Interoperability Test and Evaluation: A System of Systems Field Study

Dr. John Colombi, Maj. Brannen C. Cohee, and Maj. Chuck W. Turner
Air Force Institute of Technology

Effective operational test and evaluation (OT&E) is an essential part of successful systems and software engineering. But increased program dependencies, network-centric operations, and growing interoperability requirements have greatly complicated test and evaluation. This article examines the policy, process, and practice of the Air Force (AF) test and evaluation programs, such as Force Development Evaluations (FDEs), particularly during the sustainment of systems. Several observations are made regarding the current process and five areas are emphasized for improvement.

An increasing challenge is facing the OT&E community when operating across multiple weapon systems at various stages of development. After studying the policy, process, and practice associated with an AF Major Command's (MAJCOM's) OT&E program, this article concludes that the AF, while espousing a testing philosophy of *seamless verification*, still needs to transition to a more integrated system of systems (SoS) approach to planning and executing OT&E. Too often, a fielding decision for a single system modification is the goal for smaller test events. This system-centric focus can be misplaced. Indeed, with the dawning of network-centric operations, the SoS imperative is even greater for the successful integration, test, and evaluation of warfighting capabilities. This article highlights several observations and offers some areas for process improvement.

To confirm these challenges, this analysis focused on a geographically dispersed network of ground stations that work with AF and DoD surveillance and reconnaissance (S&R) platforms to provide data, information, and knowledge services for the joint commander and forces in the field. A decade ago, these S&R platforms and their attending ground stations existed in isolation. Since then, however, operational necessities and technological opportunities have birthed a system of increasingly interdependent hardware and software systems, spanning sensors, platforms, data links, communication networks, and software-intensive ground processing resources. The evolution of this SoS has produced remarkable advances in the warfighting capability, but this integration has also created a host of systems engineering (SE) and enterprise management challenges, such as OT&E planning and execution.

The SoS Challenge

The very nature of an SoS makes the enterprise management of traditionally system-centric support processes, such as OT&E, difficult. As Mark Maier points out, even though an SoS operates synergistically, systems in an SoS can operate and are managed independently [1]. This is a premise that is expected to continue into the foreseeable future; therefore, a single organization or program manager will not suddenly take complete managerial control of all systems within the applicable SoS. One reason is that a system may participate in multiple mission threads interacting with a variety of joint organizations and weapon systems. The "Systems Engineering Guide for Systems of Systems" recognizes that an SoS is usually not born of a single development effort but emerges as complex combinations of newly acquired and legacy systems—each with their own management, operations, and support communities—that evolve over time [2]. Annette Krygiel notes that the purpose and capabilities of an SoS change as functions are added, removed, and modified [3]. Compounding the complexity of SoS SE, Pin Chen and Jennie Clothier observe that component systems in an SoS are often systems of systems themselves [4]. All of this complicates the current test and evaluation approaches.

Despite these challenges, operational demands are forcing the AF and DoD to co-evolve historically system-centric processes like OT&E to support the development of SoS and net-centric capabilities [5]. Therefore, to better understand the need for and obstacles to this co-evolution, this study focused on a particular OT&E process and the extent to which it supports an SoS approach. The OT&E process chosen, called an FDE, is managed at the MAJCOM level to make fielding decisions for operational weapon systems as incremental upgrades are made during sustainment. It should be noted that an FDE is one of several types of OT&E called out in AF Instruction (AFI) 99-3, Capability Based Test and Evaluation. Others include Initial Operational Test and Evaluation, Qualification Operational Test and Evaluation, Follow-on Operational Test and Evaluation, Tactics Development and Evaluation, the Weapons System Evaluation Program, Operational Utility Evaluation, Operational Assessments, and Early Operational Assessments.

MAJCOMs conduct FDEs for programs requiring full-rate production or fielding decisions if the AF Operational Test Center chooses not to conduct OT&E. This is typically true for Acquisition Category III programs or maintenance modifications. After a system has been fielded and has entered the sustainment phase of its life cycle, the primary type of test and evaluation used to verify and validate smaller system upgrades is the FDE. As stated in the Air Combat Command instruction, the focus of FDE is a subset of OT&E. FDEs are primarily concerned with sustainment, pre-planned product improvement, as well as tactics, techniques, and procedures development. The objective is to demonstrate the operational effectiveness and suitability of a system as evolutionary upgrades are made to sustain its relevance to the warfighter. In the Air Combat Command (ACC) FDE process, ACC Test Centers (e.g., the AF Warfare Center, the AF Information Operations Center, and the Air National Guard AF Reserve Test Center) are responsible for planning and executing FDEs.

Next, the OT&E process is examined in the context of its governing law and policy. Then, a case study is documented which focuses on a particular test event that involved a networked ground system and one of its airborne partners.

Observations From the Field Study

From Congress, direction for OT&E flows down from four sections of Title 10 of the U.S. Code: Director of OT&E; Survivability Testing and Lethality Testing Required Before Full-Scale Production; OT&E of Defense Acquisition Programs;

and Low-Rate Initial Production of New Systems. The DoD then implements the policies into directives, instructions, and regulations. The AF further clarifies its entire test and evaluation process in the 99-series of departmental instructions, such as AF Policy Directive 99-1, Test and Evaluation Process, and AFI 99-103, Capabilities Based Test and Evaluation [6]. This policy defines the purpose of the AF test and evaluation process and provides a framework for test activities. It also expands on the two major types of tests: Developmental Test and Evaluation and OT&E. The AF philosophy clearly reflects the verification and validation of mission-level capabilities, an emphasis on *seamless verification* across the developmental and operational test activities, the use of an integrated test team (ITT) for test management, and the efficient sharing of test data through a common database. Finally, AF MAJCOMs define operating procedures, as in ACC Instruction (ACCI) 99-101, ACC Test and Evaluation. This instruction further clarifies MAJCOM OT&E procedures. Two examples that stand out in ACCI 99-101 are the Electronic Project Order, for tasking organizations and resources, and the use of a yearly Test Priority List. While extensive, the test policy hierarchy is observed to provide good vision and insightful guidance and establishes a foundation for being able to handle today's SoS test challenges. Several observations on the current process are provided in the following sections.

Seamless Verification Still Has Seams

Through our policy analysis, it was discovered that the AF endorses a capabilities-based concept of seamless verification, especially by mandating the use of ITTs. The ITT consists of a cross-functional group of empowered representatives from multiple disciplines and organizations and is co-chaired by operational testers and the program manager. The problem lies in the meaning of *integrated*. While DoD policy includes both the *horizontal* integration of test and evaluation throughout a system's life cycle and the *vertical* integration of systems under interoperability testing (as shown in Figure 1), the AF sees integrated testing almost exclusively in life-cycle terms [6]. In addition, AF policy also mandates the use of open, shared databases for managing test information. This integrated information management could be how the acquisition and test communities could begin to bridge the vertical seams.

Thus, seamless verification still has seams separating test activities among interdependent weapons systems. AF policy does not mandate organizational structures and management processes that help OT&E organizations conduct their testing activities in an SoS framework. However, the prevalence of systems of systems and the evolution toward net-centricity demands test processes that are both integrated throughout the life cycle of a single weapon system and integrated across entire sets of operational capability.

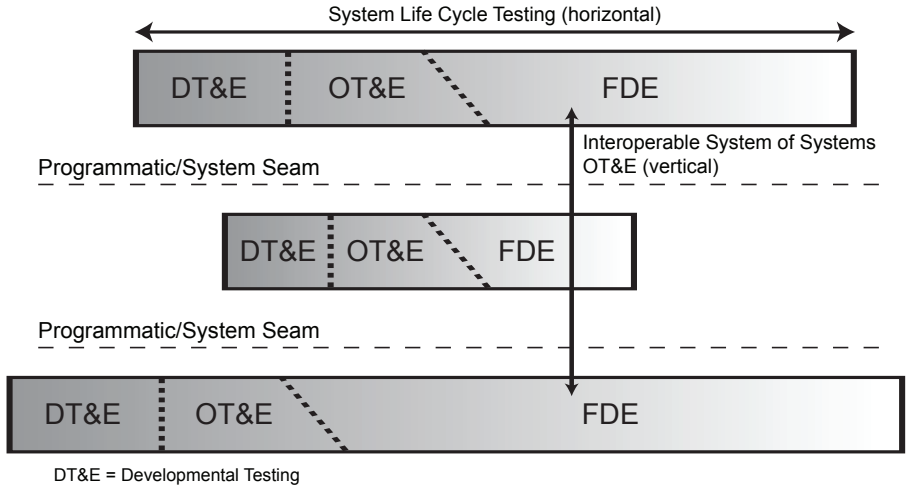

Figure 1: *System Integration During Interoperability Testing*

SoS Approach Not Built In

It was found that the FDE process is flexible in that it can be applied to both system-centric and capabilities-based SoS test events. Even so, it does not include steps to intentionally evaluate SoS capabilities rather than individual systems. Rather, it relies on the insight and foresight of the MAJCOM staff and the test center organizations to properly scope events to approximately demonstrate full warfighting capabilities.

Indeed, the test project manager, generally appointed from one of the MAJCOM's test center organizations, is the central figure in defining the scope of the test. Whether determining the composition of the planning team, developing test objectives, requesting additional support for testing, or developing the test plan, the extent to which FDE demonstrates the operational capability of an SoS depends largely on the vision and initiative of the individual project manager. The MAJCOM's process does not include functions that obligate a project manager to scope a test at the SoS level.

Increasing Load on OT&E

The studied MAJCOM has experienced a dramatic increase in OT&E requirements (from around 200 just five years ago to around 300 today) and the number of short-notice or out-of-cycle requirements (from around 10 percent of the total number of test events five years ago to approximately 40 percent today). The war on terror has certainly contributed to both the number and urgency of OT&E requirements, but one subject matter expert interviewed believes the increased number of acquisition spirals and increments along with the introduction of non-traditional software-intensive weapon systems—such as the networked ground system in our case study—has led to a more expansive, dynamic, and complex environment for MAJCOM-led testing.

Experts and senior leaders have argued that the growing interdependence of systems organized in net-centric architectures will exacerbate the increased load (i.e., the number, complexity, tempo, and expense of test events), calling it exponential growth. With testing resources either remaining static or decreasing over time, this increased load will force the MAJCOM—as well as the broader test community—to develop new methods for testing and evaluating net-centric capabilities.

System-Centric Approach Breaks Down

Our case study starts with an FDE for a modified sensor on board an airborne platform. This sensor modification was needed to support interoperability with a new data link architecture, and it included minor software upgrades to networked ground stations. The MAJCOM assigned the event to its organization responsible for testing modifications to airborne platforms. Understanding the organic relationship between the platform and the ground system, test planners knew they needed to demonstrate the end-to-end interoperability of four interdependent

systems: the sensor, airborne platform, data link, and ground station. Yet, the test was not planned as a demonstration of the operational effectiveness and suitability of an SoS capability, but as a validation of the single-sensor system with the support of these other contributing systems. While this may seem like a subtle distinction, it led to a variety of coordination and communication breakdowns throughout test planning. In retrospect, the FDE and subsequent fielding decision should have been for the combined SoS, made up of the sensor, airborne platform, data link, and ground stations, which would have necessarily involved a broader cross-section of stakeholders from the genesis of the test event. When SoS thinking is not *baked into* the overall test and evaluation process, even highly interdependent systems will have difficulty coordinating test events that are effective and relevant for the constituent systems and the SoS as a whole.

Recommendations for Net-Centric OT&E

Though important efforts have been made at all levels to promote SoS-level testing, the default focus of testing is still on individual systems as opposed to whole capabilities. As net-centric operations mature, this approach will have to change. Indeed, as our field study indicates, the DoD is already feeling the pressure that was predicted nine years ago:

> Testing systems will become far more complex since the focus will not be on the performance of individual systems, but on the performance of federations of systems. [7]

At the National Defense Industrial Association Test and Evaluation Summit in 2004, DoD officials elaborated on the net-centric challenges to traditional, system-centric testing [8]:

- The shifting focus from platforms to capabilities and SoS solutions.
- The increasing complexity of systems combined with increasing interdependencies among systems of systems.
- The increasing operational demand for broader and deeper integration among disparate systems.
- The exponential growth in functional and physical interfaces introduced by the proliferation of network participants (both newly developed and legacy).
- The increased requirements for test and evaluation initiated by the evolutionary acquisition philosophy of build-a-little, test-a-little.

As the heralds of net-centricity emphasize, the DoD's transition from Industrial Age (platform-centric) to Information Age (net-centric) operations must include a co-evolution of supporting processes. Testing is one of those supporting processes that must co-evolve with technology. The following recommendations are believed to best improve MAJCOM-level OT&E, but should be extensible to the AF and DoD OT&E. These can be considered attributes of a future process; while not meant to be exhaustive, they provide a starting point for continuing research on how to evolve today's process to complement a more interoperable net-centric environment.

Scope OT&E Events at the SoS Level

This article advocates an SoS (instead of a system-centric) approach to OT&E. However, the question arises of how to appropriately scope the boundaries of SoS test events. This is where the DoD AF and AF Enterprise Architecture (EA) could offer practical help. As the use of EAs continues maturing, they will provide effective models for assessing how weapon systems can and should interoperate in order to provide warfighting capabilities to the joint commander. These models will help planners define the boundaries of an SoS, and they will document what the MITRE Corporation's Prem Jain calls a *mission thread*:

> A precise, objective, description of an important task … a time-ordered operational event diagram that captures discrete, definable, interactions among human operators and/or technological components. [9]

Jain argues that these mission threads will support modeling and simulation (M&S) activities for net-centric test and evaluation. In the same way the AF uses high-fidelity simulators to slash the costs associated with training its pilots, the test community could use mission thread-based M&S to validate SoS and net-centric capabilities at a fraction of the cost, time, and operational impact incurred by live, end-to-end tests.

Validate SoS Interoperability

By advocating SoS testing, an endless web of end-to-end interoperability tests is not envisioned with every other possible weapon system and every configuration. Rather, changes to individual weapon systems should be evaluated according to net-readiness criteria to validate their interoperability with the rest of the SoS or net-centric enterprise. This does not mean just checking to see if a modified weapon system is IP-enabled. Net-readiness is a comprehensive concept that implies interoperability at many layers of the communications hierarchy and beyond: physical, logical, syntactic, and semantic [2]. For nearly 30 years, both government and industry have actively explored research on interoperability measurement with the goal of creating a straightforward way of measuring, reporting, and then improving the interoperability of complex networks of people, equipment, processes, and organizations. Researchers have used more than 30 definitions of interoperability and have documented more than 60 distinct types of interoperability, numerous interoperability attributes, and 14 foundational interoperability measurement models and methodologies [10].

For SoS testing, a set of net-readiness objectives (see Table 1) based on the DoD's Net-Centric Data Strategy [11] is proposed. Incorporating these objectives into SoS events would ensure that modified systems continue to conform—at their interface with the network—to the convergence protocols specified in the net-centric architecture. Instead of evaluating all end-to-end relationships to validate interoperability within the SoS, a test event could confirm the integrity of the SoS simply by demonstrating the modified system's adherence to the network's convergence protocols. This technique would greatly reduce the test load on OT&E organizations while simultaneously allowing them to conduct evaluations focused on the operational effectiveness of the net-centric SoS, as opposed to the individual systems within that SoS.

Prioritize FDEs According to Operational Risk

Although a simulation and net-readiness demonstration at the network interface will help mitigate the test load associated with SoS and net-centric operations, decision makers will still expect a certain level of live, end-to-end testing in realistic scenarios to validate higher-risk capabilities. There will always be too much to test, forcing the operational and test communities to develop a reasonable means of prioritizing test events. Complicating this issue is a fundamental property of net-centric operations: New transactions, interdependencies, missions, and capabilities as additional (even unanticipated) sensor, shooter, and command and control nodes join the network. The very nature

of net-centric operations implies that the DoD will never completely anticipate all of the relevant operational nodes and precisely how those nodes will interoperate to accomplish a mission.

Thus, the crucial criterion for prioritizing and scoping SoS OT&E events is operational risk. For low-risk development or sustainment efforts, operational decision makers may need to be satisfied with developmental test results validated in an M&S-based operational test. For medium-risk projects, testers may use a synthetic test strategy that employs a small number of distributed operational events in an M&S framework. The test and evaluation community may need to reserve traditional end-to-end events for the highest risk efforts. The key will be for operational decision makers to set an appropriate risk threshold for each development or sustainment program and seek the best value test option in terms of time, money, and testing/operational resources to achieve that threshold.

Objective	Description
Visible	Users and applications can discover the existence of data assets through catalogs, registries, and other search services. All data assets (intelligence, non-intelligence, raw, and processed) are advertised or *made visible* by providing metadata, which describes the asset.
Accessible	Users and applications post data to a *shared space*. Posting data implies that: (1) descriptive information about the asset (metadata) has been provided to a catalog that is visible to the enterprise and (2) the data is stored such that users and applications in the enterprise can access it. Data assets are made available to any user or application except when limited by policy, regulation, or security.
Understandable	Users and applications can comprehend the data, both structurally and semantically, and readily determine how the data may be used for their specific needs.
Trusted	Users and applications can determine and assess the authority of the source because the pedigree, security level, and access control level of each data asset is known and available.
Agile	Many-to-many exchanges of data occur between systems, through interfaces that are sometimes predefined or sometimes unanticipated. Metadata is available to allow mediation or translation of data between interfaces, as needed.
Responsive	Perspectives of users, whether data consumers or data producers, are incorporated into data approaches via continual feedback to ensure satisfaction.

Table 1: *A Template for Net-Readiness Testing Objectives*

Focus on Interfaces

Without trivializing the complexities of an SoS, perhaps more emphasis may have to be placed on the definition, development, and test and evaluation of interfaces. Interfaces and interface management have always been an important aspect of SE. Maier and Rechtin state this design heuristic: "The greatest leverage in system architecting is at the interfaces … the greatest dangers are also at the interfaces" [12]. According to the "Defense Acquisition Guidebook," this heuristic can be an interface that is:

> The [logical], functional, and physical characteristics required to exist at a common boundary or connection between persons, between systems, or between persons and systems. [13]

During an interface control document (ICD) review of one space program, we examined the impact of interface design and development. Over a three-year period following contract award, 596 program engineering items were examined. These items included: requirements changes, specification updates and clarifications, ICD changes, SE management documents, baseline schedule changes, test plans, and proposed risk-mitigation actions; ICD-related actions comprised 190 (or one-third) of the total number of actions. A second aspect of interface management was in contract costs. From a set of 77 contractual modifications after critical design review, 43 (nearly 56 percent) were in some way related to interfaces either through studies, ICD updates, or implementations of necessary requirement changes. More interestingly, ICD-related issues resulted in nearly $31.5 million (or 44 percent) of the cost impact to this program. Although this is just one example, it is an indication of interface management challenges during development and clearly represent an area that will require similar effort during OT&E.

Unfortunately, interface management alone may be insufficient to understand the complexity within an SoS. For network-centric information systems, one must examine the internal properties and behaviors of the logically connected systems and their users who find, fuse, modify, and ultimately use shared data.

Employ Integration Environments

The heavy use of M&S and synthetic testing to supplement traditional test and evaluation presupposes the use of integration environments to build SoS and net-centric operational capabilities. Annette Krygiel calls an integration environment:

> … a concept, not an organization. It is the environment of people, processes, and infrastructure used by a team consisting of acquisition and operational personnel to manage the integration before the product is deployed for an operation or an experiment and to sustain it afterward. [3]

Thus, an integration environment is a concept that transcends not just OT&E but is an essential SE infrastructure for the early and continuous integration of testing efforts in SoS development and sustainment. Thus, an integration environment is critical in squeezing the most overall value from OT&E and in reducing the amount of effort required late in the development cycle [14].

Clearly, integrated databases open to all test and evaluation stakeholders are just the tip of the iceberg in terms of the tools needed to achieve seamless verification. Employing integration environments would allow test teams to achieve synergy throughout the life cycle of component systems and across the networked SoS. It would allow early, comprehensive, and ubiquitous test and evaluation throughout the development of SoS capabilities, reducing the test load on test center organizations and giving them the freedom to focus on adding value where it counts for MAJCOM decision makers: in reducing the operational risk of fielding and employing warfighting capabilities.

Summary

This field study of the policy, process, and practice of FDE concludes that the testing community must shift from system-centric testing to a more SoS approach. The

DoD's ongoing transformation to net-centric operations makes the co-evolution of SoS test and evaluation an even greater imperative. While the five areas for improvement were derived from MAJCOM FDE practice, they reflect similar concepts for more formal OT&E. Improving the operational realism of network-centric environments, ensuring timely performance of operational information, and facilitating adequate test and evaluation resources continue to be priorities of the OT&E director [15]. Likewise, AF OT&E continues to be challenged with SoS test planning and execution. The software engineering community must continue to design and test capabilities with an SoS focus and develop advanced modeling and simulation capabilities to enable affordable SoS testing in a net-centric environment. Likewise, the test community, while emphasizing seamless verification, needs to make continued progress in capabilities-based interoperability testing across information-intensive SoS.◆

References

1. Maier, Mark W. "Architecting Principles for Systems of Systems." Systems Engineering. 1.4 (1998): 267-284.
2. Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics. Systems Engineering Guide for System of Systems. Vers. 1.0. Aug. 2008 <www.acq.osd.mil/sse/docs/SE-Guide-for-SoS.pdf>.
3. Krygiel, Annette. Behind the Wizard's Curtain. Washington, D.C.: DoD Command and Control Research Program. 1999 <www.dodccrp.org/files/Krygiel_Wizards.pdf>.
4. Chen, Pin, and Jennie Clothier. "Advancing Systems Engineering for Systems-of-Systems Challenges." Systems Engineering. 6.3 (2003): 170-183.
5. Alberts, David S. Information Age Transformation. Washington, D.C.: DoD Command and Control Research Program. 2002 <www.dodccrp.org/files/Alberts_IAT.pdf>.
6. USAF. "Capabilities Based Test and Evaluation." Air Force Instruction 99-103. Washington, D.C.: USAF, 6 Aug. 2004.
7. Alberts, David S., et. al. Network Centric Warfare. Washington, D.C.: DoD Command and Control Research Program. Aug. 1999 <www.dodccrp.org/files/Alberts_NCW.pdf>.
8. Lamartin, Glenn F. The Role of T&E in the Systems Engineering Process. Proc. of the National Defense Industrial Association T&E Summit. Washington, D.C.: 17 Aug. 2004 <http://proceedings.ndia.org/487F/lamartin.pdf>.
9. Jain, Prem. "Mission-Model Driven Process: Test and Evaluate Net Centric Capabilities." The MITRE Corporation. Technical Paper. 2007.
10. Ford, Thomas C., et al. Survey on Interoperability Measurement. Proc. of the 12th Annual International Command and Control Research and Technology Symposium. 2007 <www.dodccrp.org/events/12th_ICCRTS/CD/html/papers/096.pdf>.
11. Stenbit, John F. DoD Net-Centric Data Strategy. 9 May 2003 <www.defenselink.mil/cio-nii/docs/Net-Centric-Data-Strategy-2003-05-092.pdf>.
12. Maier, Mark W., and Eberhardt Rechtin. The Art of Systems Architecting. 2nd ed. New York: CRC Press LLC, 2002.
13. DoD. Defense Acquisition Guidebook. 12 Dec. 2004 <https://akss.dau.mil/dag>.
14. Brown, C. David. Transformation of Army Test and Evaluation. National Defense Industrial Association T&E Summit, Washington, D.C. 17 Aug. 2004 <http://proceedings.ndia.org/487f/brown.pdf>.
15. McQueary, Charles. Proc. of National Defense Industrial Association SE Conference. 23 Oct. 2007 <www.dtic.mil/ndia/2007systems/SEGS/McQueary. pdf>.

About the Authors

John Colombi, Ph.D., is an assistant professor of SE at the AF Institute of Technology (AFIT). He teaches graduate courses and leads sponsored research in support of the SE program. Retiring after 21 years in the AF, Colombi led Command, Control, Communications, Computer (C4), Intelligence, and Reconnaissance systems integration activities including SE for the Airborne Warning and Control System at Hanscom AFB.

Dept. of Systems and Engineering Management
AFIT/ENV
Wright-Patterson AFB, OH 45433
Phone: (937) 255-3355 ext. 3347
Fax: (937) 255-4981
E-mail: john.colombi@afit.edu

Maj. Brannen C. Cohee is director of operations at the 315th Training Squadron at Goodfellow AFB, Texas. He supervises the training of AF-enlisted and officer intelligence specialists. He is currently deployed as the intelligence planner with the Air Component Coordination Element in Kabul, Afghanistan. Cohee received his commission from the USAF Academy and a master's degree in public policy from Harvard University before entering intelligence officer training.

315th Training Squadron
154 Canberra ST
Goodfellow AFB, TX 76908-4002
Phone: (325) 654-5649
DSN 477-5649
E-mail: brannen.cohee@goodfellow.af.mil

Maj. Charles "Chuck" W. Turner is chief of the Current Operations Section for the U.S. Central Command's C4 (J6) Directorate. He recently graduated from AFIT, completing a course of study in C4 and Information Systems with a special emphasis in cyber defense. As a communications computer officer, Turner has served in a variety of operational, support, and staff positions involving communications systems.

HQ CENTCOM/CCJ6-CO
7115 S Boundary BLVD
MacDill AFB, FL 33621-5101
Phone: (813) 827-6458
Fax: (813) 827-2211
E-mail: turnerc@centcom.mil