A FRAMEWORK FOR SYSTEM OF SYSTEMS EVALUATION
WITHIN AN AIRBORNE INTELLIGENCE, SURVEILLANCE, AND RECONNAISSANCE ENVIRONMENT

Rick S. Thomas, N. Clark Capshaw, and Paul M. Franken

Federal *test and evaluation* agencies, particularly those associated with the U.S. military, are grappling with the challenge of evaluating *system of systems (SoS)* or a *family of systems (FoS)*—in short, developing methods whereby the contribution of *individual systems* can be evaluated *when operating in combination with other systems,* and determining the effectiveness when various subcomponents are added or removed from the overall SoS. In this article, the authors present a proposed framework for conducting such evaluations through integrating developmental testing, operational testing, and operational performance data into the evaluations. A recent example of the evaluation of a suite of aerial *intelligence, surveillance, and reconnaissance (ISR)* systems is also discussed, relating the aerial ISR evaluation to the proposed framework.

image designed by Nina Costine »

		Form Approved OMB No. 0704-0188

Report Documentation Page

1. REPORT DATE **2010**	2. REPORT TYPE	3. DATES COVERED **00-00-2010 to 00-00-2010**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
A Framework for System of Systems Evaluation Within an Airborne Intelligence, Surveillance, and Reconnaissance Environment	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Defense Acquisition University,9820 Belvoir Road,Fort Belvoir,VA,22060-9910**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**14**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

The Game Changer

Federal test and evaluation agencies, particularly those associated with the U.S. military, are grappling with the challenge of evaluating system of systems (SoS) or a family of systems (FoS)—in short, developing methods by which the contribution of *individual systems* can be evaluated *when operating in combination with other systems*, and determining the effectiveness when various subcomponents are added or removed from the overall SoS.

This is particularly challenging when trying to assess airborne intelligence, surveillance, and reconnaissance (ISR) sensors employed as an SoS, due not only to the abundance of sensors (Imagery Intelligence [IMINT], Signals Intelligence [SIGINT], Measurement and Signals Intelligence [MASINT], etc.), but also the myriad ways in which these sensors can be used in combination with one another to achieve mission effects. Further complicating such an evaluation is the notion that ISR sensors are employed at different levels, both in the sense of mission focus—national, strategic, operational, and tactical (N, S, O, T)—and physical altitude.

The aim of this article is to develop and present a framework for how an SoS evaluation might be realized. It begins by noting the special challenges of evaluating an SoS, introduces an analogy to aid in further discussion, and concludes by relating this analogy to a recent U.S. Army Test and Evaluation Command (ATEC) SoS evaluation of a suite of aerial ISR systems. The article also relates this framework to other recently published evaluation frameworks.

In the examples discussed in this article, ATEC used a systems-level methodology that concentrated on an effects-based approach of layers of sensors distributed across multiple aircraft. In focusing on multiple individual systems, ATEC focused on the effects of the systems and not the overall element or family; accordingly, this article will reference SoS versus FoS.

A Conceptual Model for an Airborne ISR SoS

Consider the model for airborne ISR assets presented in the Table. Note that each level presents different mission sets and capability gaps that must be addressed.

As illustrated in the Table, not only are there many kinds of airborne ISR sensor types (IMINT, SIGINT, and MASINT, etc.), but also many different mission sets, instantiations (satellites, manned aircraft, unmanned aerial vehicles), and levels of focus characterize the individual sensors. The Table is misleading in one respect; it shows an apparent clear demarcation from a mission set at one level (e.g., tactical) to the next (e.g., operational). In practice, these clear lines

A Framework for System of Systems Evaluation Within an Airborne Intelligence, Surveillance, and Reconnaissance Environment

October 2010 | 4 3 9

TABLE. A TAXONOMY OF AIRBORNE ISR SENSORS

Level	Focus	Examples of Mission Sets	Instantiation	Sensor Type
National	Very wide	Monitor nuclear weapons development and testing	Satellites	IMINT still photography, experimental/advanced and esoteric sensors
Strategic	Wide	Monitor troop massing at border areas for potential attack	High-altitude aircraft	IMINT, SIGINT, MASINT
Operational	Medium	Monitor troop movements in an existing theater of operations; monitor activities of personnel associated with enemy networks	Medium-to-low altitude aircraft and unmanned aerial vehicles (UAVs)	IMINT-wide area focus, full-motion video, aerial photography
Tactical	Narrow	Observe specific enemy attacks and activities, such as emplacement of improvised explosive devices, ambushes, etc.	Low-altitude aircraft and UAVs	IMINT full motion video

are at best fuzzy, and often nonexistent. Increasingly, the national level wants to know every bit of intelligence gained, even down to the tactical level; in turn, tactical operations incorporate as much national-level intelligence as possible. Trying to determine what asset contributes most effectively and its position within the Table is a tough problem—one which a rigorous SoS evaluation should address—but how?

Evaluating the Sum of the Parts or the Capability of the Whole?

By definition, an SoS is an amalgam of individual systems, each of which is designed to perform a specific function. When individual

systems are combined into a greater whole, this can conceivably change their character and function.

Testers often find it much easier to conduct evaluations of individual systems because the parameters, threats, and variables that are part of the individual system's tests are not complicated or influenced by other systems that could either augment or degrade the individual system's inherent capability.

In contrast, when evaluating an SoS, multiple additional challenges surface that evaluators need to consider. These challenges could be mitigated during an individual system test, so long as the evaluation teams are aware of force structuring of the SoS before conducting the individual test. The discussion that follows addresses some of the considerations that increase the complexity of effective SoS testing.

Optimization of Systems When Integrated With Complementary Systems

When individual systems are integrated as part of a larger SoS, an evaluation strategy must account for how the individual systems work to the complement or to the detriment of the other systems. To effectively evaluate these relationships, systems should be tested in an iterative fashion, first evaluating effectiveness of individual systems through isolated tests, and then determining the capability of the entire SoS while looking for synergistic benefits that may present themselves. Compatibilities can be determined by identifying all of the potential relationships the individual systems have with one another, distinguishing which relationships are most critical to the SoS, weighting the importance of these relationships, and then building these relationships into the SoS evaluation strategy. Duplicative capabilities or gaps in a particular capability must also be identified and factored into the overall assessment.

Force Structure for an SoS

The nature of SoS employment requires that certain force structure characteristics, which are not inherently part of the existing units, be in place to synergize the SoS. An operational unit is more readily capable of integrating an individual system into its structure because managing the capabilities and limitations of one system is a skill and discipline that soldiers, down to the individual level, are trained to do. However, integrating multiple systems with multiple capabilities and limitations, and the complex relationships between them, is significantly more challenging for a soldier in an operational environment to assimilate. Additionally, battle command training and leader development for the employment of individual systems, as compared to an SoS, is significantly different. With SoS employment,

A Framework for System of Systems Evaluation Within an Airborne Intelligence, Surveillance, and Reconnaissance Environment

October 2010 | 441

all stakeholders must be trained to understand the often-complex relationships that exist between the systems' sensors to ensure that each system sensor is employed as efficiently and effectively as possible. For example, a specific SoS includes assets that support all echelons of battle command, yet would be employed at the tactical level, thus requiring units to have a mechanism in place to maximize the capabilities, understand the echeloning of the capability, and synergize the entire SoS.

SoS In-Theater Evaluation Versus Domestic Test Range Evaluation

The logistics required to create a test event for an individual system are relatively simple compared to the logistics for an SoS test event. Coordinating the presence of an individual system at a test range with its respective personnel is a feasible task, even during a time of war. Conversely, coordinating the presence of multiple systems and their respective personnel at a domestic test range, all at the same time, for a test event during a time of war is extremely difficult, if not impossible. Quite often, the systems that comprise the SoS have been evaluated individually and are already deployed in theater.

The foregoing constraints limit SoS evaluation using traditional, domestic test range operational test and evaluation procedures, thus making in-theater evaluation a more attractive alternative. Yet, when conducting an in-theater evaluation, additional factors must be considered. The following discussion presents two factors that surfaced during ATEC's most recent in-theater SoS evaluation.

RED FORCE CONSIDERATIONS

The red force component that would be used in a domestic operational test would be based upon the intelligence community's best assessment of the tactics and capabilities of the enemy. Notwithstanding the best efforts of the intelligence community, assessments based on recent intelligence may not accurately represent activity of the enemy at that particular time. An in-theater assessment, however, measures the SoS against the true red force, and thus provides the most accurate assessment of the SoS capabilities and limitations. Unfortunately, knowledge of red force activities in an actual theater of war is limited—a situation that equates to SoS evaluation under uncertainty. For example, consider an SoS designed to find enemy improvised explosive devices (IEDs). If the SoS finds five IEDs, did it perform well or poorly? How does the evaluator know how many IEDs were emplaced by the red force? Evaluators can count the number of IED explosions that were documented and the number of IEDs found and cleared, but they have no idea how many IEDs remain

undiscovered and undetonated, or how many exploded without documentation.

CONCEPT OF OPERATIONS/TACTICS, TECHNIQUES, AND PROCEDURES (CONOPS/TTP) CONSIDERATIONS

Blue force CONOPS and TTP are defined for individual systems, but are often informal and evolutionary for the employment of the SoS. This presents an enormous challenge for the evaluator, who must develop measures of effectiveness for the SoS based upon CONOPS and TTP that are still under development. In addition, TTP are consistently modified based on enemy response to blue force actions. This constant operational dynamism makes it difficult to formalize measures that are relevant to the environment where the SoS is employed.

Beyond the limitations inherent in range testing or in-theater evaluation, the hierarchical relationship that exists between the employment of individual systems and the supporting tasks that enable effective operational SoS-level deployment must also be considered by the test community (Figure 1). SoS evaluation should examine the entirety of factors involved in SoS employment, applying a multidisciplined methodology to achieve a multidisciplined measure of operational effect. But, doing this in practice can quickly become

FIGURE 1. SYSTEM OF SYSTEMS PARADIGM
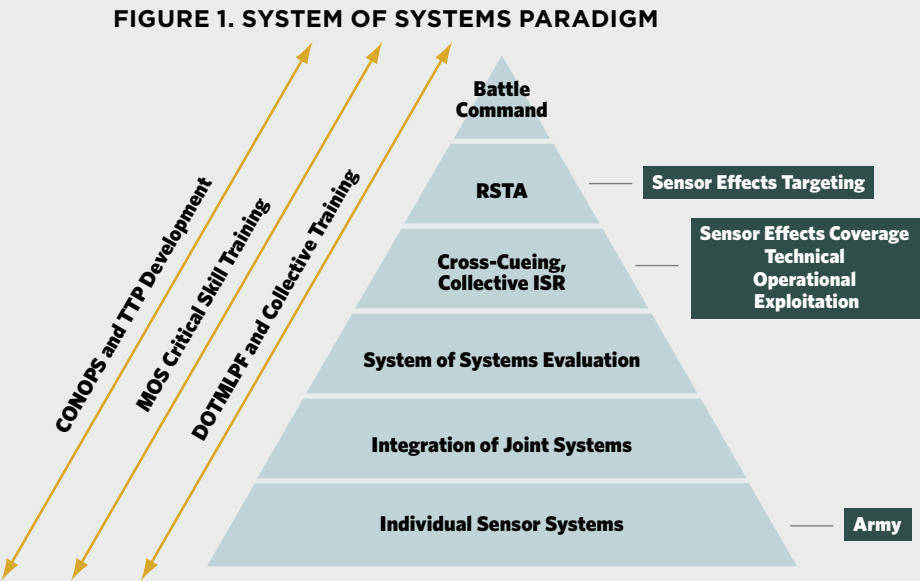
Note. DOTMLPF = Doctrine, Organization, Training, Materiel, Leadership and Education, Personnel, and Facilities; MOS = Military Occupational Specialty; RSTA = Reconnaissance, Surveillance, and Target Acquisition.

A Framework for System of Systems Evaluation Within an Airborne Intelligence, Surveillance, and Reconnaissance Environment

October 2010 | 443

an intractable problem, considering the number of variables involved and the difficulty in altering those variables one at a time.

The capabilities gained through modeling and simulation (M&S) and associated software packages, focused on insurgency operations, are essential to the overall evaluation of an SoS. M&S is needed that can ingest all of the relevant data from the system, assess the environment associated with insurgency operations, assess the immediate threat, and then fuse and maximize inputs from other systems operating under the same parameters to identify the best mix of assets at the strategic, operational, and tactical levels. The results from SoS evaluation should inform the community of not only the capability of the enterprise to engage and be successful, but how best to distribute the capabilities for the greatest overall operational effect.

The following hypothetical example taken from the world of sports may best illustrate the concept of SoS evaluation.

An Analogy for SoS Evaluation

Consider an analogy. A professional football coach is facing a decision: a first-round draft pick for a new player. This opportunity to draft a new player forces the coach to thoroughly evaluate the array of available players and their potential contributions to the team. Does he go for an offensive player or a defensive player? Does he go for a quarterback, a lineman, or a running back? If he chooses a quarterback, does he choose one from a college that has an aggressive passing game, or one from a school that has a run-based offense?

Now consider the data that the coach has for making his decision. Undoubtedly, he has accumulated data on the draft choice's speed, weight, strength, and performance in a college environment. If he is thorough, he also has some information about how the player performs in a high-stakes environment such as a championship game, his injury record, and how he fits in as part of a team. Though the coach should optimally consider individual players' skills (*system level*) in making a draft decision, his decision should be ultimately based on his evaluation of the player's potential to improve the overall performance of the team (*SoS level*).

Finally, the coach makes his decision and drafts a player. But, he doesn't put the new player in as a starter automatically. Instead, he uses training camp, pre-season games, and regular season substitutions to determine whether the new player merits a starting position with the team.

Let's suppose the coach has done all of this throughout the new player's first season, and at the end of the season the team's record is 11–5—a substantial improvement from a lackluster 8–8 perfor-

mance the previous year. Does the coach attribute the entirety of the improvement to the new player? Probably not, considering the team gained a number of new players—some from the draft, others from trades, and several through free agency. These new players, as well as factors beyond the coach's control such as schedule difficulty, have affected the team's performance. So, how does the coach determine what mixture of variables contributed to the team's improved record? How does he decide which players to retain and which ones to shed for the next season?

Each of the coach's inquiries and decisions represents a stage in the evaluation of an SoS, and it is instructive to the development of the framework to now extend this analogy to the evaluation of an SoS.

Relationship of Sports Analogy to an Airborne ISR SoS Evaluation

The sports analogy presented earlier illustrates some of the challenges of conducting an airborne ISR SoS valuation. First of all, are the capability gaps clear? Do decision makers know for what "positions" they are recruiting? Is there a gap in N, S, O, T mission coverage, or perhaps all four? Is there a reason to believe that an airborne ISR asset can fulfill one or more of these gaps?

Second, what is known about the "players"—the specific airborne ISR assets that might be added to an SoS? Much of this type of information is gleaned from *developmental testing (DT)*. DT enables evaluators to gain knowledge of the *technical* capabilities of the ISR asset (analogous to a player's weight, strength, and speed), but little or no knowledge on how the asset will perform when integrated with the "team."

Integration into the "team" begins with operational testing (or OT), which is often performed in an artificial environment, such as one of the Army's proving grounds. Use of such a test center allows evaluators to gain a better understanding of how the system undergoing testing will perform when used in an operational environment. A true operational environment will also enable evaluation of the system not only with its actual "teammates" (the Blue Force), but also in the presence of the opposing team (the Red Force).

Finally, the system is deployed and integrated with other ISR assets, all of which are dedicated to fulfilling a mission, but this mission may have components or effects at each of the N, S, O, T levels. Certain mission-related effects are measured, and things seem to be improving; coalition causalities, for example, have declined.

Now, the real dilemma begins! How is it possible to determine whether one particular combination of assets (a "team") is responsible for the improvement, or whether the improvement was due to exogenous factors (such as a troop surge or intelligence from

A Framework for System of Systems Evaluation Within an Airborne Intelligence, Surveillance, and Reconnaissance Environment

October 2010 | 445

an enemy defector)? How is it possible to determine whether certain assets (the "players") are performing well at their positions, or whether it would be better if certain of the ISR assets were "traded" for other assets?

Example of a Recent ATEC SoS Evaluation

ATEC has rich experience conducting SoS evaluations, including the recent evaluation of an Army suite of aerial ISR systems (an SoS). Figure 2 also shows how the *system, mission, test, and evaluation elements* were integrated into ATEC's recent evaluation.

A general discussion to demonstrate how the framework previously described can be applied to similar SoS evaluations follows.

ATEC began its SoS evaluation by performing thorough DT on each individual aerial ISR component of the SoS. Through DT, ATEC was able to examine the *technical* capabilities of each of the sensors. In this case, all of the sensors were IMINT sensors, though each was employed in a different way and was expected to add a unique capability to the final product.

Then, in a series of OT scenarios conducted at Yuma Proving Ground, ATEC was able to extend the *technical* results of the DT to gain a better understanding of these aerial ISR systems and how each might perform as part of an overall SoS. Although data were limited due to the operational conditions in theater, ATEC used scientific methods to complete the initial study of the SoS. In addition, ATEC developed test and threat protocols to ensure like testing across a

FIGURE 2. ATEC TFO SoS EVALUATION FLOWCHART
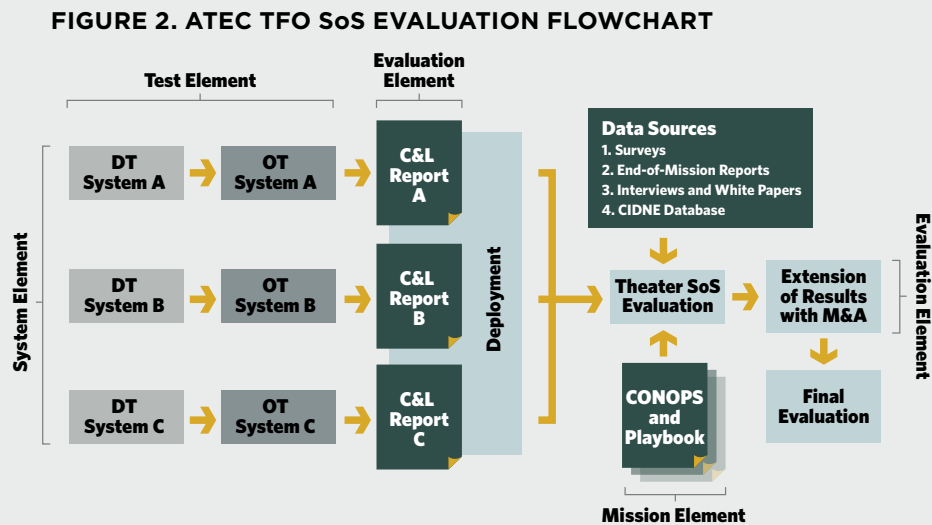
Note. C&L = Capabilities & Limitations; CIDNE = Combined Information Data Network Exchange; TFO = Transport Flow Optimization.

series of systems, and then supplemented this information with a forward operational assessment (FOA) team that collected data while systems were deployed in operational conditions.

ATEC's FOA team collected performance data from several sources: end-of-mission reports, user surveys, commander surveys, stakeholder interviews, white papers, and reports of enemy activities. Further, ATEC developed and used the Common IED Exploitation Target Set ontology and application to define conditions and standards, and model results for OT (Franken et al., 2009).

ATEC extended the results of this SoS evaluation by using modeling and analysis (M&A) methods. M&A methods were utilized to explore the possible mission effects of employing the SoS differently and in different combinations with a focus on insurgent methods, timing, and location of attack. The use of M&A created a virtual test environment for this SoS evaluation, and helped reveal a more operationally effective way (i.e., system composition, user tactics, flight schedules, etc.) to employ this SoS. The data from tests, operational assessment, and threat integration were used to conduct first-order validation of the insurgent model. More detailed verification and validation will be executed as data become available.

Relationship to Other Frameworks

The framework presented in this article is not intended to be used in isolation, but in combination with other frameworks that can help lend clarity to the evaluation of SoS. Two of these are addressed in the following discussion.

Simmons and Wilcox (2007) introduced the notion of a four-element framework for test and evaluation, encompassing system, mission, test, and evaluation elements in an integrated whole.

> The four-element framework provides a systematic approach to developing a T&E plan that evaluates mission capabilities, system effectiveness, and system suitability. The mission and system elements define what is to be evaluated. The mission-to-system interface links the elements together and ensures that the development of the evaluation and test elements always remain focused on the unit's ability to execute the mission when using the system. This provides a defined guideline for developing the evaluation measures and a roadmap for how the tests support the evaluation. (p. 66)

A National Research Council (2006) report cited "continuous process" as the framework upon which to meet the challenge of testing in an evolutionary acquisition environment.

A Framework for System of Systems Evaluation Within an Airborne Intelligence, Surveillance, and Reconnaissance Environment

October 2010 | 447

In evolutionary acquisition, *the entire spectrum of testing activities should be viewed as a continuous process of gathering, analyzing, and combining information in order to make effective decisions* [emphasis added]. The primary goal of test programs should be to experiment, learn about the strengths and weaknesses of newly added capabilities or (sub)systems, and use the results to improve overall system performance. Furthermore, data from previous stages of development, including field data, should be used in design, development, and testing at future stages. (p. 3)

The ATEC aerial ISR SoS evaluation example described earlier followed these lessons learned. The aerial ISR SoS was an evolutionary system, and all testing and evaluation activities were integrated into the overall evaluation. Additionally, M&A was used to extend the evaluation to examine and explore the other possible employment combinations and methods for this SoS in an asymmetric insurgent environment.

Conclusions

The importance and implementation of thorough SoS evaluation—distributing and measuring the effects of individual systems as they are integrated across the entirety of the SoS—poses a challenge to the test community. This challenge can be met by incorporating the lessons learned and data from multiple events, using M&A and scientific methodology to integrate and optimize the testing and providing relevant feedback to affected communities—typically soldiers employing the SoS or planners executing acquisition strategy. With the continued interdependence of SoS deployed and relied upon in operational environments, SoS evaluation must be capable of using all available assets to ensure operational realism is met in all test events, and relevant quantitative measures are applied in evaluating the SoS, ensuring a legitimate SoS evaluation, not merely an evaluation of a system within a system.

Author Biographies

Mr. Rick S. Thomas is a strategist with the Office of the Secretary of Defense (OSD) Test and Resource Management Center (TRMC) in Arlington, Virginia. He recently was the Division Chief, Air and Space ISR, during the assessment of "Suite of Aerial ISR Systems" mentioned in this article. Currently, he oversees the Department of Defense strategic plan for test resources. He has a degree in Management from the National-Louis University.

(E-mail: rick.thomas@osd.mil)

Dr. N. Clark Capshaw is an evaluator with the U.S. Army Test and Evaluation Command (ATEC), in Alexandria, Virginia. He has been the evaluator on several aerial ISR systems, including the systems reviewed in this article. He is also an online instructor for the University of Phoenix. His education includes an MS in Aerospace Engineering from the University of Dayton and a PhD in Leadership and Policy Studies from Vanderbilt University.

(E-mail: norman.capshaw@conus.army.mil)

Mr. Paul Franken is a contractor with ARES Systems Group, located in Alexandria, Virginia. ARES supports the U.S. Army Test and Evaluation Command (ATEC) with contractor support for the Test, Threat, and Evaluation communities, and provides technical, tactical, and materiel solutions for other federal and commercial clients. Mr. Franken was an active duty Army Aviator with tactical experience in Stability and Support Operations (SASO) and a veteran of Operation Iraqi Freedom (OIF 1).

(E-mail: pfranken@MAV6.com)

A Framework for System of Systems Evaluation Within an Airborne Intelligence, Surveillance, and Reconnaissance Environment

October 2010 | **4 4 9**

REFERENCES

Franken, P. M., Harrison, A. J., Holton, J. J., Macfarlan, D. L., Wowczuk, Z. S., Capshaw, N. C., … Russomanno, D. J. (2009). Development of an ontology-based tool to support the test and evaluation process for rapid acquisition of IED detection capabilities. *ITEA Journal, 30*, 300–308.

National Research Council, Oversight Committee for the Workshop on Testing of Dynamic Acquisition of Defense Systems. (2006). *Testing of defense systems in an evolutionary acquisitions environment.* The National Academies Press: Washington, DC.

Simmons, B. M., & Wilcox, C. J. (2007, September/October). The four-element framework: A mission-focused test and evaluation strategy. *ITEA Journal, 28*(3), 61–69.