to expand the range of particle sizes examined by the RAAD; however, the majority of those particles would not undergo SIBS analysis.

The RAAD should be capable of discriminating disseminated aerosols from most interferents, as the interferents should not have fluorescent and elemental spectra that are similar to biological agents. However, the RAAD will likely have difficulty distinguishing re-aerosolized particles from previous tests. That caveat aside, testing of the RAAD and the UV-APS have not found re-aerosolization to be a problem under standard test conditions. One additional consideration is that the biological fluorophores that we rely upon for detection may be labile under solar radiation and other environmental stress, but data on these phenomena are preliminary.

The RAAD is not currently in mass production, but prototype units could be available for limited testing. The cost of this prototype is likely to decrease substantially following commercial production. Current estimates place the cost of a production unit in a range from $80,000 to $90,000. This price is for the military hardened version of a RAAD, which may not be the most suitable version for T&E purposes. While the T&E version will need less physical protection, the lower volume of units that would be purchased by the T&E community would likely increase the cost. Thus, it is reasonable to assume that a T&E version of the RAAD will cost approximately the same as the military production version, making it cost-effective to use the standard military version. At this price range, the RAAD compares favorably with the UV-APS.

Fluorescent tagging of distributed aerosols

One way to discriminate a disseminated aerosol from background particles is to "tag" the aerosol with a

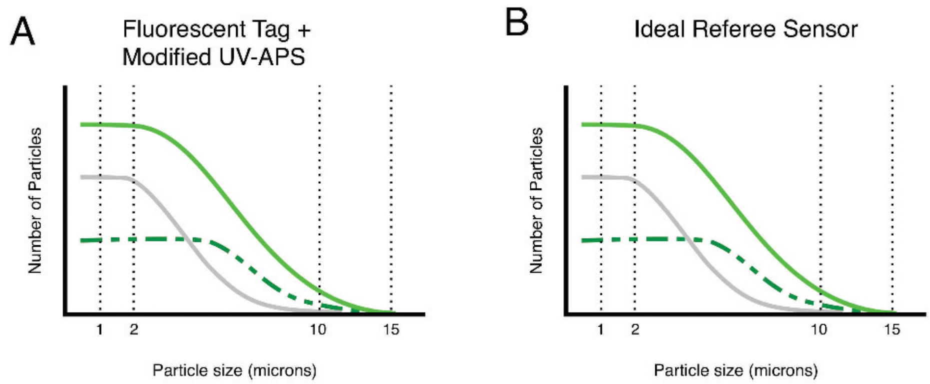


Figure 5. Fluorescent tagging can provide ideal aerosol discrimination.

fluorescent marker that makes it inherently different. A referee detector can detect the presence of this tag to determine which particles to count as disseminated aerosol and which to count as background. Both chemical and protein fluorescent tags are commonly used in biological research.

One immediate concern is that the fluorescence of the tag may affect the performance of detector technologies. For example, if a standoff detector excites the fluorescent tag in a given particle, that particle may give an enhanced signature to the detector and skew test results. Fortunately, fluorescent tags only absorb light at certain wavelengths and a large number of tags exist that do not efficiently absorb light at any of the wavelengths currently used in standoff detectors. Yellow fluorescent protein (YFP), for example, absorbs light very poorly in the infrared range *(Figure 5)* and will likely have little impact on detector performance (Nagai et al. 2002).

This flexibility in tag choice is also extremely beneficial in discriminating a tagged aerosol from background biological aerosols. The UV-APS and RAAD use innate fluorescence to determine that a particle is biological in nature. Since all biological particles (disseminated and background) have this same fluorescence, innate fluorescence cannot be used to discriminate among them. Using a tag that responds to wavelengths that are different than innate particles allows the tester to excite only particles that contain the fluorescent tag. YFP, for instance, is maximally excited at ~470 nm, well outside of the innate fluorescence of biological particles, which generally absorb light between 250 and 350 nm. Furthermore, the light emitted from a tag such as YFP has a much longer wavelength (~520 nm) than light emitted by innate fluorescence, but not necessarily longer than that captured by the detector. While there is some innate fluorescence at longer wavelengths (i.e., flavin compounds), this fluorescence is likely to be 100- to 1000-fold lower than that of a fluorescent tag, and a fluorescently enabled particle sizer (i.e., a slightly modified UV-APS or the first two stages of the RAAD) should have little difficulty distinguishing a tagged aerosol particle from biological background *(Figure 5)*.

There are two options available for fluorescently tagging a test aerosol: protein tags and chemical tags. Protein tags are widely used in biological research and fluorescently tagged versions of many simulant agents likely already exist. For nontagged agents, tagged strains could be created quickly and cheaply with standard molecular biological techniques. The largest benefit of a protein tag is that, since the tag is created by the agent itself, once a fluorescent simulant strain is established, producing tagged simulants is no more difficult than growing the cultures of untagged simulants currently used. The limitation of tagging the agent itself is that test particles that do not contain agent will not be tagged and hence will be indistinguishable from background. However, agent containing particles will be clearly marked and counted. In addition, there may be environmental concerns in disseminating large quantities of tagged agents in breeze tunnels or field tests. While these concerns could be mitigated by, for example, severely limiting the ability of the tagged simulants to reproduce in the wild, it is likely that regulatory approval would be necessary before use of protein-tagged simulants could commence. If the fluorescently tagged agent persists in the environment, there are additional concerns that re-aerosolized simulants could affect subsequent tests. These concerns may not be a problem in practice due to natural photo-bleaching of these proteins. Again, there are technical mitigations for this problem (inducible fluorescent tags, for example), but these problems will need to be addressed in order to use fluorescent protein tags during outdoor T&E trials.

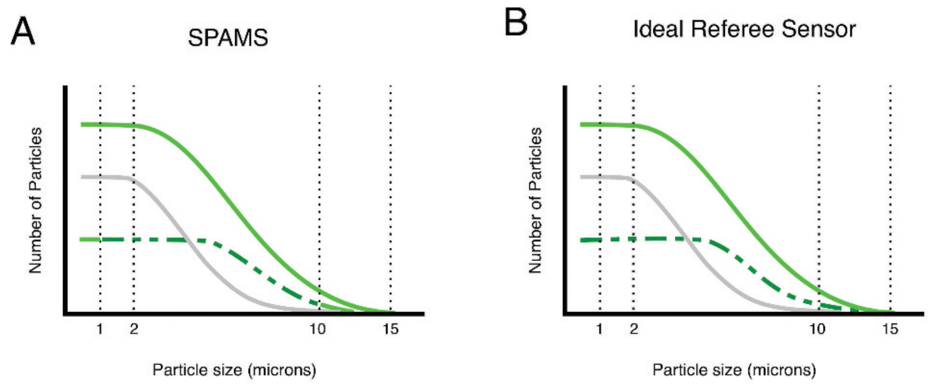


Figure 6. The Single Particle Aerosol Mass Spectrometry system provides ideal discrimination in the 1–10 µm range.

Chemical tags come in a wide range of colors and are also commonly employed in biological research. Unlike protein tags, chemical tags have to be added to each batch of simulant, making them considerably more expensive. However, the upfront cost in man-hours to create a protein-tagged strain is eliminated. It is likely that a number of chemical tags will need to be tested for effectiveness, but many are likely to be suitable (e.g., fluorescein isothiocyanate). Chemical tags can bind to a number of biological molecules (i.e., lipids, carbohydrates, DNA), providing the opportunity to use multiple tags for extra discrimination. A chemical tag will be present in all disseminated particles, meaning that the test aerosol will be easily distinguished from background particles, while agent containing disseminated particles will be indistinguishable from non-agent containing disseminated particles. Unlike protein tags, chemical tags do not require genetic alterations of simulants and are less likely to require regulatory approval before use in breeze tunnels or the field. While the chemical tag will clearly need to be nontoxic, there are commercially available dyes (e.g., food dyes) that could be safely distributed.

Overall, chemical or protein tags offer the most flexible approach to discriminating aerosols, as the fluorescent characteristics of the tag can be custom tailored to any testing environment. In fact, the ideal situation would likely employ both chemical and protein tags, a combination that would allow a modified UV-APS or RAAD-like system to not only discriminate the test aerosol from the background but also provide information on which disseminated particles actually contained agent. Furthermore, fluorescent tagging does not require the development of new referee sensors, as existing UV-APSs could be easily retrofitted to detect a fluorescent tag by swapping out a relatively inexpensive fluorescence filter. Unfortunately, with this flexibility come several drawbacks. Fluorescent tagging would require a substantial amount of development before it could be fully implemented.

Single Particle Aerosol Mass Spectrometry (SPAMS)

The SPAMS system is a biodetector technology currently under development at Lawrence Livermore National Laboratory (LLNL). The concept is similar to the RAAD in that the SPAMS initially collects the sample and then uses a series of laser beams to determine the speed, position, and size of incoming particles. Particles are then subjected to a series of fluorescence measurements (similar to a UV-APS and RAAD) to determine which of the particles might be of biological origin. Particles of interest are then subjected to single particle mass spectrometry for elemental analysis, allowing the system to identify which particles contain agent. Unlike the RAAD, the SPAMS does not use fluorescence data to identify the contents of a particle; all identification data come from mass spectrometry. However, fluorescence data could be incorporated into identification algorithms.

The extra aerosol discrimination provided by mass spectrometry gives the SPAMS system many abilities that are valuable for a referee system. First, SPAMS can differentiate between nonbiological background particles and disseminated aerosols via its fluorescence measurements. Furthermore, it can differentiate between disseminated aerosol particles and biological background via the mass spectrometry data *(Figure 6)*. In theory, these same data can also differentiate between agent containing and non-agent containing particles. In fact, testing of the SPAMS system shows that it can even differentiate to some extent between similar bacterial species. Thus, SPAMS potentially provides more discrimination information than the elemental analysis found in the RAAD.

As with the other systems, SPAMS has a number of limitations. Currently, the system only investigates particles from 1 to 10 μm in size. While this covers the most important range of particles, it does not cover the full spectrum of interest. In addition, the SPAMS is still a proof of concept instrument with no immediate plans for commercial production. A start-up company has been formed to commercialize SPAMS technology (Livermore Instruments), but they are not currently producing substantial numbers of units. Thus, it is unlikely that the T&E community could acquire multiple SPAMS units in the short term. Finally, the rate-limiting step of the system is the mass spectrometry. It can collect these measurements for up to 20 particles per second, similar to the RAAD. This number may be sufficient to accurately estimate the number of particles in the entire sample; however, as the number of particles increases, a smaller fraction of them will undergo mass spectrometry.

The cost of a SPAMS unit is not currently known; however, LLNL has several units that could be delivered to T&E facilities for testing.

Conclusions

As the earlier description of potential referee systems makes clear, there is no currently available perfect solution. However, there are a variety of options that would greatly extend current T&E capabilities. A synopsis of each technology is briefly stated below.

UV-APS

The UV-APS has several strong advantages that provide significant benefits over current refereeing technologies. First, the UV-APS can effectively subtract ambient nonbiological background from outdoor testing environments. In many test scenarios, the UV-APS will be able to provide an accurate estimate of the PSD of a disseminated bioaerosol. Second, it is already commercially available and familiar to the T&E community. This advantage allows the UV-APS to be implemented almost immediately in T&E facilities. However, the UV-APS cannot separate some interferents from disseminated aerosols and has only a limited ability to differentiate between agent containing and non-agent containing particles. Overall, implementation of the UV-APS would be a substantial short-term improvement over current referee equipment.

RAAD

The RAAD combines the fluorescence capabilities of the UV-APS with additional discrimination capabilities provided by SIBS elemental analysis. The RAAD should be able to effectively discriminate between disseminated bioaerosols and ambient biological and nonbiological background. Furthermore, it can differentiate between agent containing and non-agent containing aerosol, allowing for accurate determination of agent concentration and PSDs. Thus, the RAAD appears to be an improvement of the UV-APS. There are, however, two disadvantages. First, it only samples particles between 2 and 10 μm, thus not all particles of interest will be counted. Second, while the RAAD has been extensively tested and is slated for commercial production in 2011, production units are not currently available, though existing units could be shipped to T&E facilities in the short term to help familiarize personnel with the equipment. In summary, the RAAD would be a large step forward in T&E referee equipment and would provide the majority of information of interest to the T&E community. It would likely need to be employed in combination with other particle sizers, however, to capture information on particles outside of the 2–10 μm range. If combined in this manner, the total cost would approximate that of the UV-APS.

Fluorescent tagging

Fluorescent tagging ultimately provides the most flexible solution as a referee sensor technology. Because the tag (chemical, biological, or both) can be tailored to any given situation, it should be possible to accurately measure the disseminated aerosol as well as which particles within that aerosol actually contain agent. Using a proper detector in combination with fluorescent tagging (i.e., a modified UV-APS or the fluorescence stages of the RAAD or SPAMS system), accurate data can be obtained for all the size ranges of interest to the T&E community. However, tagging is the least technologically mature option, as considerable research needs to be done to determine which tags to employ and what detector will sense those tags. In addition, there may be regulatory hurdles involved in dispersing large amounts of genetically or chemically altered aerosols. Thus, tagging is unlikely to be implemented in the short term but could be combined with some of the shorter term solutions (i.e., UV-APS, RAAD) in the future.

SPAMS

The SPAMS system likely collects the most data of any of the potential referee systems identified. Like the RAAD and UV-APS, SPAMS collects fluorescence data to identify particles of biological origin. A subset of these particles is then subjected to mass spectrometry to determine particle identity. This extra discrimination capability should allow the SPAMS system to differentiate between nonbiological background, biological background, agent containing and non-agent containing disseminated aerosols. Thus, the SPAMS system represents a major upgrade from current T&E referee equipment. It should be noted, however, that the SPAMS system only measures particles in the 1–10 μm range, and additional sensors would be necessary to monitor particles outside this range. The largest drawback for the SPAMS system is that there are no immediate plans for commercial production. While a start-up company has been created to commercialize SPAMS technology, the system is still under development. Thus, it is not likely that SPAMS systems could be acquired by the T&E community in the short term. Overall, the SPAMS system is technologically impressive but is several years away from potential implementation as a T&E referee sensor. ❑

Dr. James J. Valdes is a senior research fellow at the National Defense University's Center for Technology and National Security Policy and the U.S. Army's scientific advisor for biotechnology. Dr. Valdes received a doctor of philosophy degree in neuroscience from Texas Christian University and was a postdoctoral fellow at the Johns Hopkins Medical Institutes. He won a bronze medal for the 2002 Army Science Conference, Best Paper (Biotechnology) at the 2008 Army Science Conference, and the 2003 Presidential Rank Award. Dr. Valdes is a cofounder of the International Society for Environmental Biotechnology and the International Society for Biological and Environmental Repositories. He has published more than 120 papers in scientific journals and was a 2009 Presidential Rank Award winner. E-mail: James.Valdes@us.army.mil

Dr. Raymond Mackay was a science advisor and the director of research and technology at the U.S. Army Edgewood Chemical Biological Center. He held the position of adjunct professor at Drexel University. Dr. Mackay earned a doctor of philosophy degree in chemistry from the State University of New York at Stony Brook. Dr. Mackay received numerous awards such as the U.S. Army meritorious Civilian Service Award and the U.S. Army Research and Development Achievement Award. He is the author of 125 publications and 28 reports. Dr. Mackay is currently a physical chemist and an independent contractor. E-mail: rmcmackay@comcast.net

Dr. Joshua L. Santarpia is a member of the senior professional staff of the Johns Hopkins University Applied Physics Laboratory, National Security Technology Department (NSTD) and the supervisor of the Aerosol Sciences section within the Applied Biology, Chemistry, and Nuclear Sciences Group. He is currently involved in the evaluation and integration of biological sensors for building protection, the design of dynamic chemical and biological agent test facilities, and the study of the fate of biological and chemical agents in the atmosphere. Dr. Santarpia received a doctor of philosophy degree in atmospheric sciences from Texas A&M University in 2005. He spent 2 years at Edgewood Chemical and Biological Center (ECBC) as a research scientist on the Aerosol Sciences Team. His work involved test and evaluation of biological aerosol detection and identification systems and method development for evaluation of biological aerosol collection and optical detection systems. Dr. Santarpia's published works include studies of physical and chemical changes observed on ambient submicron aerosol particles, bioaerosol sensor test methodology, and characterization of ambient biological aerosol. Dr. Santarpia is also an adjunct member of the graduate faculty in the University of Maryland Baltimore County Civil and Environmental Engineering department and a visiting scientist at the Edgewood Chemical and Biological Center BSL-3 laboratory.

Dr. Stephen Adams received his bachelor of science degree from Kings College, University of London, U.K., in 1981 and his doctor of philosophy degree from the University of Cambridge, U.K., in 1985. He joined the research group of Roger Y. Tsien (Nobel Prize in Chemistry, 2008), first at UC Berkeley, and then in 1989 at UC San Diego, where he is currently a project scientist in the Department of Pharmacology. He has published over 50 peer-reviewed papers including the first FRET reporter in living cells, photosensitive calcium chelators and indicators, GFP-based reporters, and small molecule genetically targeted protein tags. E-mail: sadams@ucsd.edu

Ms. Ewelina Tunia is a research assistant at the National Defense University Center for Technology and National Security Policy. She earned her master's degree in security studies from Georgetown University's Edmund A. Walsh School of Foreign Service and a bachelor's degree in political science from Hunter College, City University of New York. Previously, she held internships at Amnesty International and the United Nations and was an intelligence analyst for the U.S. Army National Guard. E-mail: tuniae@ndu.edu

Dr. Andrew Mara is the 2008–2010 AAAS science and technology policy fellow at the Center for Technology and National Security Policy. Dr. Mara earned his doctor of philosophy degree in cellular biology from Yale University where he developed cutting-edge imaging techniques to examine emergent, self-organizing systems. This work was recognized with the J. Spangler Nicholas Prize for best thesis. He was an inaugural recipient of a Department of Homeland Security (DHS) Graduate Fellowship, DHS Dissertation Grant, and Yale University's Chair Fellowship. He has multiple scientific and policy publications, which have been presented at international conferences.

References

Agranovski, V., and Z. D. Ristovski. 2005. Real-time monitoring of bioaerosols: Capability of the UV-

APS to predict the amount of individual microorganisms in aerosol particles *Journal of Aerosol Science* 36 (5–6), 665–676.

Cremers, D. A., and L. J. Radziemski. 2006. *Handbook of laser-induced breakdown spectroscopy*. Hoboken, NJ: John Wiley & Sons.

Gonser, K. R., S. D. Christensen, and J. A. Guicheteau. 2004. Chemical and biological point sensors for homeland defense II. In *Proceedings of the SPIE*, ed. A. J. Sedlacek III, S. D. Christesen, T. Vo-Dinh, and R. J. Combs, vol. 5585, 136–142.

Hindson, B. J., A. J. Makarewicz, U. S. Setlur, B. D. Henderer, M. T. McBride, and J. M. Dezenitis. 2005. APDS: The autonomous pathogen detection system. *Biosensors and Bioelectronics* 20 (10), 1925–1931.

Miziolek, A. W., and V. Palleschi. 2006. *Laser-induced breakdown spectroscopy*. New York: Cambridge University Press.

Nagai, T., K. Ibata, E. S. Park, M. Kubota, K. Mikoshiba, and A. Miyawaki. 2002. A variant of yellow fluorescent protein fast and efficient maturation for cell-biologic amplification. *Nature Biotechnology* 20 (1), 87–90.

Valdes, J. J., J. Mohr, R. Mackay, E. Tunia, and A. Mara. 2010. Total agent per liter of air with particle size distribution (TALAp): A new unit of measure for the test and evaluation of biodetectors. *ITEA Journal* 31 (4), 417–425.

Acknowledgments

The views expressed in this article are those of the authors and do not reflect the official policy or position of the National Defense University, the Department of Defense or the U.S. Government. All information and sources for this paper were drawn from unclassified materials.

ITEA Journal 2011; 32: 191–200

The Multi-Relationship Evaluation Design Framework: Producing Evaluation Blueprints to Test Emerging, Advanced, and Intelligent Technologies

Brian A. Weiss

National Institute of Standards and Technology, Gaithersburg, Maryland

Linda C. Schmidt

University of Maryland, College Park, Maryland

This article introduces the Multi-Relationship Evaluation Design (MRED) framework whose objective is to take uncertain input data and automatically output comprehensive evaluation blueprints complete with targeted evaluation elements. MRED is unique in that it characterizes the relationships among the evaluation elements and will address their uncertainties in addition to those tied to the evaluation input. These terms and their relationships are applied in an example evaluation design of an emerging technology.

Key words: Emerging technology; evaluation blueprints; intelligent technology; performance.

Intelligent system advances are continuously occurring across a range of fields in the military, automotive, and manufacturing communities. As these new technologies emerge, it becomes critical to evaluate their performance to both (a) inform the technology creators of deficiencies and obtain end-user feedback so that enhancements can be made and (b) validate the technology's ultimate capabilities so that buyers and system users know exact system capabilities. The former takes place in formative evaluations where adjustments and enhancements can be made in upcoming designs; the latter happens in summative evaluations to enable buyers and technology users to see the extent of the technology's capabilities. These two types of evaluations can be designed to comprise a few simplistic tests of key capabilities of the overall technology. Alternatively, evaluations can become highly complex events, testing numerous components and capabilities along with the entire system. Typically, tests of advanced and intelligent systems tend to be elaborate since the technologies themselves are usually complex in nature.

The National Institute of Standards and Technology (NIST) created the System, Component, and Operationally Relevant Evaluation (SCORE) framework to evaluate numerous emerging and intelligent systems at various levels (Weiss and Schlenoff 2008). Specifically, SCORE provides a set of guidelines to aid test designers in creating evaluation plans. SCORE has been effectively applied to 15 evaluations across several technologies (Schlenoff et al. 2009; Weiss and Schlenoff 2009). SCORE-enabled tests have yielded extensive quantitative and qualitative data. SCORE has proven to be valuable to technology developers, evaluation designers, potential end-users, and funding sponsors. The research described here draws upon that success to introduce a new evaluation framework that will automatically generate evaluation blueprints or "test plans." This new evaluation design tool is known as the Multi-Relationship Evaluation Design (MRED) framework. MRED's ultimate objective is to take inputs from specific groups, each complete with their own varying uncertainties, and output an evaluation blueprint that specifies all characteristics of the tests. MRED is a work in progress where the authors are identifying an adequate set of blueprint constituent elements and their interrelationships.

The overall model encompassing the MRED framework was introduced in Weiss et al. (2010) along with several evaluation blueprint elements. Since that initial work, further blueprint elements have been defined (Weiss and Schmidt 2010). In this article, we introduce the International Test and Evaluation Association (ITEA) community to the blueprint elements identified to date, the relationships among

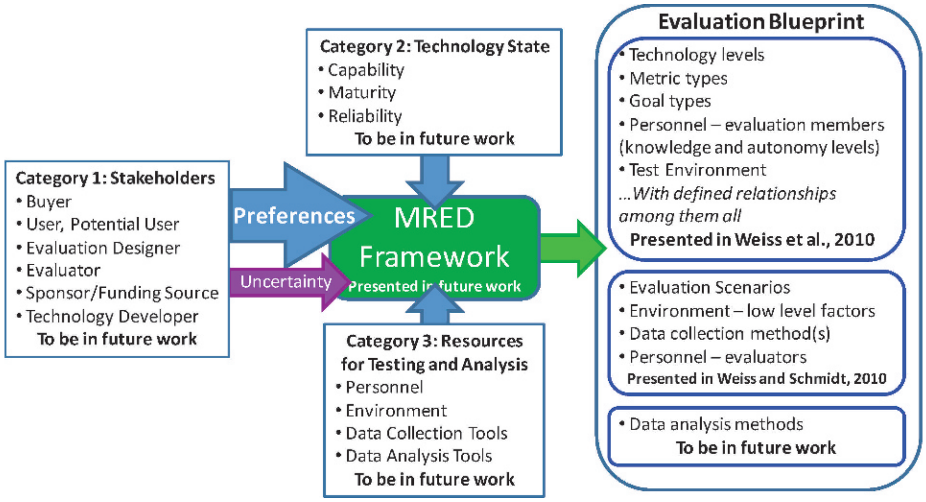


Figure 1. Evaluation design model surrounding the multi-relationship evaluation design framework including inputs and outputs.

them, and their major interdependencies. In addition, we apply the MRED model to the evaluation design of pedestrian and object tracking algorithms, whose test plans were created without the use of any formal evaluation design framework.

MRED

The MRED framework resides within a model that contains both the significant design inputs into the planner and the output, "evaluation blueprint." The overall model is shown in *Figure 1*. The evaluation blueprints have been defined previously in Weiss et al. (2010) and Weiss and Schmidt (2010). However, the three input categories (a) stakeholders, (b) technology state, and (c) resources for testing and analysis are highlighted so that the model is easier to understand.

Input categories

The MRED model recognizes three crucial input categories or personnel groups interested in a technology's evaluation. Each category is briefly described in the following subsections.

Category 1 – Stakeholders

Evaluation stakeholders are organized into six categories of parties interested in the technology's evaluation. Most (if not all) of the stakeholders are active in a technology's evaluation design *(Table 1)*. Each stakeholder's preferences regarding the test plan may evolve over time, which leads to uncertainty based upon these changing preferences.

There may be some overlap among the stakeholders. The range of potential relationships among the stakeholders can be found in Weiss et al. (2010).

Category 2 – Technology state factors

This category contains the factors that characterize the system's state at the time of its test. These factors are presented in *Table 2*. Note that all three of these factors may change between the time when testing is first discussed and planned to the moment when the tests are executed. The level of these factors determines whether or not the testing data can be used for formative or summative evaluations (Weiss and Schmidt 2010).

Category 3 – Resources for testing and analysis

This last input group is composed of various types of material, personnel, and technology to be committed to the evaluation exercise and data analysis. Resource availability (or lack thereof) and limitations can have a

Table 1. Technology evaluation stakeholder groups (Weiss and Schmidt 2010).

Stakeholders group	Who they are…
Buyers	Stakeholder purchasing the technology
Users, potential users	Stakeholder that will be or are already using the technology
Evaluation designers	Stakeholder creating the test plans by determining MRED inputs
Evaluators	Stakeholder implementing the evaluation test plans
Sponsors/funding sources	Stakeholder paying for the technology development and/or evaluation
Technology developers	Stakeholder designing and building the technology

MRED, Multi-relationship evaluation design.

Table 2. Technology state factors.

Factors	What is it…	Why it matters…
Reliability	Technology's ability to yield the same or compatible results as in previous tests	Changes will impact test comparisons, and the technology has undergone previous testing where the output test data have been used to iterate upon the design.
Capability	Technology's ability to be evaluated under certain conditions and/or used in a specific functionality	Considers if the technology is robust enough or if its current level of development is sufficient to undergo a specific test.
Maturity	Technology's state or quality of being fully developed	Considers if the technology is equipped with all of its intended functionality or if only a subset of expected features is operational at the time of testing.

N/A, not applicable; DM, decision-making.

tremendous influence on the final evaluation design. These resources are highlighted in *Table 3*.

Blueprint outputs

The blueprint outputs have been presented in detail in prior work including Weiss et al. (2010), Weiss and Schmidt (2010), and Weiss and Schlenoff (2008) but are highlighted here so that the model outputs are understood as well as the crucial inputs.

Technology levels

A technology or system is made up of constituent components and can be evaluated at multiple levels. There are several terms related to technology levels, and they are defined as follows:

- *System* is a group of cooperative or interdependent components forming an integrated whole intended to accomplish a specific goal.
- *Component* is an essential part or feature of a system that contributes to the system's ability to accomplish a goal.
- *Sub-component* is an element, part, or feature of a component.
- *Capability* is a specific ability of a technology where a system is made up of one or more *capabilities*. A capability is provided by either a single component or multiple components working together.

Table 3. Resources of testing and analysis.

Resources	Description
Personnel	Individuals who will use the technology, those who will indirectly interact with the technology, those who will collect data during the test, and those who will analyze the data following the tests
Test environment	The physical venue, supporting infrastructure, artifacts, and props that will support the tests
Data collection tools	The tools, equipment, and technology that will collect quantitative and/or qualitative data during the tests
Data analysis tools	The tools, equipment, and technology capable of producing the necessary metrics from the collected evaluation data

Metric types

There are two metric types: technical performance and utility assessments.

- *Technical performance* refers to metrics related to quantitative factors (such as accuracy, precision, time, distance, etc.). These metrics may be needed by the program sponsor to get a status of the technology's current performance, update the technology developers on their design, etc.
- *Utility assessments* refer to metrics related to the qualitative factors that gauge the condition or status of being useful and usable to the target user population. Like technical performance, these metrics may be of value to any or all of the stakeholders.

Goal types

Goal types are combinations of technology levels and desired metrics (Schlenoff et al. 2009; Weiss et al. 2008). There are five goal types that will be output from the MRED framework including three that capture quantitative technical performance data and two that capture qualitative utility assessments *(Figure 2)*.

Each of these goal types requires different features and properties within an evaluation making it possible to create a test plan that captures data to satisfy multiple goal types. This is another unique feature of MRED in that it will support the generation of a set of test plans to capture a range of goal-type data based upon the input.

MRED will be capable of outputting a set of blueprints, each able to capture data for a different goal type. Evidence of MRED's capability in creating evaluation plan blueprints is provided in Section 0 by an initial application of the model to a speech-to-speech technology. MRED's blueprints will be com-

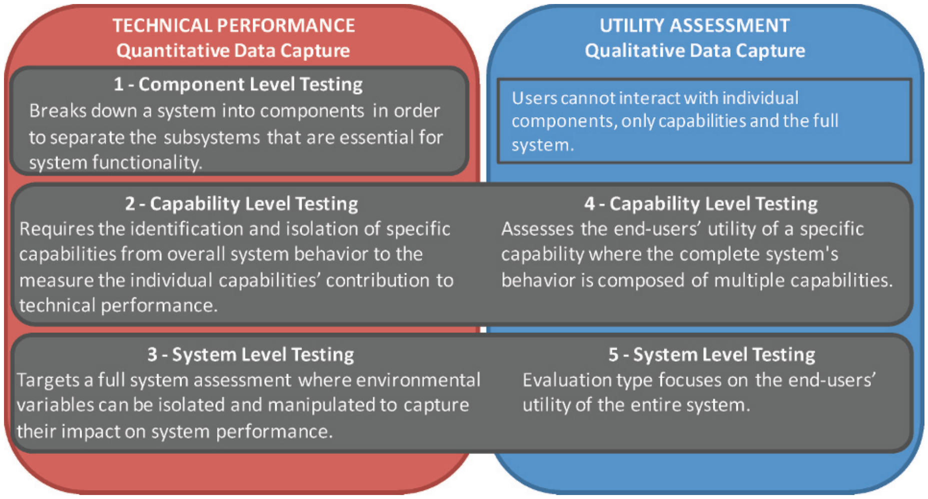


Figure 2. Multi-relationship evaluation design framework output goal types.

pared to a previous evaluation that was driven by the SCORE framework (Weiss et al. 2010).

Personnel – evaluation members

Numerous individuals and groups are necessary to produce an effective evaluation. They are classified into two distinct categories: primary (direct interaction) technology users and secondary (indirect interaction or evaluation support).

The primary technology users are identified as *tech users*. These individuals directly interact with the technology during the evaluation. They receive any training necessary to use the technology and are responsible for engaging/disengaging the technology's usage during the test event. There are multiple classes of tech users as shown in *Table 4*, which also presents the goal types that the various tech users may participate in given their characteristics. Note that the information provided in this table is intended to demonstrate the superset of possibilities given that all tests are technology-dependent.

Secondary personnel are those that indirectly interact with the technology during the evaluation, and they fall into three categories: team members, participants, and evaluators.

- *Team members* are individuals that work with tech users during the evaluation as they would to realistically support the use-case scenario in which the technology is immersed. Team members may or may not be in a position to indirectly or directly interact with the technology during the evaluation, but they are often in a position to observe a tech user's interactions with the system.
- *Participants* are individuals that indirectly interact with the technology during an evaluation. Typically, participants are given specific tasks to either interact with the tech users and/or with the environment, but not with the technology.
- *Evaluators* are members of the evaluation team present within the test environment that task the participants and/or capture data but do not interact with the technology. Depending upon

Table 4. Primary evaluation personnel using the technology during testing.

		Applicable goal types for participation				
		Technical performance			Utility assessment	
Primary personnel	Description	Component	Capability	System	Capability	System
Tech users: end users	Individuals who are the intended users for the technology	No	Yes	Yes	Yes	Yes
Tech users: trained user	Individuals selected to be tech users, yet who are not end users	Yes	Yes	Yes	Yes	Yes
Tech user: tech developer	Members of the research and development organization(s) that developed the technology under evaluation	Yes	Yes	Yes	No	No

Table 5. Relationships among personnel, knowledge, and autonomy levels (Weiss et al. 2010).

	Tech user	Team member	Participant
Technical knowledge	Low – Med – High	Low – Med – High	Low – Med – High
Operational knowledge	Low – Med – High	Low – Med – High	Low – Med – High
DM Autonomy – tech.	None – Low – Med – High	None – Low – Med – High	N/A
DM Autonomy – env.	None – Low – Med – High	None – Low – Med – High	None – Low – Med – High

Tech., technical; N/A, not applicable; env., environmental.

the test, the evaluator may interact with the tech user to capture data.

Significant relationships exist between technology levels, metric types, tech users, and test environments. These relationships among these primary and secondary technology user groups have been highlighted in previous efforts (Weiss et al. 2010). Likewise, relationships between personnel—evaluation members, knowledge levels, and autonomy levels—are documented in this article. These relationships define numerous constraints that must be satisfied in any feasible evaluation blueprint.

Knowledge levels. The tech users, team members, and participants involved in the evaluation have various levels of knowledge about aspects of the system and testing conditions within two specific areas. The levels are defined as follows:

- *Operational knowledge* is the level of practical information and experience an individual has about the *actual* environment, the intended use-case situations for the technology, and other preexisting technologies that the technology under test leverages and/or supports. Varying levels of operational knowledge can be attained through real-world experience, repetitive training, trial and error exercises, etc.
- *Technical knowledge* is the level of information and experience an individual has about the technology itself and how it should be employed to maximize success.

Autonomy levels. Additionally, the tech users and participants within the evaluation have a range of decision-making (DM) autonomy. Autonomy scope and their levels are set by MRED for each evaluation. Personnel could be fully restricted in their decision making (i.e., no DM autonomy), which requires scripted actions. Alternatively, personnel may have unbounded decision-making authority, where each participant is free to exercise their own judgment given their various knowledge levels. Specifically, there are two types of DM autonomy: technical and environmental.

- *DM Autonomy, technical,* refers to the level of authority that the tech users have in operating the technology. Depending upon the specific evaluation, some tech users could be instructed to only use certain features of a technology, while others may be given the freedom to use any or all features as they see fit.
- *DM Autonomy, environmental,* refers to the level of authority that the tech users and participants have in interacting with each other and the environment.

Autonomy levels must be equal or lower in value than their partner knowledge levels. Determination of autonomy levels is governed by multiple factors including evaluation goal types, tech user class, etc. The potential knowledge and autonomy levels for the evaluation participants are shown in *Table 5.*

Test environments

The setting in which the evaluation occurs can have a significant effect on the data since the environment can influence the behavior of the personnel and can limit which levels of a technology can be tested. MRED defines three distinct environments: lab, simulated, and actual.

- *Lab* refers to a controlled environment where test variables and parameters can be isolated and manipulated to determine how they impact system performance and/or the tech user's perception of the technology's utility.
- *Simulated* refers to an environment outside of the lab that is less controlled and limits the evaluation team's ability to control influencing variables and parameters, since it tests the technology in a more realistic venue.
- *Actual* refers to the domain of operations in which the system is designed to be used. The evaluation team is limited in the data they can collect, since they cannot control environmental variables.

Evaluation scenarios

The evaluation scenarios govern exactly what the technology will encounter and the challenges it will have

Table 6. Relationship among the scenarios, environments, knowledge and decision-making autonomy (Weiss and Schmidt 2010).

		Tech user's knowledge level		Tech user's decision-making autonomy	
Evaluation scenarios	**Test environments**	**Technical**	**Operational**	**Technical**	**Environmental**
Technology-based	Lab, Simulated	Med – High	Low – Med – High	None – Low	None – Low
Task/activity-based	Lab, Simulated, Actual	Low – Med – High	Low – Med – High	Low – Med – High	Low – Med – High
Environment-based	Simulated, Actual	Med – High	Med – High	Med – High	Med – High

to perform within the identified test environments. Three types of evaluation scenarios are identified below. Each is unique in the relationship established with the tech user considering their knowledge level, decision-making autonomy, and test environment. The three evaluation scenario types are listed below, and the relationships are shown in *Table 6*.

- *Technology-based* evaluation scenarios feature specific instructions to the user in how they should use the technology within the testing environment.
- *Task/activity-based* evaluation scenarios require the user to complete a specific task within the environment, where they may use the technology as they see fit.
- *Environment-based* evaluation scenarios enable the user to perform the relevant activities within the environment based upon advanced operational knowledge.

Further details can be found in Weiss and Schmidt (2010).

Explicit environmental factors

The explicit environmental factors are significant characteristics within the environment that impact the technology, thereby influencing the outcome of the evaluation. These factors pertain to the overall physical space (i.e., participants, structures, and any integrated props and artifacts). These factors are broken down into two constituent characteristics: feature density and feature complexity. These two characteristics combine to form the overall complexity.

- *Feature density* refers to the number of features within the test environment given the size of the test area. The greater the feature density, the more challenging it is for a technology to effectively and efficiently identify objects/events/activities, interact with, and operate within the test environment.
- *Feature complexity* refers to the complexity of various measurable features within the environment. Similar to feature density, the greater the feature complexity, the more difficult it is for the technology to accurately and appropriately operate and be beneficial to the tech user.
- *Overall complexity* refers to the global influence of feature density and feature complexity within the testing environment.

Figure 3 presents the relationships between the test environment and the explicit environmental factors. Additional information on this blueprint output and these relationships can be found in Weiss and Schmidt (2010).

It is currently expected that the overall complexity will be limited in the lab, given the nature of this highly controlled environment.

Data collection methods

Data collection methods are used to capture experimental and ground truth data depending upon

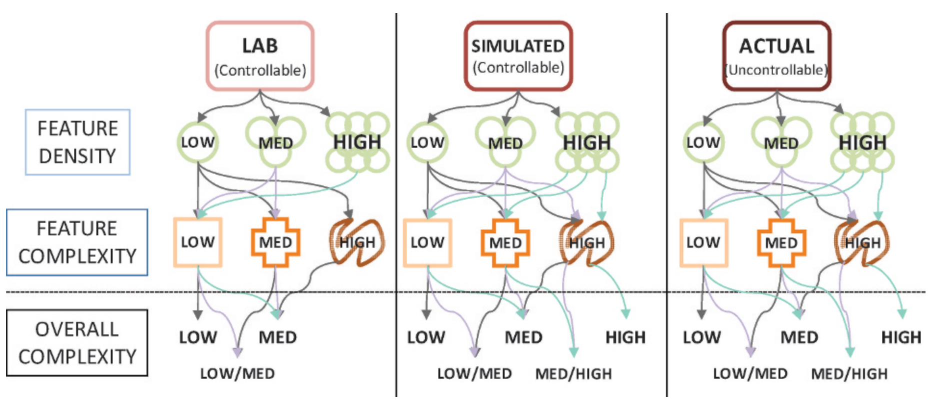


Figure 3. Relationships between the test environment and its explicit environmental factors (Weiss and Schmidt 2010).

Table 7. Example data collection methods from various collector locations and modes (Weiss and Schmidt 2010).

Data collection methods: examples		
Collector Location	**Modes – automated**	**Modes – manual**
From the technology	Sensors and/or tool collecting technical information during the evaluation Output of log files following the evaluation	Evaluator capturing data with sensors, tools Evaluator making notes of behavior
From the tech user	Sensors (e.g., helmet camera) attached to the tech user that collect data during testing	Surveys prior to and/or following the evaluation Interviews prior to and/or following the evaluation Verbal and/or physical feedback provided by the tech user during the evaluation (e.g., thumbs-up or thumbs-down at key points
From the team member	Sensors (e.g., microphone) attached to the team member that collect data during testing	Surveys prior to and/or following the evaluation Interviews prior to and/or following the evaluation Verbal and/or physical feedback provided by the tech user during the evaluation (e.g., thumbs-up or thumbs-down at key points
From the participant	None (usually)	Surveys prior to and/or following the evaluation Interviews prior to and/or following the evaluation
From a different perspective	Sensors (e.g., radar gun, thermal camera, motion detector) set up throughout the environment that collect data during testing	Evaluation personnel stationed in various parts of the environment taking notes and/or manually using a sensor and/or tool to collect data.

the technology being evaluated and the specified test environment. No matter what type of tools are used, data collection methods are characterized by factors that influence the techniques being employed. These two factors include mode and collector location and are presented below:

- *Mode* refers to the type of data collection method that will be employed. Specifically, there are two different types of modes that can collect data. Automated mode involves collecting data with a calibrated technology independent of the system undergoing testing. Manual mode features an evaluator actively managing a calibrated technology or collecting data by hand.
- *Collector location* refers to the placement of the mode relative to the technology under test.

The various collector locations and some examples of the two types of modes are presented in *Table 7*. Greater detail regarding data collection methods can be found in Weiss and Schmidt (2010).

Personnel – evaluators

There are three classes of evaluation personnel that are necessary to ensure that the evaluation proceeds accordingly to plan and that the necessary data are captured to evaluate a technology's performance. The evaluators fall into three classes: data collectors, test executors, and safety officers.

- *Data collectors* are responsible for either setting up/implementing automated collection methods and/or performing manual collection methods. This class of evaluators is also responsible for collecting experimental data directly from the technology at the conclusion of each test scenario (as necessary).
- *Test executors* are responsible for initiating the test, including instructing participants on when to engage in their specified activities within the environment.
- *Safety officers* are solely responsible for ensuring the safety of all personnel within the test environment along with protecting the technology and the environment, itself.

Greater discussion of these personnel classes can be found in Weiss and Schmidt (2010).

Application of MRED

In this section, the model's currently defined blueprint elements and relationships are applied to a technology that National Institute of Standards and Technology (NIST) personnel are evaluating. NIST and members of the Army Research Laboratory's (ARL's) Collaborative Technology Alliance (CTA) are currently testing multiple pedestrian tracking algorithms whose evaluation design and implementation are conducted jointly by NIST and CTA (Bodt et al. 2009). Each tracking algorithm uses Laser Detection and Ranging (LADAR) and video sensor data taken from a moving vehicle. This test platform moves within the test environment, where the vehicle-

Table 8. Multi-relationship evaluation design personnel matrix as matched Army Research Laboratory's Collaborative Technology Alliance evaluation.

	Tech-user: trained user	Team member	Participant
Technical knowledge	Medium	N/A	N/A
Operational knowledge	Low	N/A	Low
DM autonomy – technical	None	N/A	N/A
DM autonomy – environmental	None	N/A	Low

mounted sensors capture and send data to the on-board detection and tracking algorithms.

The NIST/CTA team have jointly planned and implemented several evaluations from 2007 through 2010. To test MRED and identify any shortcomings, MRED terminology will be used to describe an evaluation blueprint to match the completed testing plans.

Technology level, metric type, and goal type

Presently, the ARL CTA is isolating the pedestrian detection and tracking algorithms for evaluation in a manner that will yield quantitative technical performance metrics. NIST's involvement with the program has centered on conducting field exercises within the goal type of capability level testing—technical performance. These field exercises capture the technical data required to assess the performance of the CTA team's algorithm. They also provide data to support future algorithm development and produce performance analyses. These exercises are conducted by having a sensor-laden vehicle drive in a pedestrian- and obstacle-filled environment. The experimental algorithm output can be collected and measured against evaluation team–captured ground truth. Numerous quantitative technical performance metrics were produced at the conclusion of these tests. These included true positive, false positive, misclassification, first detection, persistence detection and accuracy of detected position, and velocity.

Personnel – evaluation members

For the evaluation type defined above, personnel-evaluation members and their characteristics were identified by the authors and are shown in the MRED personnel matrix in *Table 8*. Specifically, this blueprint mandated that a trained user engage and disengage the technology during the tests. It is currently premature to identify the exact intended user group since this capability is not fully developed and will be ultimately integrated into a greater system for its regular operations. Algorithms from multiple organizations were tested in parallel and activated by the same tech user to yield an objective approach.

The evaluation participants were individuals that acted as pedestrians. The pedestrians were given a specific path within the environment that they walked during the tests. Practice runs were conducted so that the walkers could determine their pace, better enabling them to complete their path in a prescribed amount of time. The specific evaluator roles will be discussed in the Personnel-evaluators section.

Test environment

These tests were performed in simulated environments that featured some Military Operations in Urban Terrain (MOUT) characteristics. The evaluation team controlled the environment during the evaluations, enabling them to get very detailed ground-truth data, including walking paths, obstacle locations, and vehicle paths.

Evaluation scenarios

The evaluation scenarios designed by NIST/CTA personnel for their test effort can be classified as technology-based when interpreted using the MRED model. The tech user is restricted in how they can interact with the technology, meaning they are only allowed to engage the technology at the beginning of a run and to disengage it at the run's conclusion. According to MRED's definition of this blueprint element, the CTA evaluation scenarios are neither task/activity-based nor environment-based since the tech user has no freedom to interact with the environment or has to complete a specific mission with the technology during the testing.

Explicit environmental factors

Assigning the appropriate MRED model explicit environmental factors of this test effort based the test events that the NIST/CTA devised is quite challenging. Since NIST personnel have not been involved in previous algorithm testing that took place in a lab or other simulated environment, it is difficult to ascertain how the feature density, feature complexity, and overall complexity compare in the current test environment. Also, since these algorithms will be incorporated into a larger technology for its envisioned use, it is challenging to anticipate what the actual environment will be, especially since the larger systems are somewhat unknown at this point.

The MOUT simulated test environment that the ARL CTA/NIST team previously tested could be classified as having an overall complexity of "medium," where feature density and feature complexity were both globally "medium." However, looking at specific artifacts and personnel activities within the environment, a case could be made that local feature density ranged from "low" to "medium," since multiple personnel were in close proximity to one another in some spots while other personnel stood by themselves in other spots. Comparably, it can be stated that local feature complexity also ranges from "low" to "medium" considering that there were various environmental features present including several rectangular buildings and about a dozen participants.

Data collection methods

In order to test the pedestrian-tracking algorithms, the CTA/NIST team deployed numerous automated data collection methods, necessary to attain the required metrics, from numerous collector locations. Specifically, an Ultra-Wide Band (UWB) tracking system was deployed to capture position ground truth data of the test vehicle, key environmental features, and pedestrians that were within the testing environment. This was used to capture quantitative technical performance data of the sensors and algorithms in order to generate the necessary evaluation metrics. Numerous cameras were set up throughout the test environment to collect visual position data of test vehicle, key environmental features, and pedestrians. Both the UWB and camera data were used as ground-truth.

The experimental data featured each algorithm reporting detection information such as positions and velocities of the humans at the end of each CTA algorithm cycle. The underlying assumptions for the outputs of the algorithms included the following:

- Only obstacles seen and classified as human were reported.
- Unique identification numbers were assigned to individual algorithm detections within a run.
- Algorithms demonstrated tracking of an individual by maintaining the same subject identification in successive frames.
- Algorithms also reported velocity of the detected humans.

Since the fixed test area was instrumented to capture ground-truth data, all detections were excluded if they occurred outside the test area. The correspondence algorithm found the correspondence between the detections and the ground-truth based on location and time stamp (Bodt et al. 2009). Detections were compared with all the ground-truth objects on the course to attain the desired metrics noted in the Technology level, metric type and goal type section.

Personnel – evaluators

The ARL CTA/NIST test effort featured both data collectors and test executors conducting the testing. The data collectors included personnel who deployed and calibrated the UWB tracking system and cameras prior to the test event. These same personnel also managed the UWB tracking system and cameras during the evaluation. Numerous test executors played a significant role in the evaluation. One individual signaled the start and conclusion of the tests, and another individual signaled the pedestrians to walk in their prescribed paths. A third test executor was employed to signal when the vehicle should begin its motion and when the sensors and algorithms under test should be activated. The evaluation also featured several safety officers stationed throughout the environment and at key locations along the test environment's perimeter to prevent non–evaluation personnel or vehicles from entering.

Conclusion

The test effort described in the previous section demonstrated that the MRED blueprints are broad enough to align with technology evaluations designed by NIST/ARL. In the future, additional technology test plans will be generated to further validate its usefulness. The next steps for developing MRED are to finalize the last of the blueprint element definitions and continue to verify the current model against evaluation designs created by NIST personnel. Once MRED's blueprint elements are fully identified, the model will continue to be explored through the detailed definition of the three key evaluation input categories and their interrelationships. Given that each of these input categories is nondeterministic in nature (based upon human preference, unknown technology states, or uncertain resource availabilities), uncertainty will be considered. Once the inputs are clearly stated, the details of the MRED framework itself will be outlined and specified. Once proven successful, uncertainty will be factored. It is envisioned that MRED will be an invaluable tool in creating complex evaluation designs of advanced and/or intelligent systems, allowing evaluation designers to be more effective and efficient in producing and implementing the appropriate tests. ❑

Brian A. Weiss has been a mechanical engineer at NIST in Maryland since 2002. His focus is the *development and implementation of performance metrics to quantify technical performance and assess end-user utility of intelligent systems throughout various stages of development. He has a bachelor of science degree in mechanical engineering and a professional master of engineering degree from the University of Maryland and is working towards his doctor of philosophy degree in mechanical engineering with the University of Maryland. E-mail: brian.weiss@nist.gov*

DR. LINDA C. SCHMIDT is an Associate Professor at the University of Maryland. She holds a doctorate in mechanical engineering from Carnegie Mellon University and bachelor of science and master of science degrees in industrial engineering from Iowa State University. Dr. Schmidt is active in teaching design research in theory and practice. She also co-authored texts on engineering decision making and product development. E-mail: lschmidt@umd.edu

References

Bodt, Barry, Richard Camden, Harry Scott, Adam Jacoff, Tsai Hong, Tommy Chang, Richard Norcross, Tony Downs, and Ann Virts. 2009. Performance measurements for evaluating static and dynamic multiple human detection and tracking systems in unstructured environments. In *Proceedings of the 2009 Performance Metrics for Intelligent Systems (PerMIS) Workshop*, September 21–23, Gaithursburg, MD, ed. Raj. Madhavan and Elena Messina, 174–181. New York, NY: ACM.

Schlenoff, Craig I., Brian A. Weiss, Michelle P. Steves, Gregory A. Sanders, Fred Proctor, and Ann M. Virts. 2009. Evaluating speech translation systems: Applying SCORE to TRANSTAC technologies. In *Proceedings of the 2009 Performance Metrics for Intelligent Systems (PerMIS) Workshop*, September 21–23, 2009, Gaithursburg, MD, ed. Raj Madhavan and Elena Messina, 237–244. New York, NY: ACM.

Weiss, Brian A., and Craig I. Schlenoff. 2008. Evolution of the SCORE framework to enhance field-based performance evaluations of emerging technologies. In *Proceedings of the 2008 Performance Metrics for Intelligent Systems (PerMIS) Workshop*, August 19–21, Gaithursburg, MD, ed. Raj Madhavan and Elena Messina, 1–8. New York, NY: ACM.

Weiss, Brian A., and Craig I. Schlenoff. 2009. The impact of scenario development on the performance of speech translation systems prescribed by the SCORE framework. In *Proceedings of the 2009 Performance Metrics for Intelligent Systems (PerMIS) Workshop*, September 21–23, 2009, Gaithursburg, MD, ed. Raj Madhavan and Elena Messina, 252–259. New York, NY: ACM.

Weiss, Brian A., and Linda C. Schmidt. 2010. The Multi-Relationship Evaluation Design framework: Creating evaluation blueprints to assess advanced and intelligent technologies. In *Proceedings of the 2010 Performance Metrics for Intelligent Systems (PerMIS) Workshop*, September 28–30, 2010, Gaithursburg, MD, ed. Raj Madhavan, Elena Messina, and Edward Tunstel. New York, NY: ACM.

Weiss, Brian A., Linda C. Schmidt, Harry A. Scott, and Craig I. Schlenoff. 2010. The Multi-Relationship Evaluation Design framework: Designing testing plans to comprehensively assess advanced and intelligent technologies. In *Proceedings of the ASME 2010 International Design Engineering Technical Conferences (IDETC) – 22nd International Conference on Design Theory and Methodology (DTM)*, August 15–18, 2010, 603–616. Montreal, Quebec, Canada. New York, NY: ASME.

Acknowledgments

The authors would like to thank Harry Scott, the NIST project leader for the ARL CTA effort, for his continued support throughout this work. Further, the lead author would like to thank Tsai Hong for providing additional information regarding the CTA algorithm output and corresponding metrics. Also, the lead author gives thanks to Craig Schlenoff of NIST for his support and encouragement.

Certain commercial companies, products, and software are identified in this article in order to explain our research. Such identification does not imply recommendation or endorsement by NIST, nor does it imply that the companies, products, and software identified are necessarily the best available for the purpose.

ITEA Journal 2011; 32: 201–205

Analytical Tools of Mathematics and Statistics (ATOMs)—Small Steps to Innovation

Mark H. Jones Jr.
Mathematical Christian Consulting, Murfreesboro, Tennessee

In spite of unprecedented technological capabilities, the aerospace community has faced several widely publicized failures and accidents in the recent past. This article describes a simple flight test program and uses this example to illustrate five elementary changes that the test and evaluation community must implement incrementally. These four analytical tools and one technological capability will allow us to evolve our decision making capabilities in the presence of uncertainty and data abundance and equip leaders for effective and efficient decision making in the test and evaluation industry.

Key words: Aviation; data science; decision theory; design of experiments; incremental innovation; shared information; social aggregation; test and evaluation failures; uncertainty analysis.

The aerospace community has faced a string of especially notable catastrophes in the past 2 years that suggest a weakness in our test and evaluation (T&E) methodology. For example, on March 25, 2009, the Associated Press reported that an F-22 crashed in the Mojave Desert just north of Edwards Air Force Base on what was considered by many to be a routine test mission, a profile similar to one flown many times before (Associated Press 2009). Additionally, in March 2009, Internet aviation news source AVWeb published the account of an experimental Cessna Skycatcher Light Sport Aircraft crash during spin certification testing, the second in 6 months (Niles 2009).

More recently, during the second half of last year, two Rolls Royce turbofan engines had uncontained failures: one on a test stand in Derby (Mustoe 2010) and one on a Qantas flight in an Airbus 380 (Peterson 2010); both of these were widely reported in news media worldwide. Also, the *Wall Street Journal* recently documented that midair collision avoidance systems on 9,000 U.S. aircraft are faulty, with one aircraft symbol disappearing from a display for over 40 seconds (Cohen 2010). These examples do not begin to exhaust the list of high-profile, flight test–related, alarming news.

Contrast all of that with our present situation and capabilities. Never before in the history of test and evaluation have we had the computational power and technological sophistication and broad professional expertise that we now have. Our ability to accurately model complex engineering problems is unprecedented. We have Computational Fluid Dynamics (CFD) to visualize and predict the most nonlinear and turbulent aerodynamics, very high fidelity hardware-in-the-loop simulators to rigorously certify our most modern avionics, and unprecedented ability to instrument test vehicles and monitor test data. We have near-supercomputer abilities with networked laptops, cloud computing, and rapidly advancing programming techniques. With these capabilities, why are we seeing such a dramatic breakdown in the development of these aeronautical systems and vehicles?

These examples highlight the fact that test and evaluation is failing to identify critical failure modes and design flaws and vividly characterizes the rapidly escalating cost of failures like these. Fortunately, only one of the above examples resulted in loss of life. In several instances, those involved were nothing short of extremely lucky to walk away from, much less survive, the mishap. We *will not be* that lucky every single time, and it is very easy to imagine, for example, a tragedy involving hundreds of lives lost in a midair collision implied by the failure mentioned above.

I do not propose to understand the cause of these failures. Nor do I envision a silver bullet, a grand, unified theory of everything, which will heal all of our systemic ailments. Instead, I am suggesting a fundamental change to the way we use data to make decisions. Let me be clear that this is not intended to be a lesson in statistics, though I will introduce some new techniques in data science and uncertainty analysis that deserve our attention. Instead, the crux of this argument is that we must embrace a vision of incremental innovation based on modern decision making methods. These methods are incredibly powerful when leveraged by Analytical Tools of Mathematics and Statistics, or ATOMs (Jones 2010a), and it is the purpose of this article to demonstrate real benefits from elementary applications of five such principles.

Example: A simple flight test program

I will begin by describing a simple flight test program that we will use repeatedly to illustrate the principles and their benefits. Suppose we have an aircraft that flies continuous orbits, and we have modified the design of the aircraft to achieve better turn performance in the cruise configuration while executing this mission. Because of the high operational demand for this aircraft, we have very limited time, budget, and resources to validate the design changes. This constraint will, for the purposes of our example, limit us to the absolute minimum number of tests. (This is, in fact, a very realistic limitation.)

Suppose further that after an initial joint assessment by the design and test teams, the proposed flight test program identifies two factors—gross weight and Center of Gravity (CG)—and two levels of each factor. The test flights include executing standard turns at 30° angle of bank at a single airspeed to evaluate turn performance (turn rate). These additional assumptions will limit the scope of our discussion and help illustrate the five principles.

(The astute reader will quickly see that something is heinously wrong with this proposed flight test program. We will address this fact later. For now, we will pretend to be ignorant of the principles of aircraft mechanics.)

In the past, we would have performed a series of experiments. First we would conduct an experiment, a set of trials, to analyze the effect of gross weight. Then we would conduct another experiment to analyze the effect of CG. Based on our assumed constraints, we would have had only two trials in a given experiment to determine the effects of each factor, a trial at the low level of the factor and a trial at the high level of the factor.

Design and analysis of experiments

The mathematical model used in Design of Experiments (DOE) allows us to combine the trials, or aggregate the information in a mathematically rigorous way, thus increasing the number of repetitions and quantity of statistical information available. Now, we have four total trials to determine the effect of gross weight, two trials at each level, as seen in *Figure 1.* The same four trials also simultaneously provide us with the data to analyze two CG positions. In this case, the benefit of DOE was to aggregate the number of test points available for our statistical analysis and aggregate the quantity of information for a given set of trials.

Let us consider two of the cases of possible data outcomes. First, assume that turn performance changes at each combination of factors—turn performance at point A *(Figure 1)* is different than at point B, which is different than at point C, which is different than at point D. This is the worst case in the statistical sense, because there is only one data point for each of the four trials. Thus it is more difficult to quantify the noise present in the data and to distinguish it from the effect of the factors and levels at each of these points.

On the contrary, the best case in the statistical sense would be if there was no change in turn performance throughout the test. This leaves us with four data points to characterize the variance created only by statistical noise; i.e., things like measurement errors, atmospheric disturbances, etc. (It is important to emphasize this concept: even with today's data measurement and analysis capabilities, it is certain that the same maneuver measured four times or four different ways will yield different results, even though these differences might be very small or statistically insignificant.)

Design of experiments is not a novel idea, and it has received some support in the test and evaluation community. But it has not received widespread implementation, and so its benefits have been repeated here in our exploration of decision making methodology. Additionally, it is foundational in the sense that

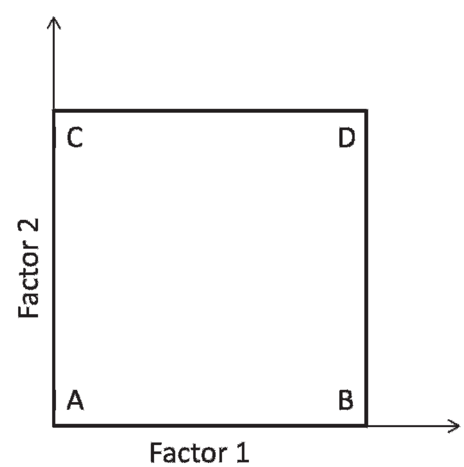


Figure 1. An example of a 2^2 experimental design.

subsequent principles build upon and dovetail nicely with it.

Decision theory

Decision theory is the second tool that we must include in our incremental innovation. It is a hybrid of economics, statistics, mathematics, and operations research. This discipline allows us to assign a cost function to decisions, and then use optimization algorithms to arrive at the best mathematical decision.

For example, we might assign a certain value to achieving a 95 percent confidence. This is a desired result, a benefit; and in decision theory, we define benefits to be the mathematical opposite of costs. If the value of this level of confidence was \$95, then the cost would be −\$95. We could assign an analogous value (cost) to each confidence level and then optimize our decision based on this cost function. The cost function would also include considerations like actual cost of flying hours, personnel, and fuel, as well as a quantitative assessment of the risk associated with drawing a statistical conclusion based on a certain amount of data. Additionally, we could combine the two, using decision theory in the above example with the DOE tool to decide how many trials we should run.

Here is a more concrete application of decision theory to our previous example. Assume that after three trials, initial data analysis suggests that the factors and levels have no effect on turn performance. Suppose also that the cost of each flight test is \$1 million and that any risk posed to the crew or the test assets is negligible. We have two possible decisions.

The first is to perform the fourth trial as planned. In this case, we might conclude after the final trial that factors and levels still have no statistically significant effect, and thus, \$1 million was wasted on the effort. Even if the data show that there is an effect, this outcome was highly improbable, and we might suspect its veracity anyway. In essence, we are evaluating whether a certain level of confidence in our data is worth the \$1 million price tag of the flight test. The alternative is not to perform the final trial. Again, we see that the cost function must balance the dollars required for executing the test and the value of the data quality.

This example is trivial, as we have simplified it immensely to facilitate the discussion, but it is easy to begin to see the wide application of this theory to the way we conduct test and evaluation. In fact, we already use these techniques in a qualitative sense, using our judgment to weigh things like cost and risk. The benefit here is added analytical rigor from assigning quantitative values to statistical properties. In other words, decision theory allows us to incorporate more of the programmatic and qualitative elements, like the value of data, in a quantitative way.

I would like to pause momentarily and rewind the discussion just a few paragraphs. We had conducted three trials in our experiment and noted that there was no statistical evidence suggesting the factors had an effect on turn performance. We were faced with the decision to continue or terminate testing based on the cost of the test, the value of the data, and the risk to the crew. (At this point, I would also like to remind the reader that earlier, I mentioned that we would revisit the heinous setup of this flight test program, and we do so now.)

Suppose that after completing three trials and before deciding about the fourth trial, the test team statistician meets a test pilot in the hallway, and the pilot mentions that turn performance is a direct function of bank angle and airspeed and that if they are held constant, gross weight and CG do not have an effect. This information totally changes the decision process. In fact, we can all agree that we should have considered this information before deciding on the setup of the test program. Let us define this to be *prior* information and show the importance of this classification.

Bayesian methods

Bayesian, or conditional, methods are part of a discipline that utilizes prior information relevant to statistical situations in a mathematically concise way. Loosely speaking, knowledge of aeronautical engineering and turn performance was the prior information that would have drastically changed the way our example flight test program was set up. Bayesian methods allow us to quantify prior information in a way that affects the nature of the data gathering and decision making process.

A key characteristic of Bayesian methodology is the incremental nature of the decision process. With classical statistical tools, we are required to decide the number of trials necessary before the experiment begins, and as noted for the DOE application this was four trials. With Bayesian methods, it is possible to update the prior information after each trial based on data gathered. This allows decision makers to see how the statistical evidence is developing as it happens.

One example illustrates the Bayesian method rather poignantly. Never in the history of flight test have engineers asked test pilots to repeat a flight maneuver for which the result was an unstable, divergent pilot induced oscillation (PIO) simply to attain the minimum sample size required for a statistical analysis. We only need a sample size of one to "prove" that a dangerous PIO exists. After the PIO has occurred, the test team subjectively updates its prior information to account for the actual occurrence of the PIO.

The length of this article prevents me from expounding on the technical details of Bayesian methods, but this is a very powerful mathematical tool that makes provisions for the incremental aggregation of information, a practice that we have adopted in almost every area of engineering except data analysis.

Time series methods

Recall from the example flight test program that the maneuver was turn performance at a fixed bank angle and speed. Traditionally the flight test technique for turn performance required more than a full turn in order to get 360° of stable data. In the event that no continuous circle was stable, engineers had the latitude to extract a portion of the maneuver and calculate the turn rate from this portion. In any case, turn rate was basically nothing more than an average.

But we are no longer limited like this, because modern instrumentation, inertial navigation units, and highly accurate time-space-position systems provide a digital stream of information at very high sample rates, and turn rate is a parameter that can be included in this data stream. This raises a very important question: How many observations of turn rate do we actually have? After a test program with four separate trials, do we have four observations of turn rate? Or, do we have four turns at 4 minutes each at 120 Hz sample rate (i.e., thousands of samples)?

Unfortunately, the answer is not clear. We do know that neither extreme is the correct answer, but time series methods can help us determine a solution. Time series is a statistical discipline based on the mathematics of stochastic processes. This tool confirms our intuition that, for turn performance, not every sample is independent. In other words, the turn rate at a given instant in time is related to the turn rate just prior to it and just after it—that makes sense intuitively. But there is no reason to expect that for a stable maneuver, some subsets of the points are, in fact, independent. It makes perfect sense that the turn rate at heading 360° (north) need not depend on the turn rate when headed due south.

Collaboration and connection: Digital and social aggregation of information

Aggregation of information has been a recurring theme in the four analytical techniques mentioned above. Aggregation is also a capability multiplied by the connectivity that we now experience. The price of digital sharing is virtually free, and the audience for digital content is growing exponentially, a phenomenon known popularly as "The Long Tail" (Jones 2010b). Additionally, cloud-based computation capabilities are improving drastically on a daily basis. For example, Google can provide us with billions of pieces of information almost instantly. This capability is rapidly expanding in the data sciences.

Social aggregation refers to the ability of a person to identify a piece of information that another person might need or find useful based on knowledge gained in a social relationship. For example, I have many friends in various sectors of the flight test industry. I might read an article tomorrow that describes a finding that one such colleague might find useful, and I could share that information easily with him. This idea is also a capability multiplied by the growth of connectivity. For example, computer algorithms based on people's preferences are an example of social aggregation, like Amazon's book recommendations based on what other people purchased. Networks multiply our capability as test professionals to share information as never before, but they also connect humans with humans, because no computer has yet replicated the brain's ability to make judgments and decisions and exercise intuition and wisdom.

There is also quantitative data that demonstrates the benefits of connection. Community is known to increase productivity, both physical and digital communities, as documented by innovation expert and author Scott Belsky (Jones 2011). Collaboration results in accountability and increased quality and increased readership (Cook 2010). And it stimulates innovation and reduces cost and timelines. We need look no further than Wikipedia to see the effectiveness of collaboration.

Based on this trend of digital and social connectivity, we have every reason to believe that the ability to aggregate seemingly unrelated pieces of information will increase in the near future, both as a result of social interactions and as a result of the connectedness and computation ability that is constantly evolving. For example, a pilot who notices an atmospheric disturbance while gathering flying qualities data in an F-16 might call his friend, the C-17 pilot, and offer him an anecdotal weather report. An hour later, as the C-17 takes off to collect performance data, its algorithms have been adjusted based on the F-16's test sortie data, which were processed by an atmospheric researcher on the other side of the world in near real time and subsequently downloaded from the cloud to the aircraft's inertial units.

Conclusions and recommendations

It is impossible to believe that each one of us will become experts in advanced statistical methodology. Nor will the statisticians among us instantly become senior program managers. I do not believe for a moment that we need either of these things *today*. There are, however, definite steps and incremental changes, that we must take *today* as a test community to improve our decision making process—five have been described here (four ATOMs and one technological capability).

The first and foremost recommendation is to share knowledge, experience, and wisdom and to share content with more people. Let me give an example. Google uses an algorithm to recommend similar content for syndicated feeds aggregated in Google Reader™, presumably similar to Amazon's book recommendations. It cannot perform this operation when there are no syndicated test and evaluation feeds and when there are no recommendations from digital readers. So "share" means that we, the T&E community, must adopt the technology of Web 2.0.

Additionally, each one of us can increase the number of our connections. We can create or join groups on popular social networks, based on our affinities and memberships in associations like ITEA. We must increase connections and collaboration between ITEA members and among sister organizations. Read a blog and make comments. Subscribe to an RSS feed, and "like" and "share" valuable content found, because the line between social and professional are blurring as fast as I type these letters.

If these steps do nothing else, they will serve to connect us with the next generation, with those who will develop the fantastic algorithms that we imagine, with those whom we can impress with the importance of the analytical tools of mathematics and statistics, and with those who will encourage us daily to inform ourselves about the essential nature of these tools and help us begin to implement them. ❑

MARK H. JONES JR. is finishing a career as the chief C-17 test pilot at Edwards Air Force Base. He is the founder of mc² (Mathematical Christian Consulting at www.multiplyleadership.com). Mark serves the Society of Experimental Test Pilots as the LinkedIn social network manager, and he publishes aerospace-social media-innovation commentary at www.planeconversations.com. He holds a master of science degree in flight test engineering from U.S. Air Force Test Pilot School and successfully defended his thesis "Minimax and Statistical Decisions in the Tactical Arrival Problem" for the master of science degree in mathematics from the University of Charleston, South Carolina. E-mail: mark@multiplyleadership.com

References

Associated Press. 2009. F-22 crashes near air force base. *Fox News,* March 25. http://www.foxnews.com/story/0,2933,510588,00.html (accessed January 11, 2011).

Cohen, Aubrey. 2010. WSJ: Collision warning devices on 9,000 planes are faulty. *Seattle PI,* December 28. http://blog.seattlepi.com/aerospace/archives/233458.asp (accessed January 12, 2011).

Cook, John. 2010. Write-only articles. *The Endeavor,* June 23. http://www.johndcook.com/blog/2010/06/23/write-only-articles (accessed January 21, 2011).

Jones, Jr., Mark. 2010a. ATOMs: multiply leadership, energize the team. *Multiply Leadership*. January 31. http://www.multiplyleadership.com/2010/01 (accessed on January 12, 2011).

Jones, Jr., Mark. 2010b. The long tail. *Multiply Leadership*. November 12. http://www.multiplyleadership.com/thelongtail (accessed January 21, 2011).

Jones, Jr., Mark. 2011. Community multiplies innovation. *Multiply Leadership*. January 10. http://www.multiplyleadership.com/community-multiplies-innovation (accessed January 21, 2011).

Mustoe, Howard. 2010. Rolls-Royce test-bed blowout shuts site used for Boeing, Airbus. *Bloomberg*. August 24. http://www.bloomberg.com/news/2010-08-23/rolls-royce-engine-blowout-on-test-bed-shuts-site-used-for-boeing-airbus.html (accessed January 12, 2011).

Niles, Russ. 2009. Second Skycatcher prototype crashes. *AVWeb*. March 19. http://www.avweb.com/avwebflash/news/SecondSkycatcherPrototypeCrashes_199991-1.html (accessed January 11, 2011).

Peterson, Barbara. 2010. A380 engine failure—Worse than originally thought. *Popular Mechanics*. November 8. http://www.popularmechanics.com/technology/aviation/safety/a380-rolls-royce-qantas-crash (accessed January 12, 2011).

ITEA Journal 2011; 32: 206–209

A Measure for the Decline of Short-Term Working Memory Capacity With Coincidental Sleep Deprivation

David Bate
Scientific Research Corporation, Charleston, South Carolina

Working memory is an important feature of the human mind's capacity to hold temporary information that must be utilized in decision making and information valuation. While multiple models for working memory exist in accepted literature, we utilize the well-known Baddeley model of working memory, which contains the short-term memory components known as the phonetic loop, the visuospatial sketchpad, the episodic buffer, and the central executive. Numerous studies have shown that there is a deleterious effect on cognitive performance as a result of sleep deprivation. A major goal of this review is to link a variety of experiments that have shown working memory capacity to have the generally accepted limit of 7 +/− 2 and to explore further work that shows sleep deprivation to carry an apparent negative effect on working memory capacity. This review attempts to derive an empirical metric or model that describes quantitatively the degradation of working memory capacity as a result of sleep deprivation. The connection between working memory and attention is also briefly investigated, since both presently appear to be based in the frontal lobe.

Key words: Attention; battlespace encoding; collaborative situational understanding; command and control; focus; memory consolidation; physiological testing in the field; warfighter.

The goal of this review is to promote evidence that electroencephalography (EEG) and functional magnetic resonance imaging (fMRI) data are relevant to monitoring prefrontal cortex (PFC) activity and that they (the EEG and fMRI) provide significant indicators of working memory activity as well as the related attentional activities. During preparation for the review, a variety of methods for evaluating sleep deprivation were covered, but few utilized the current technical capabilities available to cognitive research. An objective metric will be proposed that can be used for evaluating cognitive capabilities from the physiological data utilizing relational measurements of neural activity to assess changed capabilities. The review of working memory capacity, as it is meaningfully applied in the command and control (C2) context, further implies that this metric could offer additional validity to other cognitive tests through improved predictive and diagnostic abilities for the evaluation of cognitive state during a variety of tests involving working memory. An additional affecter of working memory capacity, skilled memory, could provide more valuable results to improve working memory performance for C2 applications. Skilled working memory is based in a natural encoding method that facilitates virtual increases in working memory capacity. The following section reviews a selection from the extensive available literature in both the areas of sleep deprivation and working memory capacity. Applications of that work as well as possible extensions are then presented for the C2 domain in the final section.

Previous work

Cognitive activity monitoring

Measuring and evaluating cognitive activity is a rapidly advancing field in which technology has recently provided new opportunities to enhance experimental test performance. Electroencephalography and fMRI of brain activity provide objective, measured values of neurocognitive state that are not otherwise available. The alternative method of evaluating neurocognitive arousal has been through the use of subjective questionnaires; however, they are susceptible to a variety of spoiling effects, most notably altered performance due to reactivity effects toward the perceived goal of the experiment and task interruption. The prefrontal cortex is largely accepted to be responsible for executive control of attention and working memory. Gevins and Smith (2003) studied changes in cortical activity related to sleep disorders in working memory tasks using electroencephalograms; while Bor et al. (2003) reported PFC activity, using fMRI to confirm increased activity during working memory path tasks. The use of EEG sensors to monitor soldiers in real-time military operations on urban terrain (MOUT) training was tested by Dorneich et al. (2007), where they found that despite noise factors such as muscle movement, the accuracy obtained from the study matched the accuracy of a pristine lab environment. Dorneich's study illustrates that the expense and portability of physiological monitoring systems are no longer preventative for technical cognitive studies in active scenarios, where previously those constraints would have mandated the experimental use of subjective questionnaires.

Limitations in short-term working memory

Working memory is the cognitive system used for temporarily storing new sensory input and managing the information to be used in complex cognitive tasks. Working memory is involved in the selection, initiation, and termination of information-processing functions such as encoding, storing, and retrieving data into more permanent long-term memory. George Miller brought together the concept of capacity-limited working memory used for short-term memory storage and encoding of individual elements into "chunks" (Miller 1956). A chunk can be defined as a group of elements combined in some meaningful way to be recalled as a single element. Miller proposed that the number of chunks that could be stored in working memory was limited to "the magical number seven, plus or minus two." Baddeley and Hitch (1974) later proposed a formal model for short-term working memory that consisted of a central executive component that controlled and manipulated two slaved storage systems: the phonological loop and the visuospatial sketchpad. Each of the slaved storage systems was thought to be limited by the accepted 7 +/− 2 chunks. A model proposed by Cowan (1995) for recall of long-term memory into short-term working memory emphasized working memory as a component of highly activated long-term memory as opposed to a separate memory system. In Cowan's model, long-term memory currently in the "focus of attention" was highly activated and therefore more rapidly accessible than the relatively inactive remainder of long-term memory. The focus of attention was described as having a limited capacity of approximately four chunks. In support of these conceptual theories, Edin et al. (2009) have used fMRI to experimentally show a physiological limit to the excitation and inhibition of neurons during working memory tasks. Bick and Rabinovitch (2009) further developed this physical discovery with a mathematical model that confirms the findings.

Sleep deprivation

Both total and chronic partial sleep deprivation (SD) are well defended in existing literature to have a negative effect on attention and working memory. Alhola and Polo-Kantola (2007) conducted an extensive review of the effects of SD and reported that SD also affects other functions, such as long-term memory and decision making. Tolerating SD can also depend on age, gender, and individual traits. Several reasons such as physiological mechanisms as well as social or environmental factors may also be involved. They concluded that many of the effects of SD were difficult to determine conclusively because of the great variation in SD studies.

Harrison and Horne have performed multiple studies showing the immediate detrimental effects of SD (Harrison, Horne, and Rothwell 2000). They found experimentally that SD preferentially reduced the functionality of the PFC in subjects of all ages. Particularly noted in one study is a decline in decision-making capability, though the specific type of PFC functionality that was diminishing was not theorized (Harrison and Horne 1999). Mograss et al. (2009), in a study of the recall of unfamiliar faces, found no evidence for decreased response times after SD when recalling previously stored information, but significantly slower identification and dismissal times when discriminating old faces from new faces. They conclude that SD does not affect access or retrieval, but rather discrimination of what is contained in memory. They posit this is due to less processing of details during storage or retrieval.

Along with the deleterious effects of SD above, Ellenbogen (2005) studied the additional loss of the benefits normally gained from sleep. Medical residents were observed to suffer apparent cognitive performance lapses even when they were administered pharmaceutical drugs to maintain a consistent level of cognitive arousal. The result indicated that along with detriments for SD, a resident missed out on benefits that were gained from regular sleep. The benefits of keeping a healthy sleep schedule included cognitive processing such as memory consolidation and insight formation on the period preceding regular sleep. Proper memory consolidation is valuable as it affects long-term memory retention and the ability to rapidly access those long-term memories.

Skilled memory

While experimenting with the theoretical bounds of working memory, Ericsson, Chase, and Faloon (1980) reported that an athletic student had been trained from a baseline of the expected seven-element recall of random numbers to a consistent recall of more than 70 elements. They coined the term "skilled memory" for the trained increase in working memory. Through interviews with the student it was determined that his short-term working memory had not grown, but instead the student had developed a method of encoding newly presented information into keys for long-term memories already reliably stored. By creating a hierarchy of these keys, the student was able to continually reduce his working memory burden to a small number of elements. This skilled memory encoding was assumed to be a subset of "chunking" that utilized specific organized, efficient methods. While Ericsson and Chase's initial study was in digit recall, skilled memory theories have been tested extensively in the domain of chess owing to the large volume of cognitive studies that have been performed in that domain and the abundance of clearly identified experts.

Building upon work by Ericsson, Gobet and Chassy (2009) showed through chess position experiments that among novices and experts studied, experts stored and retrieved working memory following some more complex method than simple chunking. In fact, when experts recalled chess positions they often seemed to intuitively recall large collections of pieces (elements) as a single chunk. Gobet called the method "templates" and presented a specific model to accurately predict the learning speed and retention of working memory by chess experts. Templates were flexible (often large) long-term memory chunks that experts utilize or modify on the fly. In chess, an expert may recognize a pattern in long-term memory of six pieces and store only a reference to that known pattern and some modification to the pattern. Gobet explained that the student in Ericsson's studies used his extensive long-term knowledge of track running numbers as templates, so "while [the student] drastically improved in less than 100 hours of practice, they drew upon their massive knowledge of running times, dates, etc., which took years to develop" (Gobet 2000).

Application to C2

Deprivation metric

Sleep deprivation can be generally noted from experimentation to reduce cognitive performance, but it seems to inhibit individuals in a variety of ways, and individuals have been seen to react differently to varied levels of sleep loss. In seeking the proposed metric for sleep deprivation effects on working memory, individual expertise and cognitive abilities have shown to be too complex to simplify into a single metric of the number of elements reduced from working memory capacity. The proposed metric for sleep deprivation effects on working memory is therefore to be generated instead by establishing a baseline of PFC activation where the attentional and working memory functions are believed to reside. The baseline activation level as determined with EEG sensors or fMRI should be established utilizing an appropriate standard test from the experimental domain. From this baseline, a relative scale ranging from −10 to +10 may then be utilized to objectively quantify performance as a metric describing the relational percentage it has to the baseline value. *Table 1* provides example metric values for various hypothetical cognitive situations. This metric, when measured in conjunction with other cognitive tests and evaluations could improve the predictive and diagnostic abilities by accounting for a key variable in performance.

Battlespace encoding

Collaborative situational understanding, which is the creation of a common and relevant operational picture, has been identified as a critical component of situational awareness for the modern warfighter. With the availability of this component, an individual is made aware of their combined situation with other individuals and is able to provide maximum focus and resources to the appropriate effort (Wesensten, Belenky, and Balkin 2005). In order to maintain the cognitive capabilities to manage collaborative situational awareness at the individual level, trained templates for rapid encoding would allow for more situational information to be considered and more time for decision making. Additional support for this concept of trained battlespace encoding with templates can be derived from Bor et al. (2003), who reported improved accuracy (performance) when data were presented in a prestructured format during an fMRI study of working memory path tasks.

Table 1. Example situations with activation levels in the area of interest and related metric value.

Situation	Activation level	Metric
Baseline	62%	0
In reverie	37%	−4
24 hours sleep deprived with caffeine	56%	−1
Immediately following attention focusing exercises	76%	+2.3

It can be expected that proper template creation under healthful sleep conditions will lead to improved performance in sleep-deprived or other harsh attentional environments. This expectation is derived from Mograss's conclusion that SD affects processing rather than access or retrieval of memory combined with Ellenbogen's contribution of memory consolidation as a benefit of sleep. In a sleep-deprived scenario, trained subjects already hold the majority of the template structure in "safe" long-term memory and are thus responsible for encoding fewer elements into short-term memory. The existing template provides helpful focusing hints to critical data when overloaded as well.

In a hypothetical training model, a variety of mission templates could be created with expected units and positions. These templates would then be available to the warfighter as references to previously created long-term memories that would help recognize key features in a situation and could be adapted rapidly to store their current operational situations. As proficiency is displayed, the number of units and positions presented to the warfighter should be increased, forcing the creation of more capable templates. The goal of this training would be to encourage the development of larger templates through rote training. As encoding seems most effectively performed based on relational encoding to familiar long-term memories, a unit presentation method that eased encoding into a familiar domain should be experimented with, as a successful method could increase the amount of information that is chunked into template slots. ❑

DAVID BATE received a bachelor of arts degree in history from Erskine College and a master of science degree in computer science from Clemson University. He is currently a research engineer at Scientific Research Corporation and a member of the International Council on Systems Engineering. E-mail: dbate@scires.com

References

Alhola, P., and P. Polo-Kantola. 2007. Sleep deprivation: Impact on cognitive performance. *Neuropsychiatric Disease and Treatment* 3 (5), 553–567.

Baddeley, A. D., and G. J. L. Hitch. 1974. Working memory. In *The Psychology of Learning and Motivation: Advances in Research and Theory*, ed. G. A. Bower, vol. 8, 47–89. New York: Academic Press.

Bick, C., and M. I. Rabinovich. 2009. Dynamical origin of the effective storage capacity in the brain's working memory. *Physical Review Letters* 103 (21), 218101.

Bor, D., J. Duncan, R. J. Wiseman, and A. M. Owen. 2003. Encoding strategies dissociate prefrontal activity from working memory demand. *Neuron* 37: 361–336.

Cowan, N. 1995. *Attention and memory: An integrated framework*. New York: Oxford University Press.

Dorneich, M. C., S. Mathan, P. M. Ververs, and S. D. Whitlow. 2007. An evaluation of real-time cognitive state classification in a harsh operational environment. *Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting*, October 1–5, Baltimore, Maryland, 146–150. Santa Monica, CA: HFES.

Edin, F., T. Klingberg, P. Johansson, F. McNab, J. Tegnér, and A. Compte. 2009. Mechanism for top-down control of working memory capacity. *Proceedings of the National Academy of Sciences of the United States of America* 106 (16), 6802–6807.

Ellenbogen, J. M. 2005. Cognitive benefits of sleep and their loss due to sleep deprivation. *Neurology* 64: E25–E27.

Ericsson, K. A., W. G. Chase, and S. Faloon. 1980. Acquisition of a memory skill. *Science* 208: 1181–1182.

Gevins, A., and M. E. Smith. 2003. Neurophysiological measures of cognitive workload during human-computer interaction. *Theoretical Issues in Ergonomics Science* 4 (1), 113–131.

Gobet, F. 2000. Retrieval structures and schemata: A brief reply to Ericsson and Kintsch. *British Journal of Psychology* 91: 591–594.

Gobet, F., and P. Chassy. 2009. Expertise and intuition: A tale of three theories. *Minds and Machines* 19 (2), 151–180.

Harrison, Y., and J. A. Horne. 1999. One night of sleep loss impairs innovative thinking and flexible decision making. *Organizational Behavior and Human Decision Processes* 78 (2), 128–145.

Harrison, Y., J. A. Horne, and A. Rothwell. 2000. Prefrontal neuropsychological effects of sleep deprivation in young adults—A model for healthy aging? *Sleep* 23 (8), 1067–1073.

Miller, G. A. 1956. The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review* 63: 81–97.

Mograss, M. A., F. Guillem, V. Brazzini-Poisson, and R. Godbout. 2009. The effects of total sleep deprivation on recognition memory processes: A study of event-related potential. *Neurobiology of Learning and Memory* 91 (4), 343–352.

Wesensten, N. J., G. Belenky, and T. J. Balkin. 2005. Cognitive readiness in network centric operations. *Parameters: U.S. Army War College Quarterly* 35 (1), 94–105.

ITEA Journal 2011; 32: 210–214

NASA Task Load Index for Human-Robot Interaction Workload Measurement

Joshua A. Gomer, Ph.D.
Human Factor Engineering, LLC, Mt. Pleasant, South Carolina
Christopher C. Pagano, Ph.D.
Clemson University, Clemson, South Carolina

The objective of this work is to demonstrate the utility of the National Aeronautics and Space Administration–Task Load Index (NASA-TLX) as a metric for collecting workload data in human systems technologies. Past work has demonstrated that the NASA-TLX is a useful measure for assessing subjective workload data with robot operators. In the current experiment, participants completed a battery of spatial ability measures. They then completed four test conditions under which they operated a robot through low- and high-difficulty courses, under direct line of sight and tele-operation. Prior to and after each of these four conditions, participants completed a computerized version of the NASA-TLX. Performance was determined by examining course completion time, the total number of collisions made during operation, and, in some instances, the accumulation of points achieved by stacking cylinders with a gripper arm. Analysis of the NASA-TLX data confirmed a difference in subjective workload between high- and low-difficulty course conditions. In addition, higher aggregate spatial ability was correlated with superior operator performance under each of the experimental conditions. The NASA-TLX is a valuable tool that can be used in many research environments to assess subjective workload data. Furthermore, better understanding of the relationship between spatial abilities and robot operation can assist in the selection and training of future operators, as well as in the design of superior human system technologies.

Key words: Human systems technology; NASA Task Load Index; robot operators; spatial ability; workload.

Human-centered test and evaluation (T&E) is quickly becoming the standard for modern industry and defense development. Whether incorporated (ideally) from product inception or used to integrate multiple technologies into a complex system, the value of focusing on the interaction between the user and the product is becoming more recognized. Human factors (HF) research domains including driving, aviation, Web development, automated systems, and robotics have all benefitted from user testing, leading to more usable products. As the usability of a product increases, the potential functionality, success, and marketability of that product also increases.

Robots and their operators have received much attention in the last decade, especially in areas of national defense and urban search and rescue (USAR). Further, within the specific domain of robotics, HF has been playing an increasing role in design. Research has focused on a better understanding of human-robot interaction (HRI). The present study resulted in the successful use of the National Aeronautics and Space Administration–Task Load Index (NASA-TLX) (Hart and Staveland 1988) for assessing subjective workload in an HRI environment and provides support for the use of this measure in other human system technologies.

Human factors professionals have many tools that can be used as metrics to analyze the performance of a human–machine system. Likert scales, time to complete a task, accuracy, eye movements, and questionnaires are just a few examples. The NASA-TLX is a specific tool that can be used to determine the subjective workload of a user after they have completed a task. This index assesses perceived workload by the user's analysis of weights and ratings of six different dimensions: mental demand, physical demand, temporal demand, performance, effort, and frustration. Particularly effective when coupled with other objective workload measures, the NASA-TLX requires 1–5 minutes of user time, is easy to administer, and can be used as a primary or secondary measure. When not being used to determine a primary dependent variable, the NASA-TLX may also be used to determine the strength of the experimental design. For example, this tool can be used to determine if manipulated levels of difficulty are perceived appropriately by the participants. In the current study, this measure was successfully implemented as a secondary tool used to determine the success of varying high and low levels of task difficulty.

Past research has revealed that higher spatial abilities (SpA) of operators correlate with superior dependent measures of performance (Lathan and Tracey 2002; Bowen 2004; Long et al. 2009; Gomer 2010). The construct of SpA is composed of two major subsections: visualization and orientation. Spatial visualization is a measure of the power necessary to solve complex mental problems, while spatial orientation is associated with the speed involved in solving mental calculations (Pelligrino, Alderton, and Shute 1984).

The present study investigated the relationship between SpA and performance measured with the dependent variables of time to complete tasks, the number of collisions made within a remote environment, and in some instances the accumulation of points achieved by stacking cylinders with a gripper. Both studies were completed to examine the relationships of SpA and performance using different types of robot platforms. Participants were required to complete tasks as quickly as possible while making as few collisions as possible under both viewing conditions of direct-line-of-sight (DLS) and tele-operation (TO).

Method

Participants

One hundred eighteen Clemson University students and employees (55 males, 63 females) participated in the study after providing informed consent. Each participant tested at a high contrast visual acuity of at least 20/25 binocularly from 6 meters and reported full use of their neck, arms, and hands.

Materials

Visual SpA measures

In order to analyze correlations of performance and SpA, two SpA measures were administered to each of the participants. Together, these measures took approximately 15 minutes and were counterbalanced by testing prior to or following experimental conditions.

Spatial visualization was assessed using the Paper Folding Test (VZ-2), which has demonstrated a test–retest reliability estimate of 0.84 (Ekstrom et al. 1976). This measure comprises two 10-item pages. Each item presents an image of a piece of paper, and two to four additional images demonstrating a sequence of the piece of paper being folded. On the final folded image, a circle indicates a hole punched through the paper. The object of the task is to select one of five correct image representations of the unfolded piece of hole-punched paper.

Spatial orientation ability was measured using the Cube Comparison Test (S-2), which has also demonstrated a test–retest reliability estimate of 0.84 (Ekstrom et al. 1976). This inventory consists of 42 items. Each item presents two cubes, each with three sides in view. Based on the limited images presented, the participant must decide whether the two cubes are the same or different.

Tasks and measures

The robots used in this study included off-the-shelf remote-controlled entertainment models of a tracked vehicle, a wheeled vehicle, a wheeled vehicle with a robotic arm gripper, and a stationary arm with a gripper. These robots were chosen because of their affordability and their design similarities with larger and more expensive industry, defense, and USAR models.

The camera used for the remote viewing condition was a Grandtec USA EyeSeeAll wireless camera with a field of view of 90° that transmitted live video feed to a computer monitor. There was no perceptible delay in transmitting video, and the angle of the camera mounted atop the robot was fixed to allow visibility of the front of the robot during operation. The camera feed was displayed in a 19.1-cm^2 window on a 38.1-cm LCD monitor via a radio-frequency complementary metal-oxide semi-conductor (RF-CMOS)-USB Tx transmitter and Rx receiver.

Participants operated the robot under both DLS and TO conditions. Similar to the operational definition created by Sheridan (1992), the TO conditions of the current study refer to operating the robot while viewing a monitor that is fed by a camera affixed to the robot. Under the DLS conditions, participants operated the vehicle while facing the robot and the courses. Prior to the TO conditions, participants viewed a training video, which demonstrated what this task would look like. Under the TO conditions, participants controlled the robot through the same courses while viewing the video feed on a computer monitor. The wheeled and tracked vehicle courses involved navigating through a maze with obstacles. Difficulty for these courses was increased by adding obstacles into the maze and reducing the amount of space for navigation. The stationary arm task involved picking up cylinders from the left hand side of the robot and stacking them on the right hand side of the robot. The wheeled vehicle with gripper task involved picking up cylinders from one location in the course and stacking them in a different location in the course. Difficulty for the stacking tasks was increased by going from a single level of cylinders to stacks of two cylinders. Following an initial training session and after each of the four conditions, the NASA-TLX was administered.

Design

The experiment used a within-subjects design for factors of environment (DLS vs. TO) and difficulty (low vs. high). The study also had a between-subjects design for platform (stationary arm with gripper vs. wheeled vehicle with gripper vs. wheeled vehicle vs. tracked vehicle). Prior to the experimental conditions, there was a brief session for training participants with the robot controls. SpA measures were counterbalanced, as were the four experimental conditions to minimize practice effects.

Procedure

Prior to the experiment, participants were given a demonstration by the experimenter on how to control the robot. They were then allowed 1 minute to practice and were tested until they demonstrated proper machine operation.

During the experiment, participants operated the vehicle under both a low and high level of course difficulty for each viewing condition. Participants had either a navigation task or a task that involved stacking cylinders. For the navigation task, they were instructed to navigate the robot from its starting location to the furthest location in the course as quickly as possible, trying not to hit anything. Once they reached the end of the course, timing was stopped, and the robot was rotated 180° for returning to the starting point. When participants were ready, and any fallen or moved obstacles had been reset, timing started again and they navigated the robot back to the start. For the cylinder tasks, they were instructed to recreate structures with their grippers. In other words, if three stacks of cylinders were to the left of the robot, participants were instructed to recreate those three stacks of cylinders to the right of the robot. Participants were informed that an experimenter would be recording the time to complete the task as well as the number of collisions. Collisions were documented by three levels of severity and a deduction of one, two, or three points: a low level (touching a wall or object without modifying the course), a medium level (modifying the course by moving objects or walls), and a high level (becoming stuck and requiring experimenter assistance).

Results

Data reduction

The number and severity of collisions were reported as a single overall composite score. The composite score was created by assigning values to the different levels of collisions and adding them together. For the cylinder task, points were based on stacking cylinders in holes (1 point per cylinder placed appropriately) and alignment of two-cylinder stacks (1 point for each cylinder placed appropriately and 2 points per cylinder for stacks within 1 mm of being perfectly aligned). Scores on both SpA measures were calculated according to the method recommended by Ekstrom et al. (1976) and were also standardized into *Z-scores* for analysis.

Performance and workload

Higher performance was defined as faster course completion time, fewer collisions, and higher cylinder staking point totals. Spatial abilities were analyzed aggregately and independently. An aggregate visual SpA score was created for each participant by averaging the standardized scores of the VZ-2 and S-2 tests.

Replicating the findings of previous research (Lathan and Tracey 2002), SpA correlated with superior operator performance across all dependent variables.

Results of NASA-TLX data revealed that the manipulation checks of low versus high difficulty were successful. In other words, for all robots, participants perceived the high difficulty conditions as requiring more workload than the low difficulty conditions.

All comparisons were significant for DLS conditions $t(117) = -10.32$, $P < .001$, and TO conditions $t(117) = -8.53$, $P < .001$, confirming that low difficulty differed from high difficulty for all participants. Paired-samples t tests were also computed and these confirmed that participants within each platform perceived similar differences between low and high difficulty experimental conditions.

Discussion

In each of these two studies, the NASA-TLX was used as a quick and easy-to-administer computer-based tool prior to and after each experimental condition to determine if participants perceived the appropriate levels of workload. Considering a combined test administration time of approximately 15 minutes, results presented here illustrate the potential benefit of screening tele-operators for these two types of SpA. Implications of these studies further illustrate the potential benefit of utilizing the NASA-TLX for subjective human workload measurement in T&E applications and human system technologies.

Both studies revealed that visualization and orientation SpAs were significant predictors of robot operation performance (Chen and Terrence 2008; Sekmen et al. 2003; Lathan and Tracey 2002). However, there is a debate over the merits of ability versus experience, as certain professions currently screen for SpA (dentistry), while others rely on experience (medicine) as a predictor of performance. Robot operators with higher SpA scores in these studies performed better than those with lower scores under both DLS and TO conditions. Evidence of this relationship presented itself with faster course completion times, fewer collisions, and higher point totals.

Looking forward, a dissertation by Bowen (2004), demonstrated that individual differences in spatial visualization can affect the type of information utilized from different types of computer interfaces. Previous research has also focused specifically on strategies, such as reducing the number of mental rotations involved in interpreting tele-operation interfaces, thereby making navigation less mentally taxing (DeJong, Colgate, and Peshkin 2004). Additionally, there is evidence that training can improve spatial orientation performance (Jaeggi et al. 2008; Kass, Ahlers, and Dugger 1998). However, it is unknown whether such training effects could extend to robot operation or other human–machine systems. Future studies should further investigate such effects, as well as robot-operator selection, and should develop superior training strategies to compensate for different levels of ability. A better understanding of the individual differences of SpAs and this relationship to performance may play a large role in training and selection and can improve the design of appropriate interfaces for modern systems. In line with this research and other HF investigations, the NASA-TLX continues be a valuable tool for assessing subjective workload data in complex human system technologies. ❑

Dr. Joshua A. Gomer is originally from the Orlando, Florida, area. He worked as a Morse interceptor/analyst and reporter for the National Security Agency at Kunia, Hawaii, before beginning his undergraduate degree in psychology at the University of Central Florida. Continuing his education, some of his graduate studies at Clemson have investigated the role of spatial abilities in tele-operation, uncoupled motion, and physical functioning in older adults. Dr. Gomer holds a doctor of philosophy degree in human factors psychology and is currently investigating human systems automation/T&E as the vice president of Human Factor Engineering, LLC. E-mail: josh_gomer@hfellc.net

Dr. Christopher C. Pagano earned his doctor of philosophy degree in experimental psychology with an emphasis in ecological psychology from the University of Connecticut in 1993. His dissertation dealt with proprioception and the kinesthetic perception of limb orientation. During his graduate studies he also studied the haptic perception of hand-held objects. Dr. Pagano served as a postdoctoral fellow in the cognitive science program at Indiana University from 1993 to 1995, where he studied visual depth perception. Since 1995 Dr. Pagano has served as a professor in the Psychology Department at Clemson University, and he is an associate of Clemson's Human Factors and Ergonomics Research Institute. His current work focuses on the augmentation of depth perception in artificial displays and the design of interfaces for the tele-operation of robotic systems and for minimally invasive surgery. E-mail: cpagano@clemson.edu

References

Bowen, S. 2004. The Effects of Spatial Ability on Performance with Ecological Interfaces: Mental Models and Knowledge-based Behaviors. *Doctoral Dissertation*, University of Central Florida, Orlando, FL.

Chen, J. Y. C., and P. I. Terrence. 2008. Effects of tactile cueing on concurrent performance of military and robotics tasks in a simulated multitasking environment. *Ergonomics* 51 (8): 1137–1152.

DeJong, B. P., J. E. Colgate, and M. A. Peshkin. 2004. Improving teleoperation: Reducing mental rotations and translations. In *Proceedings of the 2004 IEEE Conference on Robotics and Automation*, April, New Orleans, LA, 3708–3714. Piscataway, NJ: IEEE.

Ekstrom, R. B., J. W. French, H. H. Harman, and D. Dermen. 1976. *Manual for Kit of Factor-Referenced Cognitive Tests. Letter Sets*. Princeton, NJ: Educational Testing Service.

Gomer, J. A. May 2010. *Spatial perception and robot operation: The relationship between visual spatial ability and performance under direct line of sight and teleoperation. Doctoral dissertation*. Clemson, SC: Clemson University.

Hart, S. G., and L. E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In *Human mental workload*, ed. P. A. Hancock, and N. Meshkati, 139–183, Amsterdam: The Netherlands.

Jaeggi, S. M., M. Buschkuehl, J. Jonides, and W. J. Perrig. 2008. Improving fluid intelligence with training on working memory. In *Proceedings of the National Academy of Sciences of the United States of America*, 105, 6829–6833, Washington, D.C.: National Academies Press.

Kass, S. J., R. H. Ahlers, and M. Dugger. 1998. Eliminating gender differences through practice in an applied visual spatial task. *Human Performance* 11 (4): 337–349.

Lathan, C. E., and M. Tracey. 2002. The effects of operator spatial perception and sensory feedback on human-robot teleoperation performance. *Presence* 11 (4): 368–377.

Long, L., J. Gomer, K. Moore, and C. Pagano. 2009. Investigating the relationship between visual spatial abilities and robot operation during direct line of sight and teleoperation. Poster presented at the 53rd Annual Human Factors and Ergonomics Society (HFES) Conference, October 19–23, 2009. San Antonio, Texas.

Pelligrino, J. W., D. L. Alderton, and V. J. Shute. 1984. Understanding spatial ability. *Educational Psychologist* 19 (3): 239–253.

Sekmen, A. S., M. Wilkes, S. R. Goldman, and S. Zein-Sabatto. 2003. Exploring importance of location and prior knowledge of environment on mobile robot control. *International Journal Human-Computer Studies* 58: 5–20.

Sheridan, T. B. 1992. *Telerobotics, automation and human supervisory control*. Cambridge, MA: MIT Press.

Acknowledgments

The authors would like to thank Lindsay Long for her contribution to the design and collection of data for this project. In addition, we would like to thank Kerry Gretchen and Kristin Moore for their editing advice.

Chapter News

George Washington Chapter

January Meeting The George Washington Chapter held its monthly luncheon January 20th at the Army Navy Country Club in Arlington, VA. The speaker was **Dr. John B. Foulkes**, Director of the Test Resource Management Center (TRMC) who discussed Initiatives and activities of the Center.

Foulkes said the current ten-year Strategic Plan was signed December 1st, bringing it in line with the Program Objective Memorandum (POM) cycle. It expands planning for cyber capability, including the Joint Interoperability Operations Range, the joint environment for

John Foulkes *receives memento from chapter president Mike Wetzl*

biometric identification testing, multi-service implementation of C-band , artificial intelligence T&E structure, and handling of encroachment issues in the frequency spectrum and renewable energy.

He went on to tell of current significant efforts. One is the Army execution order for combined/integrated testing of network programs of record which will begin testing in March. Another is the Air Force Space Command Launch Enterprise transformational modernization. “Others are the Navy Distributed Engineering Plant (an RDT&E Enterprise solution) and the Defense Information Systems Agency Net-ready Key Performance Parameter certification methodology”.

Foulkes Described Live-Virtual-Construction modeling activities including integrated Developmental/Operational testing for the Battlefield Communications Node, which is an urgent need for the warfighter, as well as proof of concept for cyber operations on the Joint Mission T&E and execution using the Joint Mission Environment Test Capability (JMETC) and the InterTEC tool, and the Joint Close Air Support

Chapter Locations

NORTHEAST REGION
Vacant, Vice President

CONNECTICUT & RHODE ISLAND
Narragansett Bay Chapter
Vacant

MASSACHUSETTS
New England Chapter
Michael E. Keller, Sr., President
Boston, MA

NEW JERSEY
South Jersey Chapter
John Frederick, President
Atlantic City, NJ

EAST REGION
Robert A. Vargo, Vice President

EUROPE
European Chapter
Steve Lyons, President
United Kingdom

ISRAEL
Israeli Chapter
Aaron Leshem, President
Haifa, Israel

OHIO
Miami Valley Chapter
Stephen Tourangeau, President
Dayton, OH

MARYLAND
Francis Scott Key Chapter
https://www.sksi.net/fskitea/index.html
John B. Schab, President
Aberdeen, MD

Southern Maryland Chapter
Bill Darden, President
Patuxent River, MD

DC/NORTHERN VIRGINIA
George Washington Chapter
http://gw-itea.org
Michael Wetzl, President
Washington, DC

VIRGINIA
Tidewater Chapter
Jeanine McDonnell, President
Suffolk, VA

SOUTHEAST REGION
Mike McFalls, Vice President

ALABAMA
Rocket City Chapter
Leigh Christian, President
Huntsville, AL

FLORIDA
Central Florida Chapter
David P. Grow, President
Orlando, FL

Emerald Coast Chapter
http://itea-ecc.org
Robert A. Crist, President
Eglin AFB, FL

GEORGIA
Atlanta Chapter
http://iteaAtlanta.org
Alvan Kimbrough, President
Smyrna, GA

SOUTH CAROLINA
Charleston Chapter
Philip Charles, President
Hanahan, SC

Volunteer Chapter
Nickolas Frederick, President
Arnold AFB, TN

SOUTHWEST REGION
Gregory D. Lamberth, Vice President

COLORADO
Rocky Mountain Chapter
http://www.itea-rmc.org
Christopher Mayette, President
Colorado Springs, CO

ARIZONA
Huachuca Chapter
Jonathan Woodruff, President
Sierra Vista, AZ

Valley of the Sun Chapter
Robert A. Olson, President
Scottsdale, AZ

NEVADA
Southern Nevada Chapter
Steve L. Moraca, President
Las Vegas, NV

NEW MEXICO
Roadrunner Chapter
Stephen “Buzz” Sawyer, President
Albuquerque, NM

White Sands Chapter
Douglas D. Messer, President
White Sands, NM

UTAH
Great Salt Lake Chapter
Jefferey D. Peterson, President
Dugway, UT

WEST REGION
Jack Sears, Vice President

CALIFORNIA
Antelope Valley Chapter
http://www.iteaavchapter.org
Michael Berard, President
Edward AFB, CA

Channel Islands Chapter
Christopher J. Weal, President
Point Mugu, CA

China Lake Chapter
Bettye R. Moody, President
China Lake, CA

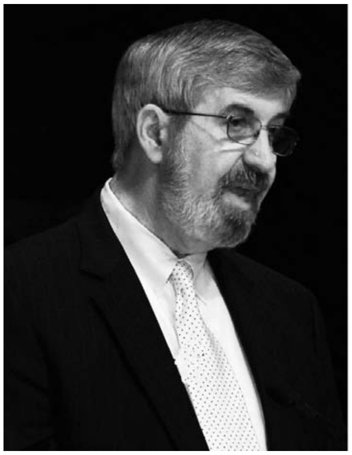
Greater San Diego Chapter
Daniel Phalen, President
San Diego, CA

WASHINGTON
Pacific Northwest Chapter
Vacant
Seattle, WA

PACIFIC REGION
Stewart V. Burley, Vice President

AUSTRALIA
Southern Cross Chapter
Peter G. Nikoloff, President
Edinburgh, South Australia

HAWAII
Mid-Pacific Chapter
Sandy D. Webster, President
Kekaha, HI




(Left) ***George Ryan*** *addresses February luncheon. (Right) Chapter president Mike Wetzl thanks speaker George Ryan*

(JCAS) event conducted at the Test Week 2010 Conference.

He said "the T&E Executive Agent Board of Directors is re-energizing the RELIANCE initiative for inter-service range investment planning". The Board is re-examining the number, type, and membership of RELIANCE panels, improving synergy with the Strategic Plan, consolidating products, and better coordinating planning with POM Program Reviews.

In a vigorous Q&A period, Foulkes said that in addition to working with the Services and Defense Agencies with respect to strategic planning, the TRMC is actively interacting with the non-Defense agencies including the Department of Homeland Security, (DHS), the National Aeronautics and Space Administration (NASA), and the Federal Aviation Administration (FAA), that TRMC has much interface with the Joint Forces Command especially concerning test and training, and that if the increased Defense emphasis on design of experiments increases the amount of testing needed it will have an effect on TRMC range planning. After his talk, Dr. Foulkes received an ITEA globe from chapter president **Mike Wetzl** as a memento of the occasion.

February Meeting At its monthly luncheon at the Army Navy Country Club in Arlington, VA on February 17th, the George Washington Chapter heard **Mr. George R. Ryan**, Special Advisor to the Department of Homeland Security (DHS) Director of Operational Test and Evaluation describe DHS testing.

Ryan started by contrasting DHS to Department of Defense (DOD). Noting that DHS has 22 components to DOD's four, and that while DOD reports to 4 Congressional committees, DHS reports to 86. He said there is a basic difference in the tested items because DOD tests systems mostly meant to defeat and destroy things while DHS test systems intended to protect and save things. The Department was established in 2003, and since 2006, much effort has gone into maturing acquisition processes including T&E. Testing organization and T&E policy are modeled so much like the DOD that DHS sends personnel to Navy and Army T&E training, and the Navy conducts some DHS testing and some testing is done at Army facilities such as White Sands Missile Range and Dugway Proving Ground, and the Navy modeled quite a bit like the DOD and is now largely in place. A major difference from DOD testing is that a significant step in planning is determining who will be the operational test agency, because unlike DOD there are no such established T&E agencies. The newly appointed Director of OT&E and Standards, Mr. Gary Carter, comes from the Navy and many staffers have experience in from the military services. There are now DHS personnel to plan, conduct, and report testing but staffing for review and approval is still in development. DHS OT&E results in a Letter of Assessment similar to DOD's Beyond LRIP Report.

He went on to say the Department's largest systems are the Electronic Baggage Screen System and the Coast Guard's new National Security Cutter WMSL Program which are among about 85 systems at the two highest acquisition levels. Most of the development testing is done by contractors, except in the Coast Guard, but operational test directors are government personnel. Some systems are operationally tested before the buy decision, which requires significant Memorandum of Agreements with a host facility for the test. Like DOD there are some test limitations, such as the restrictions on weapons grade nuclear materials and biological testing being conducted in public spaces. Ryan concluded by saying the Department is making progress in developing its T&E organization, procedures, and resources.

In a robust Q&A period Ryan said DHS is developing standards in the same office as T&E, that oversight of nonmajor systems T&E is not yet fully developed partly due to resources, that DHS is responsible for the civilian side of cyber security and has started its T&E, and that DHS has a requirement not known to DOD to “not impact commerce”, an example of which is the TWIC CARD with a 3-second criterion. After his talk, George Ryan received an ITEA globe from chapter president Mike Wetzl as a memento of the occasion.

March Meeting At its monthly luncheon at the Army Navy Country Club in Arlington, VA on March 24th, the George Washington Chapter heard **Mr. Rick Quade** of the Office of the Department of Navy Deputy for Test and Evaluation discuss the Tri-Service T&E Board of Directors’ Initiative to revitalize the T&E “RELIANCE” process

Quade said T&E Reliance is the collaborative effort of the Services and Defense Agencies working together to promote effective T&E infrastructure and investment management with the goal of providing cost-effective and efficient operation without regard to ownership, which can be summarized as “no unwarranted duplication.” He said the process is managed by the Test Resource Advisory Group (TRAG) which owns the involved resources and recommends infrastructure requirements and investment priorities. It is implemented by eleven Reliance Panels who review, validate, and prioritize needs and provide a report covering ten years of critical needs, impact if not met, and overall assessment of infrastructure status.

Five of the Panels are Air, Land, Sea, Space, and Electronic combat. Others address Armaments and Munitions, Range Instrumentation, Cyber warfare, Test Environments, and Targets. A new Panel is for Science and Technology Future Focus Areas.

(Left) ***Rick Quade*** *addresses March luncheon. (Right) Chapter president Mike Wetzl thanks speaker Rick Quade.*

Quade said changes in the Reliance process that contribute to revitalization include separating needs and solution phases, increased participation of OSD and Defense Agencies, increased responsibilities of Reliance Panels, increased transparency of submitted needs, institution of a continuous process improvement approach, streamlining of the process, and increase alignment with leadership.

Quade said the next major event will be a Test Capabilities Executive Review during the Defense Test Resources Management Center Test Week, June 14th - 16th, at which the Panels will brief T&E executives and provide summary assessments. He went on to describe top priorities for T&E infrastructure in 2011 and the top issues and concerns, including testing of next generation weapons systems, sensor fusion, and complex integrated mission systems packages.

In conclusion Quade said Defense and Services leadership is committed to revitalizing T&E Reliance, improvement of the process will result in a streamlined process that produces only mandatory products, and the revitalization process is well underway. After the talk, chapter president Mike Wetzl presented Rick Quade an ITEA globe as a memento of the occasion

■ ■ ■

South Jersey Chapter

The ITEA South Jersey Chapter awarded a $250 scholarship to high school student, **Logan Hicks** from Egg Harbor City NJ, on March 19, 2011 for an excellent science project that demonstrated good testing practices and skills at the Jersey Shore Science Fair.

■ ■ ■

Tidewater Chapter

The Tidewater Chapter would like to thank our 2010 President, **Mr. Brian Ensogna**, for a job well done. Carrying on this duty for 2011 is the newly elected chapter officers:

Ms. Jeanine McDonnell
2011 President
Mr. Kevin Long
2011 Vice President

Mr. Erwin Sabile
Secretary
Mr. Jim Spinelli
Treasurer

■ ■ ■

Southern Nevada Chapter

The ITEA Southern Nevada Chapter is ready to host the 2011 Instrumentation Workshop (May 9th-12th) in partnership with the Antelope Valley Chapter and the Ridgecrest Chapter. On May 9th you have a choice of expanding your technical knowledge in the tutorials or enjoying the sun soaked golf course honing your skills on the links.

The ITEA Southern Nevada Chapter has recognized local Engineering Students at UNLV with scholarship funds in the past year, but now we are digging a little further into the future. We are looking to the local magnet high schools to spark the interest of students early enough to promote getting them on a education track towards Engineering and Sciences. With all the cuts to education going on in our State to meet the budget woes, our chapter is working with the EAST TECH magnet school to provide test equipment to support the new programs and to make it a relevant hands-on program for the students. As we all choose our course in life it starts with the encouragement of a good teacher, the tools of the trade and a sense of what the Engineering career fields can offer.

■ ■ ■

ITEA-Central Florida Chapter

The ITEA Central Florida Chapter is coordinating with local INCOSE Chapter on establishing joint meetings as introductory chapter functions. We are also planning to begin plant tours of local firms, with Dot Decimal being first plant tour (not yet scheduled). Interviews with small businesses and programs throughout the Central Florida region are underway for posting on the emerging web site. We have identified multiple members for nomination to ITEA-CF board at our next board meeting. Also, we are supporting an International Conference in areas of education, golf (jointly with Women in Defense's local chapter), promotion, and AV.

■ ■ ■

Emerald Coast Chapter

The Emerald Coast ITEA Chapter is honored to announce six winners of a $1,000 scholarship. The following 2011 winners competed against a field of talented local high school seniors and were recognized for their academic achievements, extracurricular activities, and their interest in scientific/engineering disciplines that may lead to careers associated with test and evaluation:

Mr. Luke Ausley
Collegiate High School
Niceville, FL
Mr. Kyle Fields
Milton High School
Milton, FL
Mr. Connor Gray
Choctawhatchee High School
Ft Walton Beach, FL
Ms. Abigail McCool
Niceville High School
Niceville, FL
Mr. Andrew Neale
Paxton High School
Paxton, FL
Mr. James Pool
Milton High School
Milton, FL

In other news, the Emerald Coast ITEA Chapter hosted an open forum discussion on 29 March 2011 at Eglin AFB. The guest speaker was **Mr. George Axiotis**, Deputy Director, Air Warfare, Office of Director, Development Test and Evaluation (DDT&E), who presented an open discussion on "The mission of DDT&E" and how it relates to the Eglin's test and evaluation and acquisition infrastructure. Mr Axiotis stressed that the key points for DDT&E in the acquisition life cycle are the major milestone decisions (MS A, MS B, and MS C). DDT&E is charged with ensuring that development test and evaluation (DT&E) is adequately planned for, resourced, and executed to support the decisions made at these milestone decision points. From MS A to MS B DDT&E will understand and guide requirements so that the concept of operations is understood. At that point the scope, cost, and risk of the program can begin to be quantified. At MS B the overall T&E strategy will be captured in the TEMP. DDT&E's focus at MS C is the assessment of the performance and analysis against the key performance parameters/key system attributes/critical technical parameters (KPP/KSA/CTP) as defined in the capability development document (CDD). The designated Milestone Decision Authority (MDA) requires adequate information at MS C to commit the government to buy and produce the system in test. Generally speaking the only test data that has been collected at MS C is from the contractor's in-house testing, ground tests (arena, sled test, wind tunnel, etc), and DT flight tests. It is likely very little or no operational testing will have occurred at this stage in the program, therefore the criticality of the information gathered during DT cannot be underestimated. DDT&E must analyze the scope of testing conducted and the performance maturity, and determine what the risk is, if any, for the DoD to proceed to production. In closing, Mr Axiotis stressed the concern throughout DoD that the importance of a well understood, properly resourced DT&E has been somewhat lost in recent years. Recognizing that importance and ensuring adequate focus on DT&E from a program's inception is a key tenet of DDT&E. Equally important is ensuring T&E is efficient and meaningful and the T&E infrastructure is available to answer the question of system performance.

On a final note, the Emerald Coast ITEA Chapter is pleased to announce that the 4th quarter winner of "Tester of the Quarter" was the **Suite 7E Test Team** (24 members) of the Operational Flight Program/ Combined Test Force (OFP/CTF) at Eglin AFB. The Suite 7E Test Team is responsible the testing and fielding the new Suite 7E OFP for the F15E's. Under a very aggressive schedule the Suite 7E Test Team plans and executes hundreds of test missions with thousands of test points to ensure the F15E Suite 7E OPF is ready to safely provide the war fighter with the latest technology.

■ ■ ■

White Sands Chapter

The White Sands ITEA Chapter hosted the 16th Annual Live-Virtual-Constructive Conference the week of January 24, 2011 in El Paso, Texas. **Mr. Gilbert Harding**, Director of Systems Engineering at the White Sands Missile Range, served as the Conference Chair and **Mr. Doug Messer**, President of the Chapter, served as the Technical Chair and was responsible for the planning of the conference until his responsibilities supporting test efforts in the Pacific caused him to have to transition his role to **Mr. Charles Garcia**, NewTec President and CEO, representing NewTec and its parent company, the TRAX International Company for the execution of the conference.

On Monday, January 24, while many of our attendees took the opportunity to attend educational courses to support their training needs; others started the conference with a shot-gun start at the annual golf tournament. Regardless of how they started the day, everyone joined in at the kickoff reception spotlighting the exhibitors.

Ms. Stephanie H. Clewer, ITEA President, welcomed the group and spoke about the new initiatives ITEA is undertaking in 2011. She was followed by **Gilbert Harding** and **BG John S. Regan**, WSMR Commanding General, as they welcomed the audience to the southwest and the LVC Conference. **Dr. John B. Foulkes,** Director, Test Resource Management Center, and Office of the Under Secretary of Defense (OUSD), Acquisition Technology and Logistics (AT&L), addressed the 200 attendees as he kicked off the conference as the keynote speaker.

Recognizing the BRAC's impact at Ft. Bliss and the TRIAD (WSMR, Holloman AFB, and Fort Bliss) the White Sands team secured **Brigadier General Stephen M. Twitty,** Deputy Commanding General at Fort Bliss as their luncheon speaker. BG Twitty captivated the audience with his insight to the integration, growth, and vision of the training capability using LVC principals. Special guest speaker, **Mr. George Rumford**, T&E/S&T Program Manager, Test Resource Management Center, was exceptional as he presented, *LVC Integration 2020.*

The conference also included three panel discussions. The *VV&A of LVC for T&E* panel moderated by **Mr. Timothy Menke**, Technical Director, AFMC ASC/XRA, included **Mr. Michael D. Crisp**, Deputy Director, Air Warfare, DOT&E, **Dr. Marion Williams**, Adjunct Research Staff, Institute for Defense Analyses, **Mr. John Carlile**, The Boeing Company, and **Maj. Gen. David Eichhorn**, Commander, AFOTEC. The panel on *Distributed Testing*, moderated by **Mr. Chip Ferguson**, Program Manager for the Joint Mission Environment Test Capability (JMETC) included **Mr. Steve Kreider**, USA/ASC, **RADM David Dunaway**, Commander, COMOPTEVFOR, **Mr. Richard Clarke,** JITC, and **Mr. John Kriz,** NAVAIR. Finally, the third panel, *Solving the Standards Puzzle*, moderated by **Mr. John W. Diem**, DMSO included **Dr. Amy E. Henninger**, Technical Advisor, Center for Army Analysis, **Mr. Dennis W. Reed,** Deputy Division Head, Battlespace M&S Division, NAVAIR, and **Mr. Timothy Menke**, AFMC ASC/XRA.

The technical tracks throughout the week were on VV&A of LVC Distributed Environments for Testing, Network Support to the LVC Environment: LVC Support, Theory of VV&A, Application of LVC VV&A: Best Practices, LVC VV&A Synthesis and Way Forward, Methods and Techniques in LVC M&S Environment, Systems Engineering/Systems of Systems Engineering, and a town hall meeting style discussion on distributed testing and the successes and challenges from the programs perspective.

As always, the Chapter was very proud to announce their scholarship program and was pleased to provide scholarship checks to students from the New Mexico Institute of Mining and Technology University (New Mexico Tech), New Mexico State University (NMSU), the University of Texas at El Paso (UTEP), and New Mexico Military Institute (NMMI).

Hosting a conference of this nature provides multiple benefits. It not only provides a meeting that exceeds the expectations of the attendees technically, but it also allows other business to be accomplished. By setting aside a room for dignitaries and delegates to meet and address on-going business throughout the week, we not only create a good working environment, but due to the combination of testers and trainers in this forum, we fostered increase synergism between the two domains. And finally, ITEA is made up of dedicated volunteers who are committed to providing the T&E community with a forum to learn, share, and advance. We congratulate the team from White Sands for their untiring efforts to host this annual conference from planning through execution and we look forward to hearing about the Conference in 2012.

■ ■ ■

Association News

ITEA WANTS YOU! Call for Volunteers

Volunteers are at the heart of our association, and our strength is a result of the time and effort provided by those who volunteer to serve. ITEA offers many opportunities at the local, regional, and international levels for its member to contribute their time, talents and energy by giving back to the test and evaluation industry, helping us advance the profession, and supporting the association. If you want to become part of the ITEA volunteer team, now is the time to let us know.

There are many volunteer opportunities to volunteer at either your ITEA Chapter or to serve on an ITEA Board Committee, such as: Awards Committee; Chapter and Individual membership Development Committee; Corporate Membership Development Committee; Communications Committee; Education Committee; Elections Committee; Event Committee; History Committee; Publications Committee; Rules and Bylaws Committee; Strategic Planning Committee; Technology Committee; and the Ways and Means Committee. If you have experience, passion and a willingness to share, please consider participating as a volunteer on one of these committees. One of our goals is to have volunteer groups that represent the diversity of backgrounds, experience and demographics of the test and evaluation community.

If you have a desire to contribute to ITEA's ongoing development and management of the Association's governance, services, and guidance for the profession, you are encouraged to learn more about our volunteer opportunities.Please contact **James Gaidry**, ITEA Executive Director, by e-mail at jgaidry@itea.org, or at 703-631-6222, for more information.

ITEA WELCOMES JAY WEAVER AS NEW DIRECTOR OF EDUCATION

The International Test and Evaluation Association (ITEA) is pleased to announce that **Mr. John "Jay" Weaver** has joined the Executive Office staff as the Association's Director of Education. Jay's career highlights include military service with the U.S. Air Force and over 25 years in leading professional development, adult education, and individual certification programs. Jay previously also served in senior staff positions within the non-for-profit association community and was an instructor at both the U.S. Air Force Academy and the University of Maryland University College.

The ITEA Director of Education will work to ensure that ITEA is recognized as the premier provider of a 'continuum of learning' and provides a variety of education opportunities containing high value content that is timely, relevant, and easily accessible, and that meets the needs of both ITEA members and the T&E industry. Specific responsibilities for this position include:

- CURRICULUM DEVELOPMENT: Work with T&E subject matter experts and volunteers to identify the education needs of both the Membership and the T&E professional, and develop ITEA's overall professional development strategy and education curriculum.
- REPURPOSE CONTENT: Develop high value education content that can be repurposed through diverse and varied

learning opportunities (live courses, tutorials, webinars, on-demand web-enabled content, ITEA Journal articles, textbooks, handbooks, quick reference guides, social networking, video/audio recordings, packaged content for Chapter meetings, etc.).

- PROMOTION: Promote the value of continuous learning to members and the industry.
- ACADEMIC ADOPTION: Encourage the adoption/integration of ITEA's education curriculum by academic institutions.
- ACADEMIC SCHOLARSHIPS: Work with ITEA Chapters to facilitate their academic scholarships programs.
- GRANT FUNDING: Secure grant funding for education programs that support the mission of ITEA.
- CERTIFICATION PROGRAMS: Develop and manage accredited certification programs that enhance the image of the T&E professional.

Please join us in welcoming Jay to the ITEA family and the T&E community.

ITEA Welcomes New Corporate Members

Air Academy Associates
Air Academy Associates has spent more than two decades assisting clients in reducing variation, improving efficiency and creating better, faster and lower cost products and services. We support organization with our practical experience and tools to implement change and generate return on their investment. Based on industry proven strategies and tools our programs target improvements in development, design, manufacturing, testing and operations.

The key element of our strategy is our "Keep it Simple Statistically" (KISS) approach to gaining knowledge. This strategy allows non-statisticians in industry, government and the service sector to embrace the philosophy and effectively apply continuous process improvement needed to optimize the bottom line. Our foundation lies in our passion for efficient, effective and accurate data collection and analysis through the use of Design of Experiments (DOE). We have helped a variety of groups within the DoD realize the need for improved testing and validation through DOE and Combinatorial Testing.

Our network of more than 35 highly qualified and certified consultants have extensive experience in government, business and industry. This background, combined with targeted statistical software applications, allows us to meet the demands of a diverse client base. In addition to structured training, we provide on-site project mentoring, independent data analysis, experimental design analysis, and advanced testing solutions. Our training aids, textbooks and powerful, user-friendly software packages are integrated in our materials and set us apart from our competition. For more information on our products and services, see our Website at www.airacad.com or call 800-278-1277 for more information.

Avion Solutions, Inc.
Avion is a Veteran and Employee-Owned Small Business that provides high quality Specialized Engineering, Software Development, Logistics and Technical Support services. Formed in January 1992 in response to the engineering and analytical needs of the U.S. Army Aviation community providing technical expertise support in the following areas:

Engineering Services, Logistics Services, & Software Development
Software Systems Engineering/Development/Fielding
Automated Information Systems (AIS)/Automated Identification Technology (AIT) Army Aviation Logistics Support - Enterprise Asset Management, Technical Data and Configuration Management
Critical Asset Management: Qualification, Engineering Testing, Product Assurance
Comprehensive Life Cycle Engineering and Logistics Services
Test and Evaluation Management and Planning

Avion supports the Department of Defense (DoD) Army customers including: OSD, Test Resource Management Center (TRMC), Program Executive Office, Aviation (PEOAVN) Project Management Offices (PMO); Research Development and Engineering Command (RDECOM) Aviation System Integration Divisions; Special Operations Aviation; and Aviation and Missile Command (AMCOM) functional elements and related organizations. Additionally Avion has a customer base in the State and local Government organizations as well as in the commercial sector. As an ISO 9001:2008 certified company, the business processes are well established and documented. With more than ten business locations across the country and approximately 150 employees Avion has the right expertise and personnel to respond to any task in the defined above technical areas.

CETEST - Test and Analysis Centre
CETEST, a non-profit organization situated in the North of Spain, is a testing laboratory certified to ISO 17025:2005 by the Spanish National Certification Agency that provides test and consulting services in the world of railways and capital equipment. As part of the Basque Science, Technology and Innovation Network, CETEST's activities are centered on validation and approval of new vehicle designs (metros, tramways, commuter trains, freight trains and very high speed trains) during the manufacturing process. Production processes for manufacturing state-of-the-art technology railway vehicles and capital goods requires specific testing to validate the designs. Similarly, when high-speed vehicles are being developed the safety and comfort requirements are more stringent; and this entails lengthy and complex product approval processes. CETEST sets out to be a one-stop laboratory for validating the designs of railway vehicles and their auxiliary gear, as well as capital equipment. Please visit us at www.cetest.es.

Lockheed Martin Corporation
Headquartered in Bethesda, Md., Lockheed Martin is a global security company that employs about 133,000 people worldwide and is principally engaged in the research, design, development, manufacture, integration and sustainment of advanced technology systems, products and services. The Corporation's 2009 sales from continuing operations were $44 billion. For additional information, visit our website at www.lockheedmartin.com.

■ ■ ■

ITEA Corporate Members

Advanced Test Equipment Rentals
ACRA CONTROL, Inc.
Advanced Sciences and Technologies
Advanced Systems Development Inc. (ASD)
AEgis Technologies Group, Inc.
Agency for Defense Development
Air Academy Associates
AI Signal Research, Inc.
Alion Science and Technology
AMERICAN SYSTEMS
AMPEX Data Systems Corp.
Applied Resources, Inc.
Arcata Associates, Inc.
Argon ST, Network Systems
ARINC, Inc.
Astro-Med, Inc.
Avion Solutions, Inc.
AVW Technologies, Inc.
BAE Systems, Aerospace Solutions
Battelle
Boeing Company, The
Booz Allen Hamilton, Inc.
CALCULEX, Inc.
Calspan Corporation
COLSA Corporation
CETEST-Test and Analysis Centre
CSC, Applied Technology Group (ATG), Training and Range Support Services
Cubic Defense Applications
Data Systems and Technology, Inc.
Defense Acquisition University (DAU)
Department of Defence – RANTEAA
Dewetron, Inc.
DRS Training and Control Systems, LLC
Dynamic Science, Inc.
Epsilon Systems Solutions, Inc.
ERC, Incorporated
EWA Government Systems, Inc.
Fabreeka International, Inc.
General Dynamics C4 Systems
Georgia Tech Research Institute
Glacier Technologies, LLC
Herley Industries, Inc.
Imagine One Technology & Management, Ltd.
InDyne, Inc.
ITT Test & Support Systems
Jacobs Technology
JT3, LLC
L-3 Global Security and Engineering Solutions
L-3 Telemetry – West
Life Cycle Engineering
Lockheed Martin Corporation
MacAulay-Brown, Inc.
ManTech International Corporation
MEI Technologies, Inc.
MIL Corporation, The
NetAcquire Corporation
NMSU/PSL, 21st Century Aerospace
Northrop Grumman Corporation
Photo-Sonics, Inc.
PURVIS Systems, Inc.
QinetiQ North America – Systems Engineering Group
QUADELTA, Inc.
Qualis Corporation
Rockwell Collins, Inc.
RoundTable Defense, LLC
RT Logic
Sabre Systems, Inc.
SAIC
Science Applications International Corporation
Scientific Research Corporation
Sierra Lobo, Inc.
SPARTA, Inc. d.b.a. Cobham Analytical Solutions
Spectrum Sensors and Controls
Spiral Technology, Inc.
SRA International
SURVICE Engineering Company
SYMVIONICS, Inc.
System Development Center – CSIST
Syzygy Technologies, Inc.
Tactical Information Exchange Integration Office
TASC, Inc.
TRAX International Corporation
Trideum Corporation
U.S. Army Developmental Test Command
Weibel Scientific A/S
Westech International, Inc.
Windmill International, Inc.
Wyle

Last updated: June 2011

The ITEA Journal
Ad Sales

Ad Rates

The ITEA Journal Ad Sales Contact

Bill Dallas
Ph: 703.631.6226
wdallas@itea.org
4400 Fair Lakes Court,
Suite 104
Fairfax, VA 22033
Fax: 703.631.6221

2011 Production Schedule

Issue	Space Reservation	Ad Material
March	1/12	1/25
June	4/12	4/25
September	7/12	7/25
December	10/12	10/25

Journal Specifications

Trim Size: 8 1/2 x 11 inches. Journal trims 1/8 inch off top, bottom, and outside edge. Live matter should be a minimum of 1/2 inch inside the trimmed edges, and a minimum of 1/2 inch should be allowed for the bind.

Graphics should be a minimum of 350 dots per inch. Add 1/4 inch for bleeds.

Formats: We accept TIFF or EPS format for both Macintosh and PC Platforms. We also accept files in the following native application formats: Adobe Acrobat (.pdf), Adobe Photoshop (.psd), Macromedia FreeHand (.fh), Canvas (.cvs), InDesign (.id), QuarkXPress (.qxd), Illustrator, (.ai) CorelPoint (.cdr), PowerPoint (.ppt), and Pagemaker (.pmd).

***

Disclaimer: All claims for errors in advertisements must be made in writing and received within ten days of publication and will be considered only for the first insertion of the advertisement.

AD Type	ITEA Partner Rates (ITEA Corporate Members, Sponsors, and Exhibitors)			Standard Rates		
	1x	2x	4x	1x	2x	4x
Cover 2	$2500	$2200	$2000	$3000	$2700	$2500
Cover 3	$2420	$2120	$1920	$2900	$2600	$2400
Cover 4	$2740	$2440	$2240	$3300	$3000	$2800
	4-Color					
Full page	$2100	$1800	$1600	$2500	$2200	$2000
1/2 page	$1700	$1400	$1200	$2000	$1700	$1500
1/4 page	$1300	$1000	$ 800	$1500	$1200	$1000
2-page spread	$3300	$3000	$2800	$4000	$3700	$3500
	Black/White					
Full page	$1620	$1320	$1120	$1900	$1600	$1400
1/2 page	$1300	$1000	$ 800	$1500	$1200	$1000
1/4 page	$ 900	$ 600	$ 400	$1000	$ 700	$ 500
2-page spread	$2580	$2280	$2080	$3100	$2800	$2600

Chapter Locations

NORTHEAST REGION
Vacant, Vice President

CONNECTICUT & RHODE ISLAND
Narragansett Bay Chapter
Vacant

MASSACHUSETTS
New England Chapter
Michael E. Keller, Sr., President
Boston, MA

NEW JERSEY
South Jersey Chapter
John Frederick, President
Atlantic City, NJ

EAST REGION
Robert A. Vargo, Vice President

EUROPE
European Chapter
Steve Lyons, President
United Kingdom

ISRAEL
Israeli Chapter
Aaron Leshem, President
Haifa, Israel

OHIO
Miami Valley Chapter
Stephen Tourangeau, President
Dayton, OH

MARYLAND
Francis Scott Key Chapter
https://www.sksi.net/fskitea/index.html
John B. Schab, President
Aberdeen, MD

Southern Maryland Chapter
Bill Darden, President
Patuxent River, MD

DC/NORTHERN VIRGINIA
George Washington Chapter
http://gw-itea.org
Michael Wetzl, President
Washington, DC

VIRGINIA
Tidewater Chapter
Brian C. Ensogna, President
Virginia Beach, VA.

Jeanine McDonnell
Suffolk, VA

SOUTHEAST REGION
Mike McFalls, Vice President

ALABAMA
Rocket City Chapter
Leigh Christian, President
Huntsville, AL

FLORIDA
Central Florida Chapter
David P. Grow, President
Orlando, FL

Emerald Coast Chapter
http://itea-ecc.org
Robert A. Crist, President
Eglin AFB, FL

GEORGIA
Atlanta Chapter
http://iteaAtlanta.org
Alvan Kimbrough, President
Smyrna, GA

SOUTH CAROLINA
Charleston Chapter
Philip Charles, President
Hanahan, SC

Volunteer Chapter
Nickolas Frederick, President
Arnold AFB, TN

SOUTHWEST REGION
Gregory D. Lamberth, Vice President

COLORADO
Rocky Mountain Chapter
http://www.itea-rmc.org
Christopher Mayette, President
Colorado Springs, CO

ARIZONA
Huachuca Chapter
Jonathan Woodruff, President
Sierra Vista, AZ

Valley of the Sun Chapter
Robert A. Olson, President
Scottsdale, AZ

NEVADA
Southern Nevada Chapter
Steve L. Moraca, President
Las Vegas, NV

NEW MEXICO
Roadrunner Chapter
Stephen "Buzz" Sawyer, President
Albuquerque, NM

White Sands Chapter
Douglas D. Messer, President
White Sands, NM

UTAH
Great Salt Lake Chapter
Jefferey D. Peterson, President
Dugway, UT

WEST REGION
Jack Sears, Vice President

CALIFORNIA
Antelope Valley Chapter
http://www.iteaavchapter.org
Michael Berard, President
Edward AFB, CA

Channel Islands Chapter
Christopher J. Weal, President
Point Mugu, CA

China Lake Chapter
Bettye R. Moody, President
China Lake, CA

Greater San Diego Chapter
Daniel Phalen, President
San Diego, CA

WASHINGTON
Pacific Northwest Chapter
Vacant
Seattle, WA

PACIFIC REGION
Stewart V. Burley, Vice President

AUSTRALIA
Southern Cross Chapter
Peter G. Nikoloff, President
Edinburgh, South Australia

HAWAII
Mid-Pacific Chapter
Sandy D. Webster, President
Kekaha, HI

ITEA Corporate Members

Advanced Test Equipment Rentals
ACRA CONTROL, Inc.
Advanced Sciences and Technologies
Advanced Systems Development Inc. (ASD)
AEgis Technologies Group, Inc.
Agency for Defense Development
Air Academy Associates
AI Signal Research, Inc.
Alion Science and Technology
AMERICAN SYSTEMS
AMPEX Data Systems Corp.
Applied Resources, Inc.
Arcata Associates, Inc.
Argon ST, Network Systems
ARINC, Inc.
Astro-Med, Inc.
Avion Solutions, Inc.
AVW Technologies, Inc.
BAE Systems, Aerospace Solutions
Battelle
Boeing Company, The
Booz Allen Hamilton, Inc.
CALCULEX, Inc.
Calspan Corporation
COLSA Corporation
CETEST-Test and Analysis Centre
CSC, Applied Technology Group (ATG), Training and Range Support Services
Cubic Defense Applications
Data Systems and Technology, Inc.
Defense Acquisition University (DAU)
Department of Defence – RANTEAA
Dewetron, Inc.

DRS Training and Control Systems, LLC
Dynamic Science, Inc.
Epsilon Systems Solutions, Inc.
ERC, Incorporated
EWA Government Systems, Inc.
Fabreeka International, Inc.
General Dynamics C4 Systems
Georgia Tech Research Institute
Glacier Technologies, LLC
Herley Industries, Inc.
Imagine One Technology & Management, Ltd.
InDyne, Inc.
ITT Test & Support Systems
Jacobs Technology
JT3, LLC
L-3 Global Security and Engineering Solutions
L-3 Telemetry – West
Life Cycle Engineering
Lockheed Martin Corporation
MacAulay-Brown, Inc.
ManTech International Corporation
MEI Technologies, Inc.
~~Modern Technology Solutions~~
NetAcquire Corporation
~~NewTec~~
NMSU/PSL, 21st Century Aerospace
Northrop Grumman Corporation
Photo-Sonics, Inc.
PURVIS Systems, Inc.

QinetiQ North America – Systems Engineering Group
QUADELTA, Inc.
Qualis Corporation
~~Raytheon Missile Systems~~
Rockwell Collins, Inc.
RoundTable Defense, LLC
RT Logic
Sabre Systems, Inc.
SAIC
Science Applications International Corporation
~~Sandia National Labs Interactive Systems Simulation & Assessments~~
Scientific Research Corporation
Sierra Lobo, Inc.
SPARTA, Inc. d.b.a. Cobham Analytical Solutions
Spectrum Sensors and Controls
Spiral Technology, Inc.
SRA International
SURVICE Engineering Company
SYMVIONICS, Inc.
System Development Center – CSIST
Syzygy Technologies, Inc.
Tactical Information Exchange Integration Office
TASC, Inc.
TRAX International Corporation
Trideum Corporation
U.S. Army Developmental Test Command
Weibel Scientific A/S
Westech International, Inc.
Windmill International, Inc.
Wyle

Last updated: March 2011