Positioning Test and Evaluation for the Digital Paradigm

Dr. Laura Freeman  
Deputy Director Virginia Tech National Security Institute

Geoffrey Kerr  
Senior Research Associate Virginia Tech National Security Institute

Dr. Peter Beling  
Director, Intelligent Systems Division Virginia Tech National Security Institute

Dr. Alejandro Salado  
Associate Professor, Department of Systems and Industrial Engineering, University of Arizona

Dr. Kelli Esser  
Associate Director, Intelligent Systems Division Virginia Tech National Security Institute

Dr. Jeremy Werner  
Chief Scientist, DOT&E

Dr. Paul Wach  
Research Assistant Professor Virginia Tech National Security Institute

Dr. Sandra Hobson  
Deputy Director, DOT&E

Introduction

The United States (US) Department of Defense (DoD) and its supporting industry, research, and academia partners have proposed using Digital Engineering (DE) methods and tools to update traditional systems engineering (SE) and test and evaluation (T&E) practices to improve acquisition outcomes and accelerate traditional processes. The proposed transition to DE does this by taking advantage of computer science, modeling, analytics, and data sciences [1]. This DE transformation provides the opportunity for the DoD to reimagine the ways it utilizes data to plan, design, develop, test, deploy, and sustain new systems and capabilities in support of missions. Envisioned benefits span the entire acquisition lifecycle, and include: enhanced data sharing and management across stakeholders; improved employment of robust modeling, simulation and data analytic tools and methods; seamless integration of mission engineering (ME), SE, and T&E activities; and optimized use of data to inform decision making. While the trend continues toward adoption of DE, current programs are faced with practical challenges as to how to realize the vision with the current infrastructure, processes, and workforce.

In this paper we present a phased approach to help T&E overcome some of the early adoption and implementation challenges of DE transformation, positioning the community to leverage the benefits of this new paradigm. Our approach advocates for expanded use of DE and model-based systems engineering (MBSE) methods and tools, creating a specialized area of practice referred to as “MBSE for T&E”. We present an “MBSE for T&E” Roadmap with three phases, and outline a path forward for further research and measurement of progress toward achieving DE objectives within T&E.

Building Blocks for the DE Transformation

DoD and its supporting industry, research, and academic partners are in a transformative period of realizing a DE transformation to overcome current limitations associated with acquiring new or
enhanced system capabilities. One of the greatest limitations in current acquisition is the time required for system design and development. Current systems may spend a decade or longer in development before transitioning to operational use, with acquisition activities tending to occur in a sequential manner with data stovepipes. This traditional acquisition model puts the DoD at risk of being unable to maintain mission and operational superiority unless changes are made to better accommodate the increasingly rapid advances in emerging technologies and complex system capabilities.

The DE transformation is expected to reduce the time required to plan, design, develop, test, and deploy complex systems, while simultaneously increasing overall system performance and effectiveness by providing a digital feedback loop on how design changes translate into operational capability. Key enablers of this transformation are enhancing data sharing; improving the use of modeling, simulation, and data analytics; integrating across ME, SE, and T&E activities; and optimizing the use of data for decision making.

The building blocks for realizing these achievements were laid over a decade ago by the T&E community, with progress accelerating over the last five years. Below is a timeline of key initiatives and accomplishments from 2010 – today:

**2007 – 2008: Mandate for Integrated Testing and Definition of Integrated T&E**

The concept of Integrated T&E was first mandated by USD (AT&L) and DOT&E in 2007 in the 5000.02 and quickly followed up in 2008 with a joint memo defining integrated testing as:

*Integrated testing is the collaborative planning and collaborative execution of test phases and events to provide shared data in support of independent analysis, evaluation and reporting by all stakeholders particularly the developmental (both contractor and government) and operational test and evaluation communities.*

Notably, the definition focused on collaborative planning, but independent evaluations. The result has been a lack of emphasis on integrating information from data from across the acquisition life cycle due to concerns about real or perceived lack of independence in reporting.

**2010: The Integrated T&E Continuum, the Key to Acquisition Success**

In 2010 [4], in an ITEA editorial, DDT&E advocated for use of an integrated flow of information from DT&E to OT&E to overcome limitations in traditional acquisition process. The paper expanded the concept of integrated testing to include the continuum of T&E. This Approach spans the acquisition lifecycle with integration of developmental test and evaluation (DT&E) and operational test and evaluation (OT&E) facilitate ongoing learning and sharing of knowledge on system requirements, development, and performance spanning from pre-milestone A through post-milestone C [4]. The concept of test and evaluation across the acquisition lifecycle has been widely understood and accepted over the years, progress has been slow, largely due to implementation cost and the absence of an interoperable DE environment, interoperability of tooling, and unified data strategy to support complete end-to-end knowledge and data sharing throughout the T&E phases and acquisition lifecycle.

**2018: Launch of Digital Engineering**

In 2018, DoD leadership accelerated the push for better integration and data sharing through its release of the DoD Digital Engineering Strategy [1]. This strategy defines DE as “an integrated digital approach
that uses authoritative sources of system data and models as an acquisition lifecycle across disciplines to support lifecycle activities from concept through disposal [1]." It also outlines the need to modernize how DoD “designs, develops, delivers, operates, and sustains systems [1].”

**2022: Vision for T&E Transformation**

In June 2022, the DoD Director, Operational Test and Evaluation (DOT&E) built on the DoD Digital Engineering Strategy and its proposed DE transformation in its DOT&E Strategy Update 2022 [2], outlining a vision for transforming T&E infrastructure, processes, concepts, tools, and workforce to keep pace with the ongoing rapid evolutions in technologies, threats, and operating environments. Key aspects of this T&E transformation that are consistent with the DE transformation include:

- Harmonizing technical and programmatic data to create effective and efficient decisions and communication;
- Conducting operationally relevant assessments earlier in the development lifecycle (shift left);
- Innovating T&E data management tools and methods to measure and evaluate data-oriented operational performance;
- Developing and implementing credible digital tools and automation to propel better T&E, shorten T&E timelines, and illuminate performance shortfalls earlier;
- Leveraging contractor and development data in support of operation assessment;
- Integrating physics-based modeling and simulation, analytical models (e.g., discrete event simulations), and test data from multiple test phases to enhance understanding of technical risk and uncertainty projections.

**2023: Enabling T&E as a Continuum**

In March 2023, the Office of the Secretary of Defense (OSD) Developmental Test, Evaluation, and Assessments (DTE&A) in partnership with Naval Air Systems Command’s Capabilities Based T&E Office published “T&E as a Continuum” [5] with recommendations toward implementing a more risk-based, capability-driven, integrated T&E approach for complex system acquisition. The paper builds on the DoD Digital Engineering Strategy, and provides a “T&E as a Continuum” framework centered on three key attributes and three key enablers. Table 1 below lists the key attributes and key enablers.

**Table 1. “T&E as a Continuum” Key Attributes and Enablers.**

<table>
<thead>
<tr>
<th>KEY ATTRIBUTES</th>
<th>KEY ENABLERS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Capability and Outcome Focused Testing</td>
<td>Robust Live, Virtual, Constructive Testing</td>
</tr>
<tr>
<td>Agile, Scalable Evaluation Framework</td>
<td>Model-Based Environment</td>
</tr>
<tr>
<td>Enhanced Test Design</td>
<td>“Digital” Workforce</td>
</tr>
</tbody>
</table>
Amongst the three key enablers, the model-based environment was identified as the most critical based on its role in providing the digital backbone for modeling, simulation, data analytics, data sharing, and ME-SE-T&E integration throughout the T&E acquisition lifecycle.

### Moving from Vision to Implementation

Since the release of the DoD Digital Engineering Strategy (2018 [1]), DoD T&E leadership has embraced the pursuit in DoDI 5000.89 (2020 [12]) and is seeking how to gain the best value for achieving T&E transformation in concert with the larger DE transformation. The pathway toward implementation is not well defined. The DOT&E Strategy Update (2022 [2]) publication offers a framework, but further deliberation and planning by the T&E community is needed to deliver actionable recommendations for implementation over a prescribed timeline.

To this end, DOT&E, in partnership with the Acquisition Innovation Research Center (AIRC), is developing implementation recommendations relating to the following T&E transformation elements of the DOT&E Strategy Update (2022 [2]):

- Conducting operationally relevant assessments earlier in the development lifecycle (shift left).
- Developing and implementing credible digital tools and automation to propel better T&E, shorten T&E timelines, and illuminate performance shortfalls earlier.
- Integrating physics-based modeling and simulation, analytical models (e.g., discrete event simulations), and test data from multiple test phases to enhance understanding of technical risk and uncertainty projections.

These three elements are identified as priority based on their interdependencies and contributions for integrating ME, SE, and T&E. Collectively they represent the maturing of model-based systems engineering (MBSE) tools, methods, and data throughout the T&E acquisition lifecycle. Though this document is focused on MBSE, we recognize the value of non-MBSE digital engineering solutions such as relational databases. For example, both DOT&E and the Department of the Navy, are developing digital solutions for TEMPs and T&E planning, execution, and analysis that use relational databases.

### Current State of MBSE for T&E

AIRC has conducted literature research and engaged in numerous interviews with practitioners to characterize the current state of “MBSE for T&E”, which represents the integration of MBSE methods and tools with T&E activities, in alignment with the overall DE transformation.

### Defining “MBSE for T&E”

The Undersecretary of Defense, Research and Engineering defines MBSE as “[a]n integrated digital approach that uses authoritative sources of system data and models as an acquisition life cycle across disciplines to support life cycle activities from concept through disposal [1].” Logically following MBSE for T&E is simply the use of this integrated digital approach to support test planning, test execution, and the analysis of test results, to include integrating information from model-based data sources as appropriate. The DoD is pursuing a DE transformation that includes digital design and tooling interoperability across all aspects of the development process. MBSE is the SE portion of this transformation. As SE explores both the system level development as well as the system level Verification and Validation (V&V), the DoD
is also seeking to integrate the T&E digital processes and tools in the overarching DE transformation. Logically following the definition of MBSE, MBSE for T&E is the use of the integrated digital approach to support test planning, test execution, and the analysis of test results, to include integrating information from model-based data sources into analysis and evaluation as appropriate.

Fundamentally, MBSE enables systems engineers and mission engineers to perform the same functions including mission and system requirement definitions, requirements allocation to subsystems, identification of verification approaches to requirements, and tracing verification efforts through the various tiers of a system design as they are realized. However, MBSE enables the systems and mission engineering workforce to perform these tasks in a digital environment, which aids in dealing with the complexity of today’s systems. Use of digital environments support identification of interoperability requirements between system components, analysis of designs for conflicting requirements or technical gaps, and maintaining configuration management of the various components as a system design matures. MBSE applies digital design toolsets to define system and subsystem relationships and behaviors. The toolsets support maintaining linkage from mission objectives down into physical and software design environments. This linkage is maintained through the underlying data. MBSE enables the test community to plan for integration and test activities very early in the lifecycle. Early involvement of the T&E community is enabled by the use of toolsets automatically capturing the impact of changes in system design, supply chain, and integration steps on test planning and resources, while enabling the adoption of quantitative planning methods that directly connect to the integration and test planning models [13][14].

Large acquisition programs like the Ground Based Strategic Deterrent, the T-7 Jet Trainer, and the upcoming Future Long Range Assault Aircraft are all leveraging MBSE methods.

Current Practices in “MBSE for T&E”

From October-2022 through March 2023, AIRC conducted research to identify where “MBSE for T&E” is currently being practiced in government, industry, and academia; what approaches are being taken (to include tools and methods); what benefits are being realized; and what challenges and/or limitations are being encountered. This research included literature review and stakeholder interviews, and captured the perspectives of government, industry, and academic MBSE practitioners.

Literature review: Twenty-three publications were selected from the literature search, including government policy, academic research, and government digital engineering proposals for implementing T&E on DoD development programs. These publications were reviewed to better understand T&E policy, understand where the government envisioned the community to progress, and understand what MBSE for T&E-related practices have been explored either in academic research or industry.

Stakeholder interviews: Twelve expert MBSE practitioners were interviewed with the same intent. The research team sought to uncover where practitioners had applied digital T&E practices in concert with MBSE focus development programs. The interviews also explored current tooling and revealed some challenges that the practitioners experienced in their use of digital T&E methods.

Our overall observation is that government and industry are widely adopting and implementing MBSE tools and methods to support early system design and development. On the contrary, the integration of MBSE with T&E has been almost anecdotal and limited to a few individuals within large organizations.
In an effort to bring the ME, SE and T&E communities closer together in their adoption and implementation, we have identified three enablers for “MBSE for T&E”

1. Culture / Workforce
2. Methods / Process
3. Interoperability

Figure 1 below describes the current challenges associated with each of these enablers, along with recommendations to further the advancement of “MBSE for T&E” practice.

**Figure 1: Model Based T&E Challenges and Recommendations**

**Proposed “MBSE for T&E” Roadmap**

Building on our findings in the previous section, we have developed a “MBSE for T&E” Roadmap that outlines a logical progression for moving from a paper-centric, stove-piped acquisition to the envisioned data-centric T&E acquisition lifecycle within the DE transformation. This roadmap is comprised of three phases:

**Phase 1, Model-Based T&E Planning and Control.** This phase involves replacing traditional paper-based T&E artifacts with digital artifacts that include descriptive and executitional models. This phase does not simply consist of digitizing paper-based artifacts. Instead, it includes (1) establishing dedicated formal dialects, statistical models, and modeling languages, that are consistent but not yet integrated with mission and system development formal dialects, statistical models, and modeling languages, (2) deploying shared digital repositories, and (3) updating configuration control, work processes, and workforce culture to adopt modeling strategies towards T&E planning and control.
• **Benefits:** Streamlined workflows, improved access and sharing of information, increased T&E information consistency, more efficient and effective change management, and enhanced knowledge management.

• **Impact:** Enables better traceability, visibility, and decision-making across mission outcomes, operational capabilities, system requirements, and test plans.

**Phase 2, Dynamic T&E Planning and Execution.** The modeling framework established in Phase 1 is augmented in this phase with quantitative methods that enable dynamically updating a test plan as results become available during its execution. Here, multiple test phases are not only linked descriptively or quantitatively in terms of resource consumption and schedule and cost planning, but the confidence they generate in the accomplishment of mission objectives is linked in real-time using Bayesian and other statistical modeling and analysis methodologies (e.g., [6-9]). Evaluation of test results evolves from checking passing criteria to reassess confidence and reevaluate the need of future T&E activities.

• **Benefits:** Higher efficiency of test planning and management due to improved use of current observations and past experience (only activities that are truly necessary are executed), increased accuracy of test planning due to earlier identification of potential technical risks and gaps.

• **Impact:** Optimizes the use of available T&E data and information to support decision making under various conditions (i.e., under certainty, at risk, or under uncertainty) at major acquisition review points or milestones.

**Phase 3, Coupled Mission, System, and T&E Decision Making.** The modeling framework established in Phase 2 is integrated with mission and systems engineering modeling frameworks, both of descriptive models and of executional models. It includes (1) integrating formal dialects, statistical models, and modeling languages across domains, (2) integrating shared digital repositories, (3) develop mathematical underpinnings and quantitative methods to integrate data across domains and efficiently explore the integrated ME/SE/T&E trade-space, and (4) updating ME, SE, and T&E work processes and workforce cultures.

• **Benefits:** Ability to perform total system trade-off’s (T&E drives ME and SE and vice versa).

• **Impact:** Broken silos throughout ME, SE, and T&E acquisition lifecycle; delivers insights on 2nd order relationships across missions, systems, and their T&E.

The “MBSE for T&E” Roadmap should not be viewed as a one-size-fits-all implementation approach. Nor should it be viewed as purely sequential in its execution, where one phase must be completed in its entirety before moving to the next. Instead, this roadmap provides a guide for ongoing coordination on what additional research and development is needed to support each phase, and how practitioners should anticipate changes to T&E practices moving forward. In this manner, the roadmap facilitates common understanding and agreement on how to progress from today’s current state of practice toward the future desired end state of “MBSE for T&E”. The pathway for progressing through the three phases will need to be tailored for each DoD acquisition program, taking into account factors that include: complexity of the system and its operating environment(s), program schedule, program resources, maturity and availability of a digital backbone, and workforce competency.
Figure 2 provides a notional example of progression through the “MBSE for T&E” Roadmap, identifying the methods that may be introduced during each phase.

![Figure 2: Model Based T&E Roadmap](image)

**Conclusions & Future Work**

Based on our research to date on “MBSE for T&E” current state of practice, we recommend that Phase 1 of the “MBSE for T&E” Roadmap become the new baseline for all acquisition programs with respect to the adoption and implementation of MBSE and DE tools and methods. This baseline represents minimal requirements on the use of digitized data and information, in combination with basic digital tools and models, under the DE transformation. It offers a low barrier of entry for programs to begin their progression through the “MBSE for T&E” phases, and helps further the integration of ME, SE, and T&E disciplines across the acquisition lifecycle.

Additional research is needed to further the “MBSE for T&E” Roadmap and characterize each phase. Our steps toward developing further recommendations include:

1. Continuing to assess the current state of practice for “MBSE for T&E,” via test and evaluation master plan (TEMP) or other artifacts.
2. Advancing the above roadmap to describe the different levels of application within each phase.
3. Contextualize the accomplishment of “MBSE for T&E” across the entire acquisition lifecycle, to greater alignment with early ME activities.
4. Explore an exemplar case study to share with the community of practice to aid in adoption of these methods.

**Reference**

*The ITEA Journal of Test and Evaluation, (ISSN 1054-0929), is published four times each year by the International Test and Evaluation Association, 11350 Random Hills Rd, Suite 800, Fairfax, VA 22030 ([www.itea.org](http://www.itea.org)).*  
Copyright 2023, International Test and Evaluation Association, All Rights Reserved. Reproduction in whole or in part is prohibited except by permission of the publisher.
Author Biographies

Dr. Laura Freeman is a Research Professor of Statistics, the Deputy Director of the Virginia Tech National Security Institute and Assistant Dean for Research for the College of Science at Virginia Tech. Her research leverages experimental methods for conducting research that brings together cyber-physical systems, data science, artificial intelligence, and machine learning to address critical challenges in
national security. Dr. Freeman has a B.S. in Aerospace Engineering, a M.S. in Statistics and a Ph.D. in Statistics, all from Virginia Tech.

Dr. Peter A. Beling is a professor in the Grado Department of Industrial and Systems Engineering at Virginia Tech and is director of the Intelligent Systems Division of the Virginia Tech National Security Institute. He has contributed extensively to the development of methodologies and tools for the design and acquisition of cyber military systems. His research interests include digital engineering and topics at the intersection of artificial intelligence and systems engineering. Dr. Beling serves on the Research Council of the Systems Engineering Research Center (SERC) University Affiliated Research Center (UARC).

Dr. Kelli Esser is the Associate Director for the Intelligent Systems Division of the Virginia Tech National Security Institute. She is a recognized leader and change agent with over eighteen years of experience working with national and homeland security agencies to improve the planning, management, and coordination of research development and innovation investments across public and private partners. Dr. Esser’s research has contributed to major policy and implementation changes across government in the areas of test and evaluation, capabilities-based planning, and operational assessments in support of technology development, transition, and deployment.

Dr. Paul Wach is a Research Assistant Faculty with the Intelligent Systems Division of the Virginia Tech National Security Institute. His research interests include the intersection of theoretical foundations of systems engineering, digital transformation, and artificial intelligence. Dr. Wach is associated with The Aerospace Corporation, serving as a subject matter expert on digital transformation. He has prior work experience with the Department of Energy in lead engineering and management roles ranging in magnitude from $1-12B as well as work experience with two National Laboratories and the medical industry. Dr. Wach received a B.S. in Biomedical Engineering from Georgia Tech, M.S. in Mechanical Engineering from the University of South Carolina, and Ph.D. in Systems Engineering from Virginia Tech.

Geoffrey Kerr is a Senior Research Associate at the Virginia Tech National Security Institute with expertise in Systems and Digital Engineering. Mr. Kerr is a 34-year veteran of the Aerospace and Defense industry where he has vast experience in Systems Engineering, Program and Project Management, and diverse technical leadership in developing aircraft for the US Government and allied nation states. Mr. Kerr held a prominent executive role implementing Digital Transformation across the Lockheed Martin Corporation prior to joining VTNSI.

Dr. Alejandro Salado is an associate professor of systems engineering with the Department of Systems and Industrial Engineering at the University of Arizona. In addition, he provides part-time consulting in areas related to enterprise transformation, cultural change of technical teams, systems engineering, and engineering strategy. Alejandro conducts research in problem formulation, design of verification and validation strategies, model-based systems engineering, and engineering education. Before joining academia, he held positions as systems engineer, chief architect, and chief systems engineer in manned and unmanned space systems of up to $1B in development cost. Dr. Salado holds a BS/MS in electrical and computer engineering from the Polytechnic University of Valencia, a MS in project management and a MS in electronics engineering from the Polytechnic University of Catalonia, the SpaceTech MEng in space systems engineering from the Technical University of Delft, and a PhD in systems engineering from the Stevens Institute of Technology.
Dr. Jeremy Werner, ST was appointed DOT&E’s Chief Scientist in December 2021 after starting at DOT&E as an Action Officer for Naval Warfare in August 2021. Before then, Jeremy was at Johns Hopkins University Applied Physics Laboratory (JHU/APL), where he founded a data science-oriented military operations research team that transformed the analytics of an ongoing military mission. Jeremy previously served as a Research Staff Member at the Institute for Defense Analyses where he supported DOT&E in the rigorous assessment of a variety of systems/platforms. Jeremy earned a PhD in physics from Princeton University where he was an integral contributor to the Compact Muon Solenoid collaboration in the experimental discovery of the Higgs boson at the Large Hadron Collider at CERN, the European Organization for Nuclear Research in Geneva, Switzerland. Jeremy is a native Californian and received a bachelor’s degree in physics from the University of California, Los Angeles where he was the recipient of the E. Lee Kinsey Prize (most outstanding graduating senior in physics).

Dr. Sandra Hobson was appointed to the Senior Executive Service in 2014 as a Deputy Director in the Office of the Director, Operational Test and Evaluation (DOT&E), within the Office of the Secretary of Defense. She is currently responsible for defining and executing strategic initiatives, and supporting the development of policy and guidance to meet the test and evaluation demands of the future as the complexity of Department of Defense weapons systems and multi-domain operational environments evolve. Dr. Hobson was selected to perform the duties of the DOT&E Principal Deputy Director from January 20 to December 19, 2021.

Prior to this appointment, Dr. Hobson supported DOT&E in two other capacities: first as a Research Staff Member at the Institute for Defense Analyses and then as an Aircraft Systems and Weapons Staff Specialist. During these tenures, she provided technical oversight of test and evaluation programs to enable adequate assessments of the survivability and lethality of a subset of Department of Defense aircraft and weapons acquisition programs.

Dr. Hobson earned her Bachelor of Science degree in Aerospace Engineering from the United States Naval Academy and her Doctor of Philosophy degree in Aerospace Engineering from A. James Clark School of Engineering at University of Maryland. Dr. Hobson is a recipient of the National Fellowship for Exceptional Researcher awarded by the United Nations Educational, Scientific and Cultural Organization (UNESCO), and two Secretary of Defense Medals for Meritorious Civilian Service. She currently resides in Virginia with her husband and two German Shepherd dogs.